Iowa State University
Digital Repository

Graduate Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

2013

Using laboratory experiments to study otherwise unobservable labor market interactions

Fanzheng Yang
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Labor Economics Commons

Recommended Citation
Yang, Fanzheng, "Using laboratory experiments to study otherwise unobservable labor market interactions" (2013). *Graduate Theses and Dissertations*. 13093.
https://lib.dr.iastate.edu/etd/13093

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Using laboratory experiments to study otherwise unobservable labor market interactions

by

Fanzheng Yang

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:

Tanya S. Rosenblat, Major Professor

Joseph A. Herriges

Wallace Huffman

Markus M. Mobius

John R. Schroeter

Iowa State University

Ames, Iowa

2013

Copyright © Fanzheng Yang, 2013. All rights reserved.

DEDICATION

I would like to dedicate this thesis to my parents without whose support I would not have been able to complete this work. I would also like to thank my friends and family for their loving guidance and financial assistance during the writing of this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis. First and foremost, Dr. Tanya S. Rosenblat for her guidance, patience and support throughout this research and the writing of this thesis. Her insights and words of encouragement have often inspired me and renewed my hopes for completing my graduate education. I would also like to thank my committee members for their efforts and contributions to this work: Dr. Joseph A. Herriges, Dr. Wallace Huffman, Dr. Markus M. Mobius and Dr. John R. Schroeter.

ABSTRACT

In this dissertation, we use laboratory experiments to study otherwise unobservable interactions in the labor market. The key advantage of laboratory experiments is the ability to control conditions more tightly than in any other context, so it plays a distinctive role in serving as the first link in a longer chain running from standard theory to actual outcome in the real world. Our experiments can even provide valuable information on how people behave in situations where existing theory provides little or no guide to what should happen.

In our first experiment, we study the impact of specific incentive schemes on people's behavior by systematically varying them. We create an experimental labor market where "workers" can join "companies" that pay them according to different compensation schemes: (a) piece rate, (b) revenue sharing, (c) individual tournament, and (d) team tournament. In order to disentangle incentive and self-sorting effects, our experiment forces all workers to initially complete real-effort tasks paid by four given incentives respectively, and then use compensating differentials to elicit their preferences for different incentives. Therefore, based on the lab data from our sample of Chinese university students, we are able to study their productivity response to various incentives as well as their preferences for different types of compensation.

When we analyze individual productivity under four incentives, we find that: (1) compared to the baseline performance paid by piece rate, under three team-based incentives, more competitive incentive generates higher performance improvement. (2) Feedback about relative performance reduces the performance differences between three team-based incentives. (3) Regardless of incentives and feedback information, an additional compensation in terms of sign-up bonus brings a positive and significant effect on individual performance. In addition, by eliciting subjects' preferences for different compensation schemes, we build a mapping from individual characteristics to their self-sorting outcome as follows: (1) Subjects with high relative performance always prefer individual tournament to other two team-based incentives. (2)

Risk-averse subjects are less likely to choose individual tournament if knowing the information about their relative performance. (3) Cooperative incentives attract more women than men, which is partially explained by gender-specific social preferences. (4) Compared to children with siblings, only children are less cooperative but more competitive. (5) In the absence of feedback, overconfident subjects are more likely to enter into individual tournament than those under-confident subjects with the same ability. Interestingly, the provision of information about their relative performance eliminates the impact of biased self-assessment. As a result, the feedback helps reduce the gender gap in competition as well as the difference between only child and child with siblings.

In a different study, we design a new laboratory experiment to investigate the ways that trust between strangers evolves in a setting where noisy feedback regarding mutual trustworthiness is present. We use a two-player sequential trust game where each trustor receives a sequence of noisy binary signals that reveal the trustworthiness type of the trustee. As a result, we track the evolution of trustors' individual beliefs about trustworthiness types of trustees to document that subjects process information in an asymmetric way compared to a perfect Bayesian: they react more to negative feedback rather than positive. We show that our empirical results arise naturally in a theoretical model where there exists a complementary relationship between initial trust and optimally biased Bayesian information processing. Hence, we theoretically predict that greater initial trust must be counter-balanced by more asymmetric belief updating. We then use a novel method to demonstrate this hypothesis in the following-up experiment. We match participants from two different universities (in Hong Kong and Beijing, respectively) and prime them on the social identity of their counterparts. Consequently, by the introduction of social identity, we find that both initial trust level and asymmetry of belief updating are stronger for in-group matches than out-group matches, which is consistent with our theoretical prediction.

CHAPTER 1. GENERAL INTRODUCTION

This dissertation is organized into six chapters. The current chapter presents a general introduction to the chapters that follow, and provides an outline for the organization of the dissertation. While the general theme of this dissertation is using laboratory experiments to study otherwise unobservable labor market interactions, each chapter is meant to stand alone by addressing a specific issue. We begin in the next chapter by using a laboratory experiment to study the interaction between incentive schemes and work organization in the modern workplace. Three aspects of the important findings from this labor market experiment are presented in Chapter 2, 3 and 4, respectively. And then, Chapter 5 discusses a new laboratory experiment, which is designed to investigate the ways that trust between strangers evolves in a setting where noisy feedback regarding mutual trustworthiness is present. Finally in Chapter 6, we provide a summary and several general conclusions.

Recent labor market research suggests that workers of similar ability might behave quite differently under different incentives. For example, some workers might reduce effort when working in a team because of the "free-riding" opportunity. Some might dislike competition and relative performance pay because they find it stressful. Although many authors (Niederle and Vesterlund 2007; Wozniak et al. 2010) have stressed the importance of non-cognitive characteristics such as risk preference and self-confidence on labor market outcomes, only very few studies (e.g. Dohmen and Falk 2011) use incentivized elicitation methods to measure those characteristics. In our first labor market experiment, we use a rich battery of diagnostic games to elicit individual characteristics including risk attitude, social preference and self-confidence. Hence, Chapter 2 provides the description of how we use experimental methods to elicit those characteristics, and then summarizes the important findings from our subject pool of 411 Chinese university students. First of all, it is shown that there are significant gender differences on those non-cognitive characteristics: On average, men are less risk averse and more over-confident than women; regarding the social preference, men are more likely to be either perfectly selfish or perfectly selfless, whereas women tend to be equalitarian who prefer to share evenly[1]. Moreover, thanks to the uniqueness of our subject pool, we obtain intriguing findings about the behavior of only children participants in our sample in China. For instance, only children are less risk-averse and also less over-confident than the children with siblings; however, there is no significant difference in social preference between only children and children with siblings, which provides experimental evidence against the stereotype that only children are more likely to be selfish.

Next, in order to investigate individual response to different incentives, we create an experimental labor market where "workers" can join "companies" that pay them according to different compensation schemes. Since the modern workplace puts an ever greater emphasis on teamwork, we systematically study the impact of different team-based incentives. In addition to the baseline incentive, i.e. "piece rate" compensation scheme for individual-based working environment, the following three representative team-based incentives are designed to highlight the tradeoff between cooperation and competition: (1) "revenue sharing" where all earnings are equally shared among team members; (2) "individual tournament" where only the highest performer in the team can earn a prize; and (3) "team tournament" where teams should compete first, and then prizes are shared within the winning team. In particular, the last team-based incentive, i.e. "team tournament", is defined as a hybrid of "individual tournament" and "revenue sharing", which are two contrasting "corporate cultures" one that focuses on competition and one that emphasizes collaboration.

Moreover, self-sorting might amplify the effect of an incentive scheme. In the presence of worker self-sorting, it is possible that workers with different individual characteristics feel attracted by different pay schemes and then systematically self-select into particular firms and organizations (Dohmen and Falk 2011). Therefore, in order to disentangle incentive and sorting effects, our experiment forces all workers to initially complete four real-effort tasks paid

[1] Our finding is consistent with the conclusion of Andreoni and Vesterlund (2001), in which the authors conduct a dictator game with varying income and prices. The sample of that experimental study consists of 47 female subjects and 95 male subjects from University of Wisconsin and Iowa State University in United States.

by piece rate, revenue sharing, individual tournament, and team tournament, respectively, and then elicits their preferences for different incentives through a self-selection process. Using our experimental data, we are able to study individual productivity response to various incentives as well as their preferences for different types of compensation.

In Chapter 3, we aim to explore the most efficient way(s) to improve workers' productivity by considering three potential influence factors at the same time, that is: team-based incentives, feedback about relative performance, and compensating differential in terms of extra sign-up bonus. As a result, we find that: (1) compared to the baseline performance paid by piece rate, under three team-based incentives, more competitive incentive generates higher performance improvement. (2) Feedback about relative performance reduces the performance differences between three team-based incentives. (3) Regardless of incentives and feedback, the extra sign-up bonus brings a positive and significant effect on individual performance. Hence, we suggest that compared with designing a distinctive incentive scheme and feedback treatment for each heterogeneous worker, it should be much more feasible and convenient for companies to set a flexible sign-up bonus policy while holding other elements of workplace environment fixed.

Chapter 4 studies sorting into four given incentives based on individual characteristics. An innovation in the design for this self-sorting task is that we ask subjects to not only rank those four payment schemes from most preferred to least preferred, but we also elicit their compensating differentials between any two consecutive incentives. More importantly, instead of asking several hypothetical evaluation questions in an open-ended format, the elicitation of compensating differential in our experiment adopts a variant of the Becker-DeGroot-Marschak (BDM) mechanism that is incentive-compatible. To the best of our knowledge, this experiment is the first one to implement such a mechanism in a laboratory experiment to price the cost of a mismatch between a worker and a company. As a result, the lab data from our sample of Chinese university students clearly show the relationship between their stated-preferences and individual characteristics, and the main findings include (1) Subjects with high relative performance always prefer individual tournament to other two team-based incentives; (2) Risk-averse subjects are less likely to choose individual tournament if knowing the information about their relative performance; (3) Cooperative incentives attract more women than men, which is partially explained by gender-specific social preferences; (4) Compared to children with siblings, only children are less cooperative but more competitive; (5) In the absence of feedback, overconfident subjects are more likely to enter into individual tournament than those under-confident subjects with the same ability. Interestingly, the provision of information about their relative performance eliminates the impact of biased self-assessment. As a result, the feedback helps reduce the gender gap in competition as well as the difference between only child and child with siblings[2].

Regarding the outcome in our incentive experiment, it's noticeable that majority of subjects are often reluctant to choose cooperative incentives. A possible reason for this is the lack of trust that other team members will exert maximum effort. Another reason is the lack of information about the type of partner. Since this lack of trust is pervasive in many labor market interactions, it is worthwhile to explore the evolution of trust in a more general setting. Hence, in Chapter 5, we study how trust between strangers evolves in a setting with noisy feedback about the trustworthiness of others. In our second laboratory experiment, the design uses a two-player sequential trust game where each trustor receives a sequence of noisy binary signals that reveal the trustworthiness type of the trustee. As a result, we track the evolution of trustors' individual beliefs about trustworthiness types of trustees to document that subjects process information in an asymmetric way compared to a perfect Bayesian: they react more to negative feedback rather than positive. Trust is therefore much easier to lose than to gain. Next, we show that our empirical results arise naturally in a theoretical model where individuals optimally manage their trust in others with noisy feedback. We discuss whether trustors should use Bayesian updating to calculate their trust levels or whether they should process information in a biased manner. The conclusion suggests that there exists a complementary relationship between initial trust and optimally biased Bayesian information processing. Hence, we theoretically predict that greater initial trust must be counter-balanced by more asymmetric belief updating. We then use a novel method to demonstrate this hypothesis in the following-up experiment. We match participants from two different universities (in Hong Kong and Beijing, respectively)

[2] This result is similar to the findings in the experimental study of Wozniak et al. (2010), where feedback information about relative performance helps all subjects move towards more optimal choices and then the gender gap in competitiveness is reduced.

and prime them on the social identity of their counterparts. Consequently, by the introduction of social identity, we find that both initial trust level and asymmetry of belief updating are stronger for in-group matches than out-group matches, which is consistent with our theoretical prediction.

Finally, in Chapter 6, a summary and some general conclusions are presented. In order to explore the interaction between incentive schemes and work organization, our first experiment applies simple but efficient experimental methods to elicit individual response to different team-based incentives. To the best of our knowledge, this design is the first one to implement the "compensating differential" in a laboratory experiment to price the cost of a mismatch between a worker and a company. Moreover, thanks to the uniqueness of our subject pool in China, we obtain intriguing findings about the behavior of only children participants in our labor market experiment. In our second study, we investigate the evolution of trust between strangers by both experimental method and theoretical model. As a consequence, we reject the behavioral predictions based on standard Bayesian rule and verify that individuals follow an optimally biased information processing rule to update their trust beliefs. Besides, this study also contributes to social identity research by highlighting the impact of social identity on individual initial trust as well as trust updating.

CHAPTER 2. EXPERIMENTAL METHODS FOR ELICITATION OF INDIVIDUAL CHARACTERISTICS

Abstract

Although many labor economists have stressed the importance of non-cognitive characteristics on labor market outcomes, only very few studies use incentivized elicitation methods to measure them. Hence, in this laboratory experiment, we use a rich battery of diagnostic games to elicit several non-cognitive characteristics including risk attitude, social preference and self-confidence. From our experimental data, we can verify that there are remarkable gender gaps on those non-cognitive characteristics: on average, men are less risk averse and more over-confident than women; regarding the social preference, men are more likely to be either perfectly selfish or perfectly selfless, whereas women tend to be equalitarian who prefer to share evenly. Moreover, thanks to the uniqueness of our subject pool in China, we also find some differences between only children and children with siblings: only children are less risk-averse and also less over-confident than the children with siblings; however, there is no significant difference in social preference between two groups, which provides experimental evidence against the stereotype that only children are more likely to be selfish.

2.1 Introduction

Over the last two decades, laboratory experiments have become a popular method of research to examine individual behavior and social preferences in labor markets (see a literature review in Charness and Kuhn 2011). The key advantage of laboratory experiments is researchers can control the environment under which individuals make their decisions and allow causal inferences by exogenously varying one parameter while holding all others constant.

Another major advantage of laboratory experiments derives from the convenience of eliciting endogenous determinants of individual differences which are unobservable in the real world. For instance, although many labor economists have stressed the importance of non-cognitive characteristics such as risk attitude, social preference and self-confidence on labor market outcomes, it is very difficult to measure them by field data. On the contrary, by using a rich battery of diagnostic games to elicit those individual characteristics in a laboratory experiment, our study obtains intriguing findings about the subject pool of 411 Chinese university students. We verify that there are remarkable gender gaps on individual characteristics: on average, men are less risk averse and more over-confident than women; regarding the social preference, men are more likely to be either perfectly selfish or perfectly selfless, whereas women tend to be equalitarian who prefer to share evenly[1]. Moreover, thanks to the uniqueness of our sample in China, we also find some differences between only children and children with siblings: only children are less risk-averse and also less over-confident than the children with siblings; however, there is no significant difference in social preference between two groups, which provides experimental evidence against the stereotype that only children are more likely to be selfish.

Finally, the experimental data elicited by incentivized methods should be more convincing than the self-reported data from traditional surveys. Different from the way of only asking several hypothetical questions, laboratory experiments always provide participants with real monetary rewards based on their response to decision-making questions. Therefore, compared with answering survey, people are more likely to take the questions seriously and tell truth in the laboratory experiments. Now, there is a new research strand combines incentive-compatible experimental measures with survey methods (Fehr et al. 2003; Ermisch et al. 2009; Dohmen et al. 2010). For example, the British Household Panel Study (BHPS), the new household panel Understanding Society in the United Kingdom, and the German Socio-Economic Panel (SOEP), both add experimental sessions into a traditional questionnaire to collect more specific individual characteristics in addition to the basic individuals' socio-demographic characteristics used in representative surveys.

[1]Our finding is consistent with the conclusion of Andreoni and Vesterlund (2001), in which the authors conduct a dictator game with varying income and prices. The sample of that experimental study consists of 47 female subjects and 95 male subjects from University of Wisconsin and Iowa State University in United States.

In order to systematically investigate the impacts of individual characteristics on workers' response to different team-based incentives in our experimental labor market, before the workers enter into the hiring process, we first arrange a diagnostic stage to elicit their individual characteristics including: risk attitude, social preference, ability for a specific "real effort" task and their self-assessment about the relative performance compared with other team members.

The rest of this chapter is organized as follows. Section 2 provides the description of how we use experimental methods to elicit four individual characteristics one by one. And then, we summarize the important findings in section 3. Finally, in section 4, some concluding remarks are presented.

2.2 Experimental Design

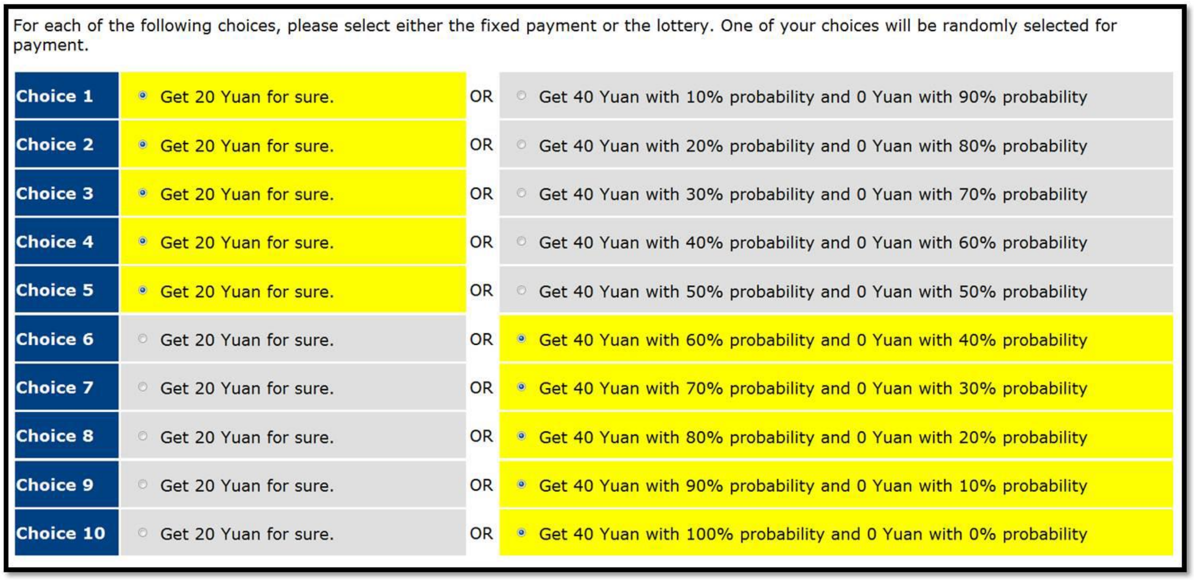
2.2.1 Four Tasks to Elicit Individual Characteristics

2.2.1.1 Task 1: Elicit Risk Preference

This task uses lottery choices to elicit people's risk preference. In fact, there is an influential and large body of research in experimental economics measures risk aversion (e.g. Tversky and Kahneman 1992; Barsky et al. 1997; Andersen et al. 2008). In many experiments in this literature, economists infer risk preferences from lists of binary lottery choices. Hence, in our experimental design, we elicited subjects' risk attitudes by this commonly used method. In Task 1, each subject was asked to answer 10 binary choices questions as shown in Figure 2.1.

In each binary choice, subjects need to decide whether they prefer a safe option or playing a lottery. The safe option is always the same in each question, which is 20 Chinese Yuan as the fixed payoff. However, the lotteries vary and the risk of lottery option decreases from question to question. In other words, the probability of winning 40 Chinese Yuan increases from 10% to 100% in the end.

Now suppose a subject has consistent risk preference on lotteries, s/he will prefer the safe option to the lotteries above a certain risk level, and then switch to preferring the lottery option in all subsequent choice questions. Therefore, the switching point of winning 40 Chinese Yuan in a lottery indicates the subject's risk attitude. That is, the higher this switching point is, the



For each of the following choices, please select either the fixed payment or the lottery. One of your choices will be randomly selected for payment.

Choice 1	Get 20 Yuan for sure.	OR	Get 40 Yuan with 10% probability and 0 Yuan with 90% probability
Choice 2	Get 20 Yuan for sure.	OR	Get 40 Yuan with 20% probability and 0 Yuan with 80% probability
Choice 3	Get 20 Yuan for sure.	OR	Get 40 Yuan with 30% probability and 0 Yuan with 70% probability
Choice 4	Get 20 Yuan for sure.	OR	Get 40 Yuan with 40% probability and 0 Yuan with 60% probability
Choice 5	Get 20 Yuan for sure.	OR	Get 40 Yuan with 50% probability and 0 Yuan with 50% probability
Choice 6	Get 20 Yuan for sure.	OR	Get 40 Yuan with 60% probability and 0 Yuan with 40% probability
Choice 7	Get 20 Yuan for sure.	OR	Get 40 Yuan with 70% probability and 0 Yuan with 30% probability
Choice 8	Get 20 Yuan for sure.	OR	Get 40 Yuan with 80% probability and 0 Yuan with 20% probability
Choice 9	Get 20 Yuan for sure.	OR	Get 40 Yuan with 90% probability and 0 Yuan with 10% probability
Choice 10	Get 20 Yuan for sure.	OR	Get 40 Yuan with 100% probability and 0 Yuan with 0% probability

Figure 2.1 Binary choices between a fixed payoff and a lottery to elicit risk preference

more risk-averse s/he is.

More importantly, instead of using hypothetical gambles to measure risk aversion, our experiment used actual decision-making questions with real payoff. In other words, according to subjects' final decisions in 10 binary choices, we would randomly select one choice out of ten to calculate their payoffs in Task 1. For instance, if Choice 2 is randomly selected to calculate the payoff, then the subject would receive 20 Yuan if s/he selected the fixed payment; if s/he selected the lottery option of winning 40 Yuan with 20% probability, then her/his payoff would depend on the realized outcome of such a lottery, which could be either 40 Yuan or 0 Yuan in the end.

2.2.1.2 Task 2: Elicit Social Preference

In order to better understand the heterogeneity in individual preferences for equitable outcomes or social welfare, we elicited subjects' social preferences by a dictator game. In this dictator game, each participant was randomly matched with another participant in the lab room, playing the role of a dictator or a receiver. As a dictator, s/he could get 30 Chinese Yuan as endowment at the beginning of the game, and could either keep all money or send some amount to another person, i.e. the receiver. On the other hand, as a receiver, s/he had no endowment, but could receive something if the dictator passed money to her/him. Moreover, whatever amount the dictator decided to send, the money was doubled by the experimenter and then added to the receiver's payoff. In order to classify each subject, everyone had to play the role of dictator and decide how much sent to the receiver from an endowment of 30 Chinese Yuan. After all transfer choices made, subjects were randomly matched in pairs and two roles within each pair were assigned by a random draw to determine their final payments in this task.

From the individual choices in this dictator game, we are able to distinguish three types of social preferences. First, the subjects with "selfish" preference, who sent an amount of money less than 10 Chinese Yuan to the receiver so that s/he would have a payoff higher than the receiver; Second, the subjects with "inequality-averse" or "fair" preference, who sent 10 Chinese Yuan to the receiver to make an equal income distribution in the end; Third, the subjects with "altruistic" preference, who sent an amount more than 10 Yuan so that the dictator would earn less than the receiver. Actually, the dictator's behavior of sacrificing her/himself to make other people better off might because they have increasing utility in others' payoffs (Andreoni 1990; Andreoni and Miller 2002), or they want to maximize total social welfare (Charness and Rabin 2002; Engelmann and Strobel 2004).

2.2.1.3 Task 3: Elicit Individual Ability

We implemented a "real effort" work in our experiment to elicit individual ability for this specific puzzle-solving task, which means subjects really had to work and were to some extent uncertain about their own productivity as well as the productivity of others. This is a realistic feature of most work tasks and leaves room for eliciting self-assessment about their relative performance in the next diagnostic task. Hence, in Task 3, all subjects were required to solve letter-puzzles as many as possible in 3 minutes. As a task, solving letter-puzzles is well suited for our purposes because it requires no previous knowledge, is easy to explain, and guarantees a sufficient degree of heterogeneity in productivity. Figure 2.2 provides an example of letter-puzzle used in our experiment.

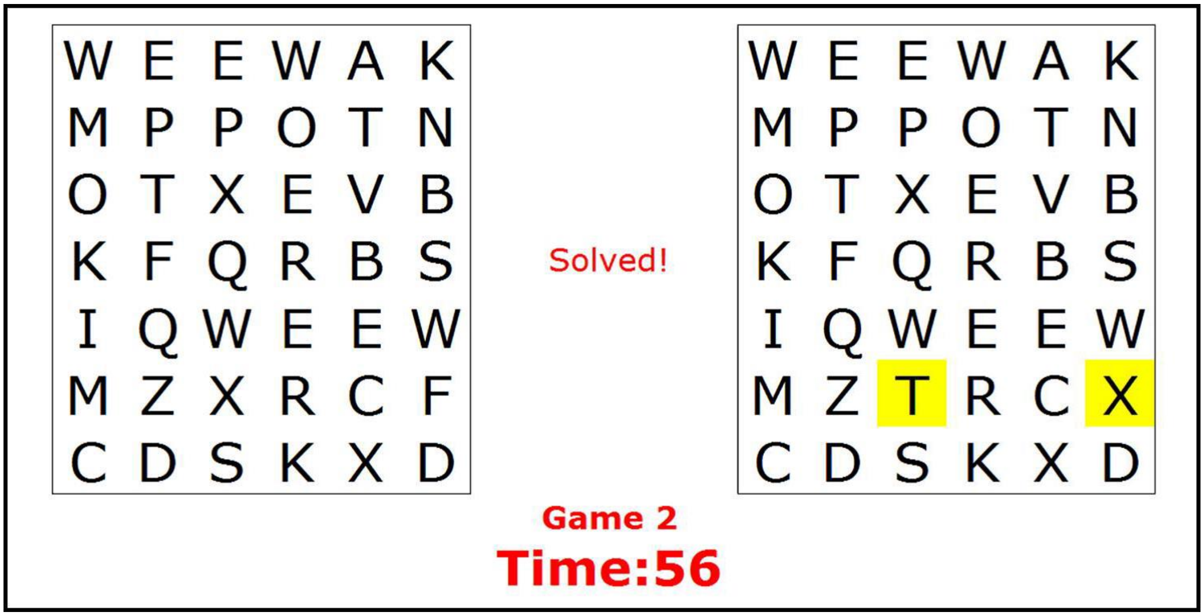


Figure 2.2 Puzzle-solving game to elicit individual ability for this task

The square with letters on the right differs from the square of letters on the left in two letters. Subjects had to find those letters and click on them to solve the puzzle. Whenever they clicked a letter other than those two letters, the remaining time was reduced by 1 second. Hence, people had to be more cautious and should try to only click letters that differ. Within a 3-minute period, each subject was asked to solve as many puzzles as possible, and s/he could increase her/his payoff by a high productivityBefore the task started formally, subjects had three practice examples to familiarize themselves with this letter-puzzle, in which no payment was involved.. A clock on the screen informed how many seconds they had used.

Once this task finished, all subjects were informed of their own performance, but nothing about the performance of others. The pay scheme in this task was “piece rate”, i.e. subjects would receive 1 Chinese Yuan for each solved puzzle. The main reason for choosing “piece rate” as the monetary incentive is because it is a compensation scheme based on individual absolute performance only. Therefore, by excluding the impacts from other people, the piece rate scheme is a good proxy for subjects’ real abilities irrespective of any team-based incentives. Besides, it is also an intrinsic motivation for everyone to exert effort in the task: the better someone performed, the more s/he would earn.

2.2.1.4 Task 4: Elicit Self-confidence

According to outcome in Task 3, this new task was designed to elicit subjects' self-assessments about their relative performance in the previous puzzle-solving task. That is, each subject was asked to estimate the probability that her/his performance was better than a randomly selected other player in the lab room. And, subjects could increase their earnings by submitting a good estimate. We used a novel crossover mechanism to elicit their self-confidence about the relative performance. In fact, Mobius et al. (2011) is the first paper to implement this crossover mechanism in an experiment. The crossover mechanism has one main advantage over the otherwise popular quadratic scoring rule that most experimental papers use, i.e. it works for a wide class of preferences: it is well known that the quadratic scoring rule only produces truthful reports if subjects are risk-neutral. If subjects are risk-averse, quadratic scoring induces biased reports. In contrast, the crossover mechanism only requires that subjects' preferences are monotonic in the following sense: among the class of lotteries that pay off $x > 0$ with probability q and 0 with probability $1 - q$, subjects strictly prefer those with higher q. Truthful reporting of subjective beliefs is then a strictly dominant strategy. Monotonicity holds for all von-Neumann-Morgenstern preferences as well as many non-standard preferences such as Prospect theory.

In the experimental instruction, we presented this mechanism in a simple narrative form. We told subjects that they were paired with a "helper robot" who would also take the puzzle-solving task and who had a certain fixed probability q of winning against a randomly selected player. This probability was between 0 and 100% but unknown to the subject. Subjects were told that they could have their winnings based on their robot's performance rather than their own, and asked to indicate a threshold level of q above which they preferred to use the robot's performance. As a consequence, when the helper robot's winning probability was higher than the subject's indicated threshold, the robot's performance would be used to compete with the subject's opponent; otherwise, subject's own performance would be used against her/his selected opponent. In the end, only if the subject's or the helper robot's performance was better than the performance of the other selected player, 10 Chinese Yuan could be added to the individual earnings. Therefore, subjects should maximize their probability of earning an extra 10 Chinese Yuan by choosing the threshold as their own subjective probability of being better than some other player.

2.2.2 Subject Pool

In 2011, we conducted this laboratory experiment at Central University of Finance and Economics in Beijing, China. We had 411 complete observations, including 184 males and 227 females, and all of them were university students. The summary statistics of our subject pool is presented in Table 2.1. There were 213 only children (and 198 children with siblings) in our subject pool. Because majority of the participants i.e. 70% were graduate students, the average college year was above 4 and the average age was 23.03 correspondingly. Only 2% of the participants were married. Finally, the average level of GPA was 3.30 (on a 4.0 scale) with a standard deviation of 0.35. It is worth noting that subjects' grades were self-reported, so we cannot guarantee the accuracy of this information and treat it as a quite limited index to personal academic performance.

Table 2.1 Summary statistics of subject pool in the labor market experiment

	Mean	Std. Dev.	Min	Max
Female	0.55	0.5	0	1
Only Child	0.52	0.5	0	1
College Year	4.62	1.39	1	9
Age	23.03	2.35	17	38
Married	0.02	0.15	0	1
GPA	3.3	0.35	2	3.95

Notes:
"Female" = 1 if the subject is a female, otherwise = 0;
"Only Child" = 1 if the subject is the only child in family, otherwise = 0;
"College Year" =1 for freshmen, = 2 for sophomores, = 3 for juniors, = 4 for seniors, = 5 for 1st year masters, = 6 for 2nd year masters, = 7 for 3rd year masters, = 8 for 1st year PhDs, = 9 for 2nd year PhDs;
"Married" =1 if the subject is married, otherwise = 0;
"GPA" is the academic score measured on a 4.0 scale.

2.3 Results

Since we designed four tasks to elicit four individual characteristics respectively, we plan to report our findings related to each task one by one.

2.3.1 Risk Preference in Task 1

For each subject having monotonous preference, given a binary choice between a safe option of fixed 20 Yuan and a lottery of winning 40 Yuan with probability $x\%$, the minimum level of the winning probability $x\%$ that makes the subject prefer lottery to safe option indicates her/his risk attitude. That is, the higher this threshold x is, the more risk-averse the subject is. Figure 2.3 shows the distribution of individual preference for risk within the overall population, then Figure 2.4 indicates the gender difference on risk preference, and finally Figure 2.5 makes the comparison between only children and children with siblings.

Moreover, we divide subjects with different risk attitudes into three types: risk-loving if subject has a threshold of winning probability lower than 50%; risk-neutral if the threshold is equal to 50%; and risk-averse if the threshold is higher than 50%. Accordingly, we summarize the proportion of those three types in Table 2.2.

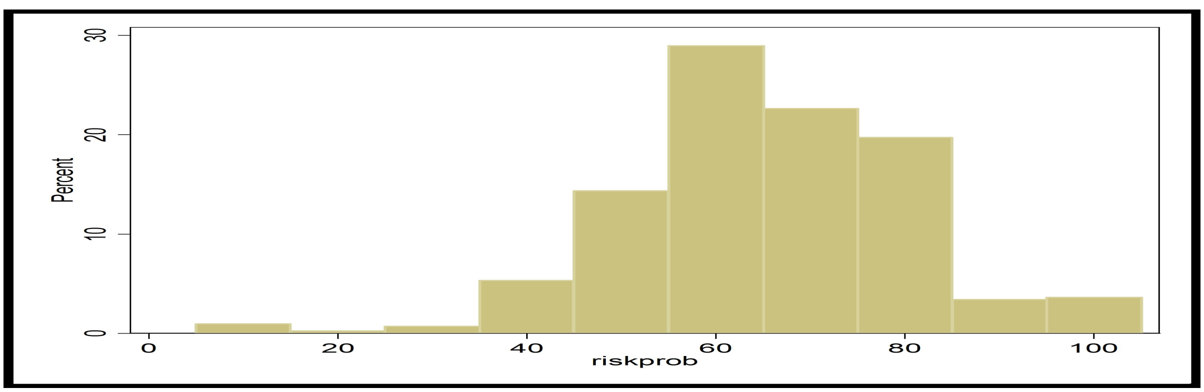


Figure 2.3 Distribution of risk preferences in the overall sample

From the results in three figures and Table 2.2, we find the majority of subjects are risk-averse. Meanwhile, consistent with the previous studies on gender difference (e.g. Dohmen and Falk 2011), male is less risk-averse than female since the difference in the mean of risk

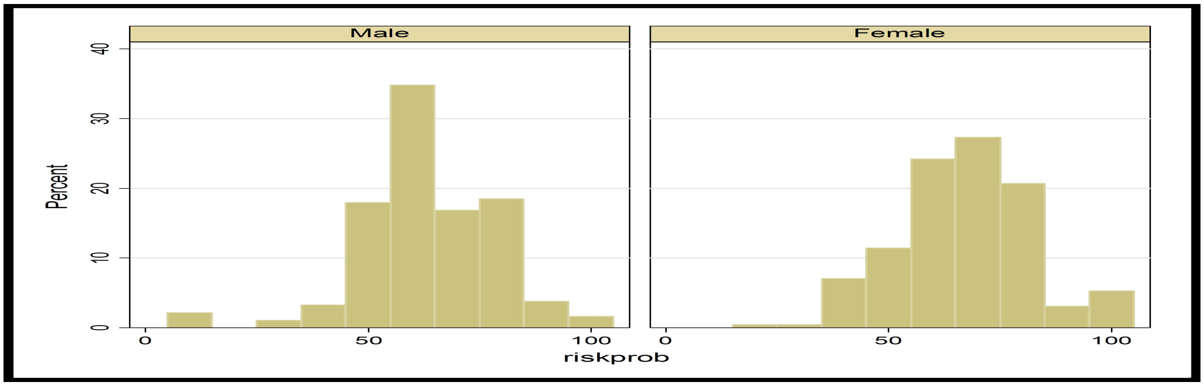


Figure 2.4 Distribution of risk preferences – male vs. female

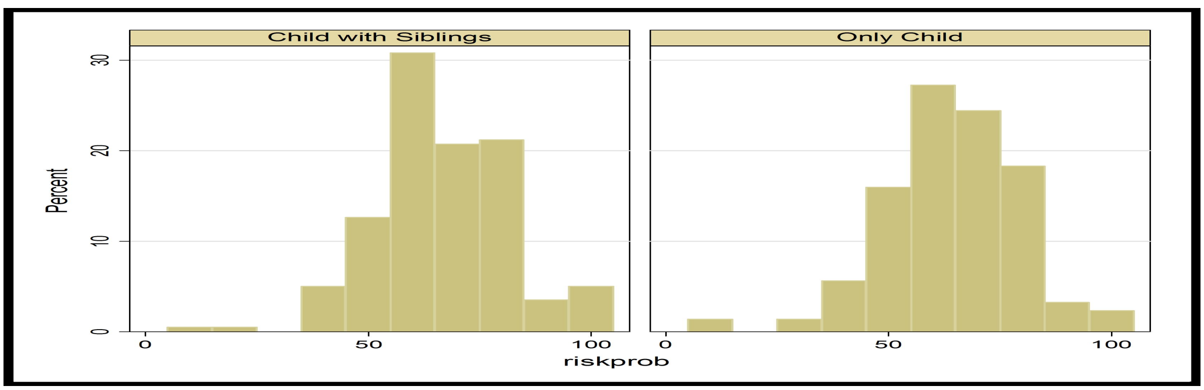


Figure 2.5 Distribution of risk preferences – child with siblings vs. only child

threshold between two groups is significant at $p = 0.013$. Similarly, by the comparison between only children and children with siblings, we also find a statistically significant difference (at $p = 0.099$) in their risk preferences, that is, only children are less risk-averse than the counterpart. It is worth noting that our finding about only child is just opposite to the conclusion in the latest paper on China's One-Child policy, i.e. Cameron et al. (2013), where the authors document that only child is more risk-averse than the child with siblings.

Furthermore, we suspect the difference of risk attitude between only child and child with siblings is embedded in One-Child policy, which is unique in China. There are some minor exceptions in China's One-Child policy. For example, the policy is stricter for urban residents than for rural residents. Parents in rural areas are allowed to have another kid if the first baby

Table 2.2 Summary statistics of risk attitude and proportions of three types

	Obs.	Mean	Std. Dev.	Min	Max	Risk-loving	Risk-neutral	Risk-averse
Total	411	65.38	15.26	10	100	7.3%	14.35%	78.35%
Male	184	63.32	15.38	10	100	6.52%	17.93%	75.55%
Female	227	67.05	14.98	20	100	7.93%	11.45%	80.62%
Child with Siblings	198	66.67	15.11	10	100	6.06%	12.63%	81.31%
Only Child	213	64.18	15.32	10	100	8.45%	15.96%	75.59%

is a girl. In contrast, parents living in the urban areas, especially those working in state owned enterprises, schools, hospitals, government and other private firms are not allowed to have the second kid. Violation will result in large amount of fines, loss of jobs and criticism. Therefore, such "biased" policy results in lower cost of "illegal" baby in rural areas than in urban areas. This leads to the fact that only child is more likely to live in urban areas and child with siblings live in rural areas.

Unfortunately, we could not observe the residence for all 411 observations. Only the sample of experiment sessions in December 2011 has rural urban origin information, resulting in a rather limited sample of 125 observations[2]. Based on this smaller sample, the correlation between dummy variable "urban" and "only child" is 0.617, which provides a strong evidence of the favor toward rural residents in the policy. Because growing environment in rural and urban area is different too, if rural kids are more likely to be risk-averse than urban kids, we could also observe only child are less risk-averse.

The regression on risk-aversion is reported in Table 2.3. From the results in the first two columns, the marginal effects of gender and only child status are significantly different from zero, consistent with the comparisons in Table 2.2. However, the effect of only child status becomes insignificant when adding the residence information in the last column. Instead, we find that urban participants are significantly less risk-averse than participants who grow up in rural areas. Besides, due to the reduced sample size in the second column, the standard errors are all increased. Therefore, it suggests that the difference in risk preference between only child

[2]There are 51 male and 74 female subjects in the limited sample. Moreover, 57 subjects are children with siblings, and 68 are only children

and child with siblings might be partially driven by their different places of residence.

Table 2.3 Regression for risk preference

Risk aversion	(1)	(2)	(3)
Female	3.793**	4.442*	5.486**
	(1.501)	(2.591)	(2.610)
Only child	-2.575*	-8.414***	-4.226
	(1.493)	(2.557)	(3.260)
Urban			-7.155**
			(3.524)
Constant	64.620***	67.470***	69.550***
	(1.346)	(2.519)	(2.691)
Observations	411	125	125
R-squared	0.022	0.11	0.14

Standard errors in parentheses.
$* * *p < 0.01, * * p < 0.05, *p < 0.1$.

2.3.2 Social Preference in Task 2

In the dictator game of Task 2, each subject was required to decide her/his transfer amount to a receiver assuming s/he played the role of a dictator. As a result, Figure 2.6 shows the overall distribution of transfer amounts from dictators, and then Figure 2.7 and 2.8 report the differences between male and female, only child and child with siblings, respectively.

As mentioned before, subjects are classified into three types of social preference based on their transfer decisions. First, a subject is defined as a “selfish” type if s/he sent an amount of money less than 10 Chinese Yuan to a receiver so that s/he would earn more than the receiver; Second, a subject is defined as a “fair” type if s/he sent 10 Chinese Yuan and share an equal income with the receiver in the end; Third, a subject is defined as an “altruistic” type if her/his transfer amount was more than 10 Chinese Yuan. Accordingly, we report the proportion of those three types in Table 2.4.

For the overall population, over half of the subjects had the social preference of fair type so that the proportion of inequality-averse subjects (i.e. fair type) was significantly more than other two types. Moreover, there were only 52 subjects (i.e. 12.65%) kept all endowments

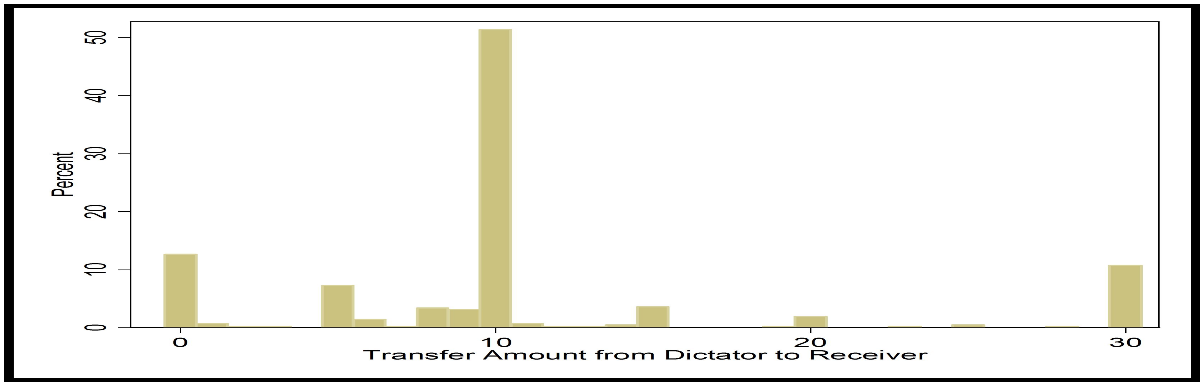


Figure 2.6 Distribution of social preferences in the overall sample

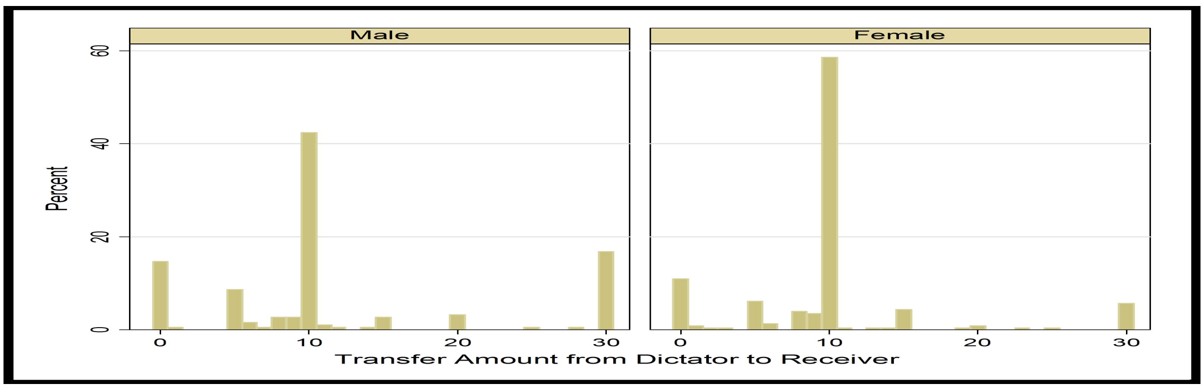


Figure 2.7 Distribution of social preferences – male vs. female

for themselves and passed nothing to their receivers in the dictator game. It implies that the absolutely self-interested individuals are quite rare in the real world, and provides a solid evidence to support the existence of other-regarding preferences.

As shown in Figure 2.7 as well as Table 2.4, 58.59% of female subjects had the social preference of fair type, whereas only 42.39% of male subjects were fair type. Moreover, the proportion of having altruistic social preference in male subjects (26.09%) was remarkably larger than the proportion in female (13.66%). Hence, we can conclude that our female participants cared more about the equalization in welfare distribution. This finding is consistent with the conclusion of Andreoni and Vesterlund (2001), in which the authors conduct a dictator game with varying income and prices, and then they find that men are more likely to be either

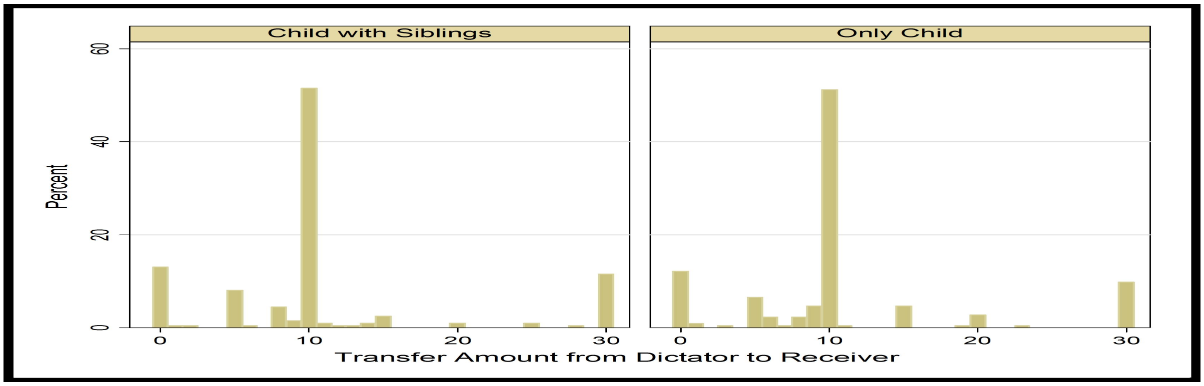


Figure 2.8 Distribution of social preferences – child with siblings vs. only child

perfectly selfish or perfectly selfless, whereas women tend to be equalitarian who prefer to share evenly.

Regarding the comparison between only children and children with siblings, there was no evidence to indicate any significant difference in their social preferences. In fact, Figure 2.8 presents almost the same distribution of transfer amounts for both groups of only children and children with siblings. And, the mean of transfer amount by only children was not significantly different from the mean of transfer amount by children with siblings ($p = 0.788$). Hence, contrary to what we expected, only children participants in our sample didn't behave more selfishly than children with siblings.

In addition, the regression on social preference is reported in Table 2.5.[3] From the results of Probit model in those three columns, only "female" has a significant effect on the probability of being inequality-averse (i.e. having "fair" type of social preference), whereas both only child status and rural urban origin information have insignificant impacts.

2.3.3 Individual Ability in Task 3

As shown in Figure 2.9, the distribution of individual productivity seems to be a normal distribution in the overall population. Specifically, it has a mean of 12.36 and a standard deviation of 2.71 (see Table 2.6).

[3]The dependent variable is the dummy of "inequality-aversion", which is equal to 1 if the subject prefers fairness so that s/he decided to send 10 Chinese Yuan to the receiver in the dictator game; and equal to 0

Table 2.4 Summary statistics of transfer amount and proportions of three social preferences

	Obs.	Mean	Std. Dev.	Min	Max	Selfish	Fair	Altruistic
Total	411	10.83	7.97	0	30	29.44%	51.34%	19.22%
Male	184	11.94	9.43	0	30	31.52%	42.39%	26.09%
Female	227	9.93	6.44	0	30	27.75%	58.59%	13.66%
Child with Siblings	198	10.94	8.24	0	30	28.79%	51.52%	19.69%
Only Child	213	10.73	7.73	0	30	30.05%	51.17%	18.78%

Table 2.5 Regression for social preference

Inequality aversion	(1)	(2)	(3)
Female	0.409***	0.318*	0.351*
	(0.125)	(0.231)	(0.236)
Only child	-0.018	-0.065	0.061
	(0.125)	(0.227)	(0.294)
Urban			-0.214
			(0.316)
Constant	-0.183	-0.184	-0.124
	(0.112)	(0.223)	(0.241)
Observations	411	125	125
Log likelihood	-279.367	-85.561	-85.332

The regression model is Probit model.
Dependent variable is “inequality aversion”.
Standard errors in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1$.

Table 2.6 Summary statistics of individual performance in the puzzle-solving task

	Obs.	Mean	Std. Dev.	Min	Max
Total	411	12.36	2.71	5	21
Male	184	11.98	2.79	5	20
Female	227	12.67	2.61	6	21
Child with Siblings	198	11.72	2.5	6	19
Only Child	213	12.96	2.76	5	21

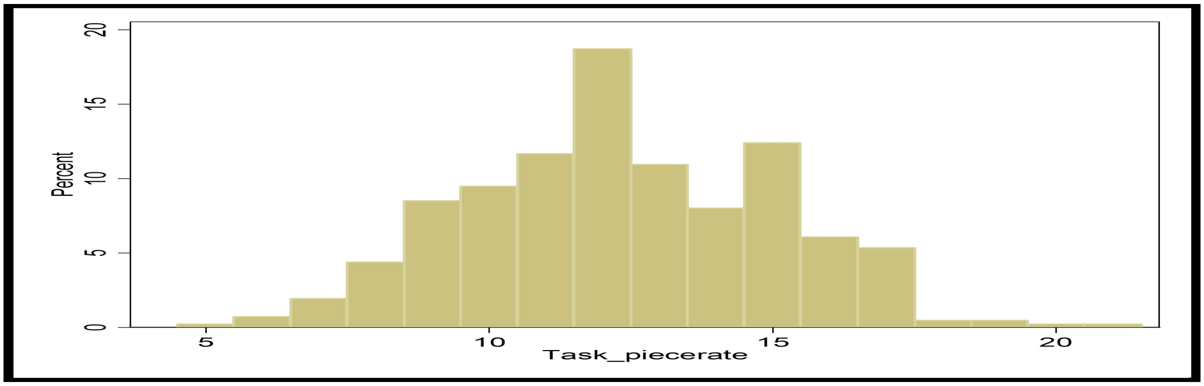


Figure 2.9 Distribution of puzzle-solving performance in the overall sample

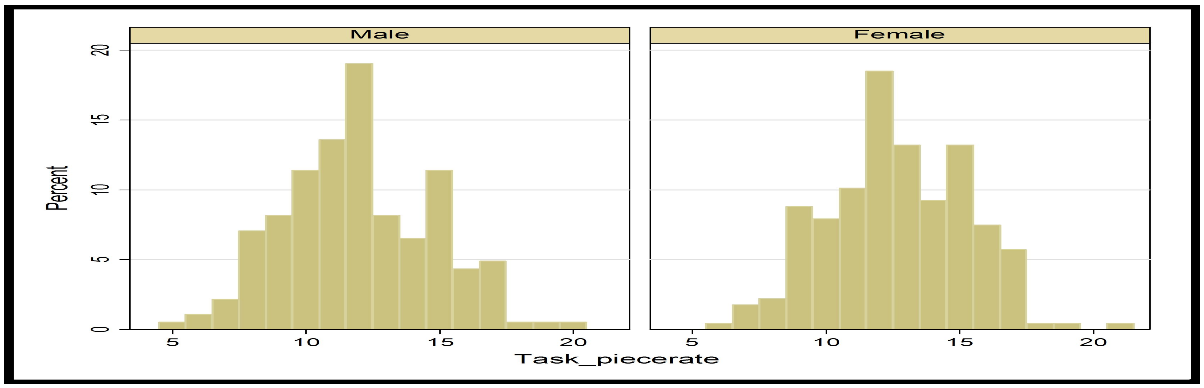


Figure 2.10 Distribution of puzzle-solving performance – male vs. female

However, if we consider male and female separately, then there was a significant gender difference on performance in this puzzle-solving task. That is, the average number of puzzle solved was 11.98 for male, which was significantly lower than the average performance of female (i.e. 12.67) at $p = 0.005$. Although we didn't expect such a gender difference in performance under this gender-neutral task, the fact shows that women might have better quick observation skills and/or because those female students took the experiment more seriously than the male participants and then put more effort during the task.

In addition, there was also a significant performance difference between only children and children with siblings. According to the intra-household allocation model of Becker and Tomes

otherwise.

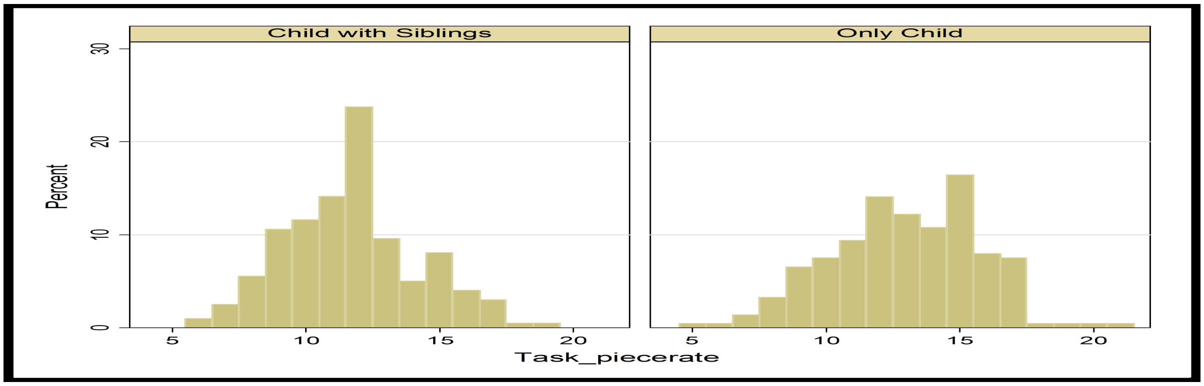


Figure 2.11 Distribution of puzzle-solving performance – child with siblings vs. only child

(1976), parents choose to allocate more resources to the more able child so as to maximize the lifetime income of all children. However, compared with larger family, parents of only child do not face such a resource allocation problem within the household. Therefore, labor market outcomes for only children are presumably ceteris paribus better than children with siblings. For instance, children in larger families are found to have lower levels of educational attainment (Steelman et al. 2002); and only children are verified to be more able than children with siblings at a math task (i.e. adding up five two-digit numbers) in the behavioral experiment of Cameron et al. (2013).

The regression results are reported in Table 2.7. In the first column, the marginal effects of gender and only child status on the individual performance in the puzzle-solving task are significantly positive, which is consistent with the comparisons in Table 2.6. However, when considering the smaller sample including residence information, the performance difference between different groups becomes insignificant.

2.3.4 Self-confidence in Task 4

Based on elicited self-assessment about their relative performance in the puzzle-solving task, we define a new variable "confidence bias" to measure the distance between subjects' self-assessments and their true percentile rankings. Hence, based on the value of "confidence bias", we can classify subjects into three types: an "under-confident" type if s/he has a negative

Table 2.7 Regression for puzzle-solving performance

Puzzle-solving performance	(1)	(2)	(3)
Female	0.657**	-0.013	-0.091
	(0.260)	(0.469)	(0.479)
Only child	1.230***	0.993	0.682
	(0.259)	(0.463)	(0.599)
Urban			0.532
			(0.647)
Constant	11.360***	12.040***	11.890***
	(0.234)	(0.456)	(0.494)
Observations	411	125	125
R-squared	0.067	0.037	0.042

Standard errors in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1.$

confidence bias; an "unbiased" type if her/his self-assessment is exactly the same as the true relative performance; and an "over-confident" type if her/his confidence bias is larger than zero.

In fact, according to the performance distribution within the overall population, the mean of actual percentile rankings should be a number around 50. However, since majority of subjects (i.e. 69.34%) were overconfident about their relative performance, the average self-assessment (i.e. 60.78) was significantly larger than the mean of actual percentile (i.e. 44.87). Actually, many social psychologists have pointed out that people always tend to be overconfident about themselves and systematically rate their own ability as "above average". Take one classic example, 88% of US drivers consider themselves safer than the median driver (Svenson 1981).

Similar to the results in Niederle and Vesterlund (2007), we find that both men and women were over-confident, but men were more overconfident about their relative performance than women ($p = 0.017$).

In addition, compared to those only children, children with siblings were more over-confident about their relative performance, and the difference between two groups was statistically significant ($p = 0.000$). Again, our finding is different from the results in Cameron et al. (2013), where the authors documented that the difference of self-confidence between Chinese only children

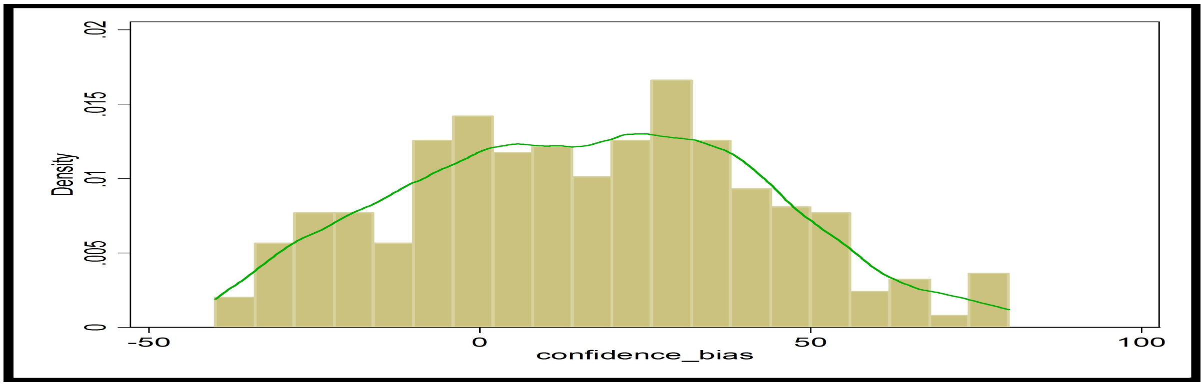


Figure 2.12 Distribution of confidence bias in the overall sample

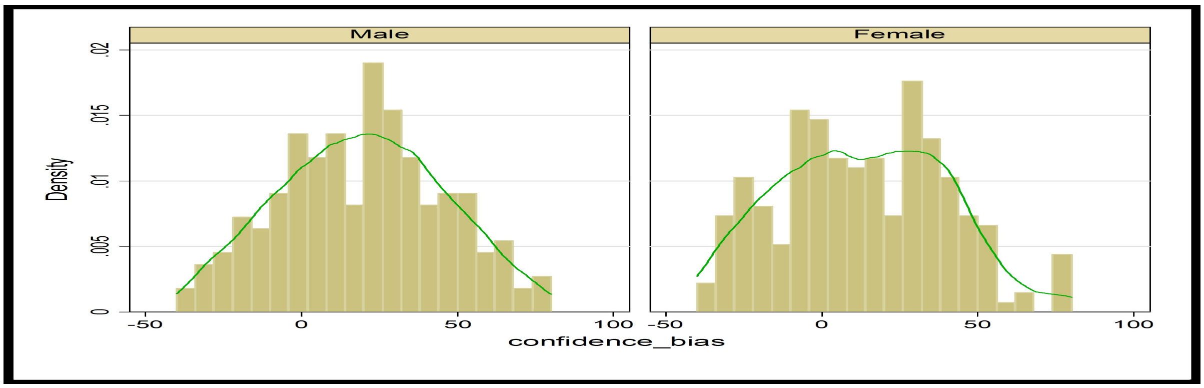


Figure 2.13 Distribution of confidence bias – male vs. female

and children with siblings is not significant.

Finally, we make regression analysis to investigate the determinants of self-confidence. We select "confidence bias" as the dependent variable, and then exam the effects of individual performance in puzzle-solving task, gender, only child status, rural/urban residence and GPA score. As a result, in Table 2.9, the level of confidence bias is negatively related to their performance in the task, which means the subjects with low performance are more likely to over-estimate their relative performance than the subjects with high performance. In addition to their performance in this specific task, subjects' confidence levels are also related to their GPA scores. As the index of students' overall academic performance, a higher GPA score provides the subjects with more confidence about themselves, which also leads a tendency to

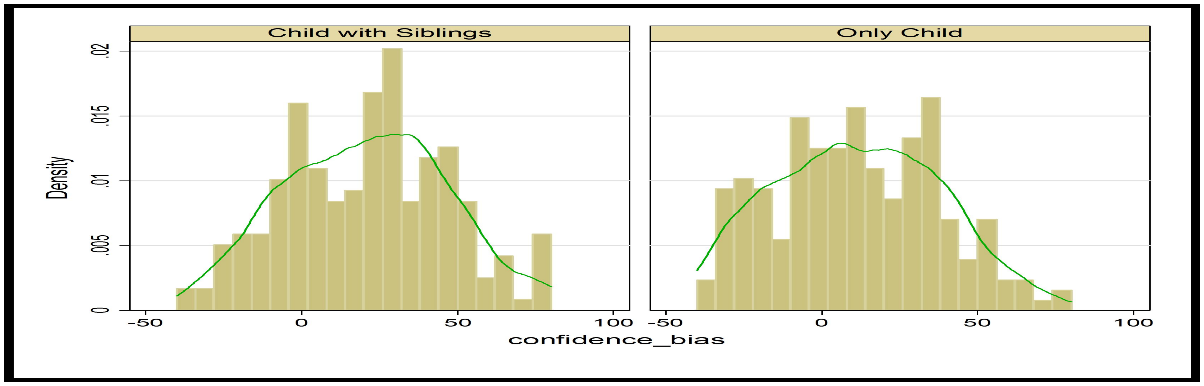


Figure 2.14 Distribution of confidence bias – child with siblings vs. only child

over-estimate their performance in the puzzle-solving task. Therefore, according to the negative relationship between task performance and over-confidence level, since the average puzzle-solving performance of women and only children are higher than their counterparts, then both women and only children are observed to be less over-confidence. However, after controlling the effect of performance difference, only the dummy of "female" still has a significant marginal effect on the confidence bias.

Table 2.8 Summary statistics of self-confidence and proportions of two types of bias

	Obs.	Actual Percentile	Self Assessment	Confidence Bias	Under-confident	Over-confident
Total	411	44.87	60.78	15.91*	29.20%	69.34%
		(29.85)	(18.49)	(26.68)		
Male	184	42.54	61.94	19.40*	25.00%	75.00%
		(29.2)	(19.63)	(26.3)		
Female	227	46.75	59.84	13.09*	32.60%	64.76%
		(30.31)	(17.49)	(26.71)		
Child with Siblings	198	38.45	59.22	20.77*	23.74%	74.24%
		(28.02)	(18.40)	(26.19)		
Only Child	213	50.84	62.23	11.39*	34.27%	64.79%
		(30.32)	(18.50)	(26.41)		

Notes: Standard deviations in parentheses;
$*$ indicates the value is significantly different from zero at $p < 0.001$.

Table 2.9 Regression for self-confidence bias

Confidence bias	(1)	(2)	(3)
Puzzle-solving performance	-6.556***	-7.419***	-7.422***
	(0.371)	(0.670)	(0.674)
GPA score	5.865**	10.170**	10.200**
	(2.903)	(5.041)	(5.084)
Female	-2.839*	-10.090***	-10.140***
	(2.034)	(3.554)	(3.653)
Only child	-1.129	-1.322	-1.498
	(1.995)	(3.488)	(4.472)
Urban			0.307
			(4.831)
Constant	79.750***	83.220***	83.070***
	(10.340)	(18.280)	(18.490)
Observations	411	125	125
R-squared	0.462	0.537	0.537

Standard errors in parentheses.
$***p < 0.01, **p < 0.05, *p < 0.1$.

2.4 Conclusions

From our experimental data, we can verify that there are significant gender differences on those non-cognitive characteristics: on average, men are less risk-averse and more over-confident than women; regarding the social preference, men are more likely to be either perfectly selfish or perfectly selfless, whereas women tend to be equalitarian who prefer to share evenly. Meanwhile, we also obtain some intriguing findings about only children participants in our sample of Chinese university students. For instance, only children are less risk-averse and also less over-confident than the children with siblings; however, there is no significant difference in social preference between two groups, which provides experimental evidence against the stereotype that only children are more likely to be selfish.

More importantly, what do our results mean in terms of the way people behave outside the experimental setting? Previous studies have shown that individual behavior in the lab correlates with behavior elsewhere. For example, prosocial behavior in the lab was associated with more charitable giving in the field (Benz and Meier 2008); decisions made in experimental lotteries were correlated with real life financial decisions (Castillo et al. 2010). Hence, the difference of those individual characteristics between women and men probably result in a gender gap in their occupational choices. And, our study also suggests that China's One-Child policy should have important impacts on the labor market because of the characteristics differences between only children and others.

CHAPTER 3. WHAT MATTERS FOR PRODUCTIVITY? THE ROLE OF INCENTIVE, INFORMATION AND SIGNUP BONUS

Abstract

People with similar abilities might perform quite differently in the workplace. We study this question through an experimental labor market in which we vary the incentive scheme, the information about relative performance, and also the additional compensation in terms of sign-up bonus. We create a laboratory experiment where "workers" can join "companies" that pay them according to different incentive schemes: (a) piece rate; (b) revenue sharing where all earnings are equally shared among team members; (c) individual tournament where only the highest performer in the team can earn a prize; and (d) team tournament where teams should compete first, and then prizes are shared within the winning team. As a result, we find that: (1) Compared to the baseline performance paid by piece rate, under three team-based incentives, more competitive incentive generates higher performance improvement. (2) Feedback about relative performance reduces the performance differences between three team-based incentives. (3) Regardless of incentives and feedback, the extra sign-up bonus brings a positive and significant effect on individual performance.

3.1 Introduction

The modern workplace puts an ever greater emphasis on teamwork: employees work in virtual teams with colleagues, suppliers, clients, and even competitors. They work in ad hoc combinations, and the interaction between incentive scheme and work organization has become increasingly prominent. Hence, in order to investigate the differences between individual-based scheme and team-based institutional structures in a controlled laboratory environment, we create an experimental labor market where “workers” can join “companies” that pay them according to different compensation schemes: (a) piece rate, (b) revenue sharing, (c) individual tournament, and (d) team tournament. First of all, we select the piece rate as our baseline pecuniary incentive for individual-based working, and subsequently design three representative team-based incentive systems to highlight the tradeoff between cooperation and competition. The first team-based incentive is “revenue sharing”, which focuses on collaboration within the team, so this payment scheme requires people to work together and all earnings are equally shared among team members; oppositely, in order to emphasize the competition within the team, the second team-based incentive is defined as “individual tournament”, which compels everyone to compete with others in the same team and only offers a prize to the best performer in the end. However, in the real world, cooperation and competition always exist on a continuum, and few of the workplaces are wholly cooperative or wholly competitive. Hence, as the third team-based incentive, we create a hybrid of “revenue sharing” and “individual tournament”, i.e. “team tournament”. This incentive requires teams to first compete with each other, and then a reward will be split among the members of winning team.

In this chapter, we aim to figure out the most efficient way(s) to improve workers’ performance by considering three potential influence factors at the same time, that is: team-based incentive in terms of cooperation and/or competition, feedback about relative performance, and an extra pay from sign-up bonus. In other words, the purpose of this chapter can be summarized by three research questions as follows:

First, do people vary their performance a lot in response to different team-based incentives? If so, which incentive is most effective for increasing productivity, and which individual characteristics affect their productivity response to various types of compensation?

Second, does feedback about relative performance matter? In particular, regarding the team-based incentives that determine the workers’ payoffs by their relative performance in a team, do workers change their performance significantly if knowing more information about their relative performance?

Third, in addition to a certain incentive per se, how does an extra compensation in terms of sign-up bonus influence individual performance?

Overall, our design is somewhat close to the approach used by Dohmen and Falk (2011), but also differentiates in several aspects. In particular, Dohmen and Falk (2011) didn't make a within-subjects comparison of individual performance between different incentives but only had a between-subjects comparison of individual performance when all subjects worked under their most preferred incentives. On the contrary, in order to disentangle the exogenous incentive effect and the endogenous self-sorting effect on individual performance, our experiment adopts a within-subjects design that forces all workers to complete four real-effort tasks paid by different incentives respectively. As a result, in the absence of self-selection, the average performance achieved the highest level under individual tournament. Moreover, compared with the baseline performance paid by piece rate, the order of performance improvement under three team-based incentives is consistent with the rank of their competitiveness; i.e. the more competitive the incentive is, the higher the improvement is generated. Furthermore, if introducing an additional compensation in terms of prepaid sign-up bonus, then no matter which incentive scheme is provided, subjects remarkably increase their performance in response to this extra pay of sign-up bonus. And, we verify that such a positive effect of sign-up bonus on individual performance dominates the effect of incentive per se.

In addition, we also study how subjects respond to information about their relative performance under those team-based incentives. Our findings show that the provision of feedback information lessens the performance differences between three team-based incentives so that the average performance under individual tournament is no longer better than the average performance under any other incentives.

Last but not least, our experiment uses a rich battery of diagnostic games to elicit several non-cognitive characteristics such as risk attitude, social preference and self-confidence before subjects start the real-effort tasks. Although many authors (Niederle and Vesterlund 2007; Wozniak et al. 2010) have stressed the importance of those non-cognitive characteristics on labor market outcomes, only very few studies (e.g. Dohmen and Falk 2011) use incentivized elicitation methods to measure them. As a result, lab data from our sample of 411 Chinese university students verifies that subjects of similar ability perform quite differently under the same incentive, because individual characteristics such as risk attitude and self-confidence influence their productivity in a systematic way. For instance, without feedback information, those risk-loving and over-confident subjects tend to make less effort under any team-based incentives; however, the effects of both risk attitude and self-confidence on individual performance become insignificant after subjects receive feedback about their relative performance.

The rest of this chapter is organized as follows. Section 2 presents a literature review. Section 3 describes our experimental design. The results are given in section 4. Finally, in section 5, further discussions and some concluding remarks are presented.

3.2 Literature Review

Recent labor market research suggests that workers of similar ability might behave quite differently in the workplace. Hence, with the help of a well-controlled environment in the laboratory, there are a lot of experiments designed to investigate the individual productivity differences between various incentive schemes. For example, Gibbons (1987), Booth and Frank (1999) and Lazear (2000) focus on the performance variance under piece rate scheme. Another frequently used incentive format is competition within a team. Rank-order tournaments may motivate employees to work harder (Bull et al. 1987; Harbring and Irlenbusch 2003). But it may also be demoralizing and create an excessively stressful workplace, which may hinder overall performance (see Lazear 1989). For the contrasting incentive in terms of collaboration, what happens under a revenue sharing is also studied (e.g. Nalbantian and Schotter 1997; Gachter and Fehr 1999). Overall, those papers examine the effective incentive(s) that can encourage workers to exsert maximum effort and also discuss the mechanism design to make an optimal contract between employees and employers. In our study, various incentives based on individual, team and relative performance are compared. Similar to the results in the previous literature (e.g. van Dijk et al. 2001[1]), we also find that compared with piece rate or any cooperative incentives, individual tournament is the most effective way to stimulate individual productivity in a competitive environment. Meanwhile, the "free-riding" problem in

[1] The results in van Dijk et al. (2001) show that individual and team payment induced the same effort levels. In team production free-riding occurred, but it was compensated by many subjects providing more effort than in case of individual pay. Effort was higher, but more variable in tournaments, while in case of varying abilities workers with relatively low ability worked very hard and drove up effort of the others.

the teamwork is not as severe as we expect.

In addition to incentive scheme per se, feedback about relative performance has been found to influence behavior even when conventional economic analysis says it should not matter. In fact, social psychologists have documented many non-monetary effects of relative performance feedback such as social pressure, fear of informal sanctions, and shame (e.g. Kluger and Denisi 1996). Economists have recently started to formally model how peer effects interact with monetary incentives (see Krakel 2008). Moreover, in the laboratory experiments, Falk and Ichino (2006), Mas and Moretti (2009) show that allowing people to observe each other's performance has a positive peer effect on an employee's performance although her/his payment is independent of the others' performance[2]. But there are also some papers have considered the influence of competitors' ability on own effort in a competition setting (e.g. Eriksson et al. 2009; Freeman and Gelber 2010; Wozniak et al. 2010). Therefore, in order to investigate how information interacts with incentives especially how relative performance feedback affect individuals productivity response to various team-based incentive, we apply a horizontal comparison between two groups, of which one half subjects receive the feedback of their relative performance after they finish their productivity test, and the other half never receive any feedback information during the experiment.

Last but not least, although the policy of sign-up bonus is very prevalent in the real world, economists have paid little academic attention to this issue. So far as I know, there are only a few papers on this topic. In Van Wesep (2009), the author provides a theoretical explanation for the claim that "the bonus is a way for high quality firms to differentiate themselves from low quality firms, attract the best employees and induce them to work hard". And, Arya et al. (2003) studies various employee incentives in a principle-agent model and then suggests that the sign-up bonus is a useful instrument for mitigating hold-up problems. Different from those theoretical studies, our experimental research is much more straightforward to verify the role of sign-up bonus on employee behavior.

[2] Falk and Ichino (2006), Mas and Moretti (2009) both use a flat way pay scheme in their laboratory experiments.

3.3 Experimental Design

3.3.1 Procedure of the Lab Experiment

In order to study how incentive, feedback and sign-up bonus differently affect individual performance, we implemented an experiment which consists of 8 tasks and is divided into 3 stages: Stage 1 includes the first four tasks that are designed to elicit each subject's risk preference (in Task 1), social preference (in Task 2), baseline productivity in a puzzle-solving task paid by piece rate (in Task 3) and then the self-assessment about her/his relative performance (in Task 4). After the first four diagnostic tasks complete, subjects were randomly assigned into two different treatment groups: half subjects were given detailed feedback about their relative performance in Task 3 of puzzle-solving game, and the other half never knew any relative performance feedback during the whole experiment. Next, the following Stage 2 consists of Task 5, 6 and 7: by those three tasks, we observed individual performances in similar puzzle-solving games but paid by three different team-based incentives, i.e. revenue sharing, individual tournament and team tournament. And then, during Stage 3, all subjects completed the last puzzle-solving task i.e. Task 8 which was paid by a certain incentive plus an extra sign-up bonus. Finally, when all tasks ended, a simple questionnaire was used to gather individuals' socio-demographic information including gender, age, marital status, major fields of study, grades, etc.

Subjects were only given a brief introduction of the experiment at the very beginning, so they knew that they had to complete several tasks but did not know what those tasks would look like. Just before starting each task, subjects were informed of the rules for that task. Besides, subjects were told that their final earnings from the experiment would be the payment for one randomly selected task out of eight, plus a show up fee of 10 Chinese Yuan.

3.3.1.1 Stage 1 (Task 1 – 4): Elicitation for Individual Characteristics

Task 1: Elicit Risk Preference This task used lottery choices to elicit individual risk preference. Subjects were asked to answer 10 binary choices questions.

In each binary choice, subjects need to decide whether they prefer a safe option or playing a lottery. The safe option is always the same in each question, which is 20 Chinese Yuan as the fixed payoff. However, the lotteries vary and the risk of lottery option decreases from question to question. In other words, the probability of winning 40 Chinese Yuan increases from 10% to 100% in the end.

Now suppose a subject has consistent risk preference on lotteries, s/he will prefer the safe option to the lotteries above a certain risk level, and then switch to preferring the lottery option in all subsequent choice questions. Therefore, the switching point of winning 40 Chinese Yuan in a lottery indicates the subject's risk attitude. That is, the higher this switching point is, the more risk-averse s/he is.

Task 2: Elicit Social Preference Subjects' social preferences were elicited by a dictator game. In this dictator game, each participant was randomly matched with another participant in the lab room, playing the role of a dictator or a receiver. As a dictator, s/he could get 30 Chinese Yuan as endowment at the beginning of the game, and could either keep all money or send some amount to another person, i.e. the receiver. On the other hand, as a receiver, s/he had no endowment, but could receive something if the dictator passed money to her/him. Moreover, whatever amount the dictator decided to send, the money was doubled by the experimenter and then added to the receiver's payoff. In order to classify each subject, everyone had to play the role of dictator and decide how much sent to the receiver from an endowment of 30 Chinese Yuan. After all transfer choices made, subjects were randomly matched in pairs and two roles within each pair were assigned by a random draw to determine their final payments in this task.

From the individual choices in this dictator game, we are able to identify the social preference for each subject. According to the game rule, if and only if the dictator sent 10 Chinese Yuan to the receiver, then both players could have 20 Chinese Yuan equally. Otherwise, the receiver obtained less than the dictator if the transfer amount was lower than 10 Yuan whereas the receiver got more than the dictator if the transfer amount was higher than 10 Yuan. Therefore, based on each dictator's wealth distribution between the receiver and her/himself, we define the "inequity aversion" as the social preference for fairness and resistance to incidental inequalities, which is equal to 1 if the dictator sends 10 Chinese Yuan and equal to 0 otherwise.

Task 3: Elicit Performance under Piece Rate The result in this task was an individual productivity indicator measuring subjects' performance independent of any team-based incentives. Therefore, all subjects were required to solve letter-puzzles in 3 minutes and paid by piece rate.

In each puzzle, two similar squares of letters are presented at the same time, but the square of letters on the right differs from the square of letters on the left in two letters. Subjects had to find those letters and click on them to solve the puzzle. Whenever they clicked a letter other than those two letters, the remaining time was reduced by 1 second. Hence, people had to be more cautious and should try to only click letters that differ. Within a 3-minute period, each subject was asked to solve as many puzzles as possible, and s/he could increase her/his payoff by a high productivityBefore the task started formally, subjects had three practice examples to familiarize themselves with this letter-puzzle, in which no payment was involved.. A clock on the screen informed how many seconds they had used.

Once this task finished, all subjects were informed of their own performance, but nothing about the performance of others. The pay scheme in this task was "piece rate", i.e. subjects would receive 1 Chinese Yuan for each solved puzzle. The reasons for choosing piece rate as the monetary incentive in Task 3 include: piece rate contract is a good proxy for subjects' productivity since it is a compensation scheme based on individual absolute performance only and excludes the competition and/or cooperation with others. Besides, it is an intrinsic motivation for everyone to exert maximum effort and elicit their real ability levels: the better someone performed, the more s/he would earn. More importantly, as mentioned before, our study emphasizes the differences between individual-based incentive and team-based incentives. In the following tasks, we will introduce three different incentives related to relative performance and team pay. Hence, the performance in piece rate can be treated as a benchmark under individual-based incentive and then compared with those three team-based incentives later.

Task 4: Elicit Self-confidence This is a task for eliciting self-assessment about their relative performance. Each subject was asked to estimate the probability that her/his performance in the previous puzzle-solving task paid by piece rate (Task 3) was better than a randomly selected other player in the lab room. In fact, subjects could increase their earnings by submitting a good estimate. We used a novel crossover mechanism to elicit their self-confidence.[3] In the experimental instruction, we presented this mechanism in a simple narrative form. We told subjects that they were paired with a "helper robot" who would also take the puzzle-solving task and who had a certain fixed probability q of winning against a randomly selected player. This probability was between 0 and 100% but unknown to the subject. Subjects were told that they could have their winnings based on their robot's performance rather than their own, and asked to indicate a threshold level of q above which they preferred to use the robot's performance. As a consequence, when the helper robot's winning probability was higher than the subject's indicated threshold, the robot's performance would be used to compete with the subject's opponent; otherwise, subject's own performance would be used against her/his selected opponent. In the end, only if the subject's or the helper robot's performance was better than the performance of the other selected player, 10 Chinese Yuan could be added to the individual earnings. Therefore, subjects should maximize their probability of earning an extra 10 Chinese Yuan by choosing the threshold as their own subjective probability of being better than some other player.

Provision of Relative Performance Feedback When the first four diagnostic tasks ended, participants were randomly assigned into two different treatment groups: half subjects were given the list of all subjects' performance in the puzzle-solving task (Task 3), and then they were clear about how well they did in that piece rate task compared with other players (see an example of feedback information in Figure /refmgraph3.1). In contrast, another half subjects were not provided with any information about their relative performance in Task 3. According to the comparison between these two treatment groups, it allows us to systematically examine the effects of feedback on individual productivity response to the following three team-based incentives.

[3] Mobius et al. (2011) is the first paper to implement the crossover mechanism in an experiment. The crossover mechanism has one main advantage over the otherwise popular quadratic scoring rule which most experimental papers use, i.e. it works for a wide class of preferences: it is well known that the quadratic scoring rule only produces truthful reports if subjects are risk-neutral. If subjects are risk-averse, quadratic scoring induces biased reports. In contrast, the crossover mechanism only requires that subjects' preferences are monotonic in the following sense: among the class of lotteries that pay off $x > 0$ with probability q and 0 with probability $1 - q$, subjects strictly prefer those with higher q. Truthful reporting of subjective beliefs is then a strictly dominant strategy. Monotonicity holds for all von-Neumann-Morgenstern preferences as well as many non-standard preferences such as Prospect theory.

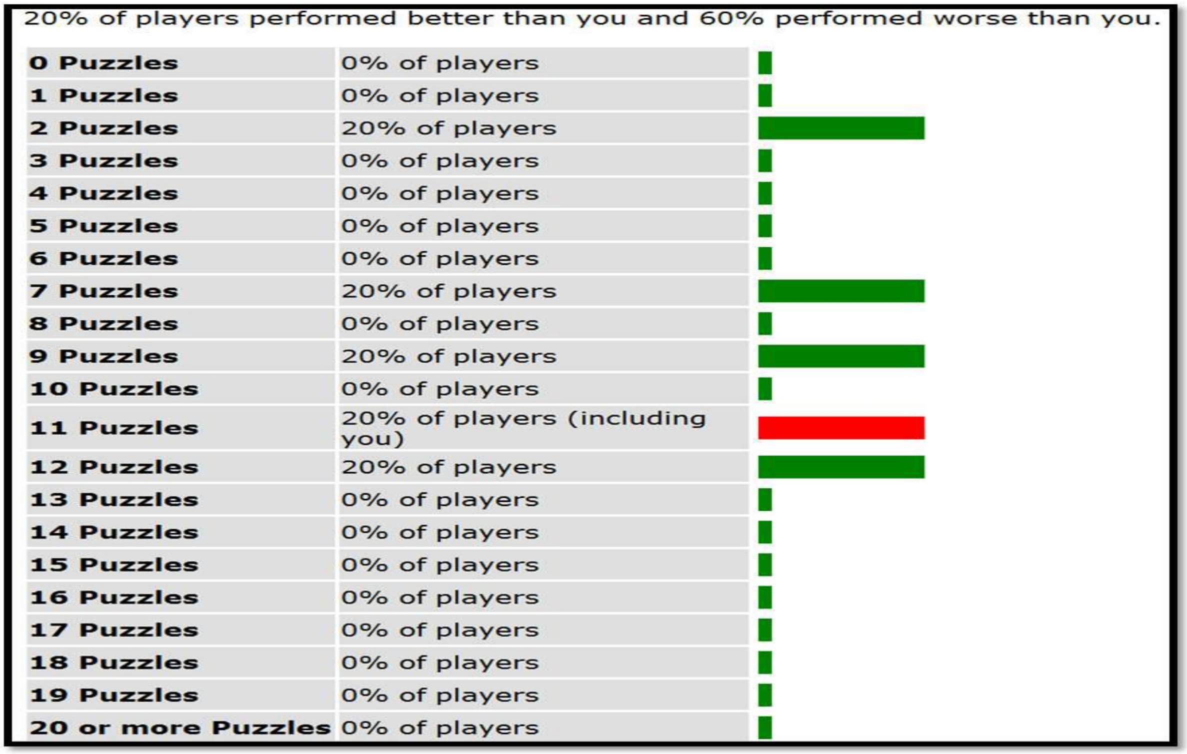


Figure 3.1 Feedback about relative performance in the puzzle-solving task

3.3.1.2 Stage 2 (Task 5 – 7): Performance under Three Team-based Incentives

During the second stage, subjects were asked to do three more tasks. In those three tasks, everyone had to complete letter-puzzles similar to Task 3 and also needed to solve problems as many as possible in 3 minutes. The only difference is the payment scheme: from Task 5 to Task 7, three distinct team-based incentives were introduced to subjects. More specifically, in our experimental labor market, each subject was assigned to be an employee of a "company" that specialized in solving puzzles. In addition, there were three companies and each company consisted of two teams of two employees but compensated employees differently.

Performance under "Revenue Sharing" The first company paid its employees through "revenue sharing" scheme. Each team of two employees would receive 1 Chinese Yuan for every puzzle that two team members solved in 3 minutes. And then, the earnings would be divided equally between two members of each team.

Performance under "Individual Tournament" The second company paid its workers through "individual tournament" scheme. The employee would only be paid if s/he solved more puzzles than her/his teammate in 3 minutes and could receive 2 Chinese Yuan for every solved puzzle. Otherwise, s/he would earn nothing.

Performance under "Team Tournament" The third company paid its workers through "team tournament" scheme. The employee would only be paid if her/his team solved more puzzles than the other team in 3 minutes. In this case, the winning team would receive 2 Chinese Yuan for every puzzle that both team members solved in 3 minutes. And then, the reward for the winning team would be divided equally between the two members.

Apparently, the subject's final payment in any of those three schemes is not only depending on an individual absolute performance, but also related to other teammates' performance at the same time. In addition, it is important and necessary for us to consider the order effects of incentives' arrangement in the experimental design. If all subjects are given the same order of incentives rather than in different orders, it cannot disentangle potential order effects from incentive effects on individual performance. Such order effects include a positive learning effect that subjects perform better once they familiarize the task after initial rounds and also a negative time effect that people behave worse due to fatigue after repeated tasks. Therefore, in Stage 2, we randomly assigned the order of three team-based incentives for each subject, which helped us control for the order effects on individual performance.

3.3.1.3 Stage 3 (Task 8): Performance Paid by Additional Sign-up Bonus

Again, subjects were required to solve puzzles as many as possible within 3 minutes. However, different from the previous stages, the system assigned a certain incentive as well as a prepaid sign-up bonus for every subject in last task. The sign-up bonus was a random integer between 0 and 10, and all workers were informed of the given incentive and sign-up bonus before the task.[4]

[4]According to our design, before starting Task 8, we elicited participants' preferences for given incentives through a self-sorting process. In particular, we asked subjects to not only rank those four payment schemes, that is: (1) piece rate, (2) revenue sharing, (3) individual tournament, and (4) team tournament, from most preferred to least preferred, but also state their compensating differentials between any two consecutive incentives. And then, based on individual stated-preferences, the system would automatically determine an incentive scheme as well as an contingent compensation of extra sign-up bonus for each subject.

Therefore, based on individual performance in the last puzzle-solving task, each worker's payoff was calculated by the assigned incentive scheme plus the sign-up bonus. In particular, if someone would be paid by a team-based incentive in this task, her/his co-workers would be randomly selected among the players in the lab room; however, the performance of other three co-workers in the same company were their past performance when they worked for this company in a previous round of Stage 2.

3.3.1.4 Follow-up Questionnaire

In the end, after eight tasks finished, subjects had to complete a questionnaire for personal data on socioeconomic characteristics (including gender, age, marital status), on educational achievement (major fields of study, GPA, score of university-entrance examination) and feedback about this experiment experience (the clearness of the instruction, the length of the experiment, the difficulty of the tasks).

All in all, the flowchart of the whole experimental design is described in Figure 3.2 .

3.3.2 Subject Pool

In 2011, we conducted this experiment at Central University of Finance and Economics in Beijing, China. We had 411 complete observations, including 184 males and 227 females, and all of them were university students. In addition, there were 213 only children (and 198 children with siblings) in our subject pool. Because majority of the participants i.e. 70% were graduate students, the average college year was above 4 and the average age was 23.03 correspondingly. Only 2% of the participants were married. Finally, the average level of GPA was 3.3 (on a 4.0 scale) with a standard deviation of 0.35. However, since subjects' grades were self-reported, we cannot guarantee the accuracy of this information and treat it as a quite limited index to personal academic performance.

A total of 21 experiment sessions took place in the same computer lab at the university, and the average size of those sessions was 19.57 participants (with a standard deviation of 6.62). Each session lasted, on average about 50 minutes. Final payments for subjects were based on one randomly selected task out of eight tasks, plus a fixed show-up fee of 10 Yuan. The average payment was 30.85 Yuan.

3.4 Results

3.4.1 Performance Differences

Since this experiment aims to discuss the impacts of three factors on individual performance, that is: incentive, information about relative performance and extra sign-up bonus. Therefore, according to our design, we have $4\times2\times2 = 16$ treatments, in the form of $x-y-z$, where $x =$PR (Piece Rate), RS (Revenue Sharing), IT (Individual Tournament), TT (Team Tournament); $y =$NF (No Feedback), F (Feedback); and $z =$NB (No Bonus), B (Bonus). As a result, Table 3.1 summarizes the average performance in different treatments.

We first consider the performance data without feedback and sign-up bonus. As compared with the baseline output paid by piece rate scheme, the incentive of revenue sharing does not affect the average performance ($p =$0.625), whereas both competitive schemes have positive effects on performance improvement, which are statistically significant at $p =$0.010 under individual tournament and $p =$0.043 under team tournament. Besides, although the average performance under individual tournament is somewhat higher than the average performance under team tournament, the difference is not significant ($p =$0.603).

Secondly, the feedback information has significant impact on the performance paid by revenue sharing. Compared with the data in the absence of feedback, providing subjects with full information about their relative performance does not affect their productivity sorting in those three team-based incentives. However, when comparing to the baseline output paid by piece rate, all three team-based schemes have positive effects on performance, which are statistically significant at $p =$0.023 under revenue sharing, $p =$0.009 under individual tournament and $p =$0.026 under team tournament. Besides, although the average performance under individual tournament is somewhat higher than the average performance under team tournament, the difference is not significant ($p =$0.603). Meanwhile, the performance differences between those three team-based incentives become insignificant. Hence, we can conclude that feedback

Table 3.1 Summary statistics of performance in different treatments

Treatment	Obs.	Mean	Std. Dev.	Min	Max
PR-NF-NB	209	12.51	2.61	6	19
RS-NF-NB	209	12.60	2.65	3	19
IT-NF-NB	209	12.97	2.92	3	21
TT-NF-NB	209	12.87	2.74	3	21
PR-F-NB	202	12.21	2.80	5	21
RS-F-NB	202	12.62	2.60	7	20
IT-F-NB	202	12.71	2.72	6	20
TT-F-NB	202	12.64	2.78	6	21
PR-NF-B	69	14.70	2.41	8	21
RS-NF-B	30	13.37	2.66	7	18
IT-NF-B	63	14.98	2.90	8	21
TT-NF-B	47	15.00	2.48	8	20
PR-F-B	64	13.83	2.86	5	18
RS-F-B	48	13.54	2.62	7	18
IT-F-B	51	15.86	2.26	9	24
TT-F-B	39	14.56	2.52	9	21

Notes:
The first four treatments compare four incentives while holding “No Feedback” and “No Bonus”;
The second four treatments compare four incentives while holding “Feedback” and “No Bonus”;
The third four treatments compare four incentives while holding “No Feedback” and “Bonus”;
The last four treatments compare four incentives while holding “Feedback” and “Bonus”.

information promotes better performance under team revenue sharing, while it cannot improve performance in both competitive incentives.

Finally, the extra sign-up bonus brings a positive and significant effect on individual performance regardless of incentive and feedback treatments. By controlling the incentive scheme and feedback, the results of pairwise comparison in Table 3.2 below shows that all the within-pair differences are significantly positive.

Table 3.2 Summary statistics of within-pair performance comparison

Treatment	Obs.	Average NB	Performance B	Within-pair Differences
PR-NF	69	12.22	14.70	2.48*
		(2.48)	(2.41)	(0.27)
RS-NF	30	11.73	13.37	1.63*
		(3.15)	(2.66)	(0.45)
IT-NF	63	13.62	14.98	1.37*
		(2.90)	(2.90)	(0.35)
TT-NF	47	13.23	15.00	1.77*
		(2.81)	(2.48)	(0.36)
PR-F	64	11.78	13.83	2.05*
		(2.64)	(2.86)	(0.34)
RS-F	48	12.19	13.54	1.35*
		(2.97)	(2.62)	(0.37)
IT-F	51	13.61	15.86	2.25*
		(2.38)	(2.26)	(0.28)
TT-F	39	13.08	14.56	1.49*
		(2.43)	(2.52)	(0.40)

Notes: Standard deviations in parentheses;
$*$ indicates the value is significantly different from zero at $p < 0.01$.

In addition to those three influence factors that are targeted, it's very important us to consider the time effects on individual performance during the experiment. Recall the procedure, the puzzle task in Stage 1 is paid by piece rate task, and then three more puzzle tasks are paid by three different team-based incentives in Stage 2. However, the orderings of those three team-based incentives are randomly assigned. For instance, the incentive used in Task 5 was possibly any one out of those three team-based incentives. So, this design allows us to control and exam the time effect on individual performance. As indicated in Table /reftable3.3, during stage 2, there is a clear time trend on the performance irrespective of specified team-based incentives and/or the provision of feedback information: First of all, the learning effect helps subjects to improve their performances from Task 5 to Task 6, which are statistically significant at p =0.000 in both "No Feedback" and "Feedback" treatments and p =0.043 with "Feedback" treatment, respectively. Afterwards, a negative effect from fatigue decreases the average performance in the last task of Stage 2 (i.e. Task 7), and the reductions are significantly from zero at p =0.000 in both "No Feedback" and "Feedback" treatments. Moreover, the difference between Task 5 and Task 7 is not significant in "No Feedback" treatment (p =0.206) but becomes statistically significant at p =0.002 in "Feedback" treatment.

Table 3.3 Summary statistics of within-pair performance comparison between consecutive tasks in Stage 2

	No Feedback			Feedback		
Task	Obs.	Average Performance	Within-pair Difference	Obs.	Average Performance	Within-pair Difference
Task 5	209	12.37	–	202	12.39	–
(1st task in Stage 2)		(2.39)			(2.29)	
Task 6	209	13.86	1.49*	202	13.65	1.27*
(2nd task in Stage 2)		(3.16)	(2.39)		(3.12)	(2.75)
Task 7	209	12.2	-1.66*	202	11.94	-1.72*
(3rd task in Stage 2)		(2.39)	(2.57)		(2.3)	(2.61)

Notes: Standard deviations in parentheses;
$*$ indicates the value is significantly different from zero at $p < 0.01$.

3.4.1.1 Regression Analysis

Now, we continue to probe the effects of three factors by regression models. Based on the data collected from Stage 1 and 2, in the absence of sign-up bonus, we discuss the incentive effects and feedback effect, respectively.

In the first fixed-effect model, we focus on the incentive effects of different team-based payments.

$$\dot{Y}_{ij} = \delta_j + \alpha_i + \varepsilon_{ij} \tag{3.1}$$

where $\dot{Y}_{ij} = Y_{ij} - Y_{i0}$ represents the change in individual i's performance from Y_{i0} paid by piece rate to the output of Y_{ij} in a team-based incentive j ($j = 1$ for revenue sharing, $j = 2$ for individual tournament and $j = 3$ for team tournament); δ_j captures the observable incentive effect; α_i is the unobservable incentive-invariant individual effect, ε_{ij} denotes the residual error.

In addition, considering the time effects on individual performance during Stage 2, we create dummy variables to identify the order effect which is independent of the incentive. So we have the second fixed-effect model as follows:

$$\dot{Y}_{ij} = \delta_j + \beta' D_{ij} + \alpha_i + \varepsilon_{ij} \tag{3.2}$$

where $D_{ij} = (1, 0, 0)$ if the incentive j is the 1st task in Stage 2; $D_{ij} = (0, 1, 0)$ if the incentive j is the 2nd task in Stage 2; and $D_{ij} = (0, 0, 1)$ if the incentive j is the 3rd task in Stage 2.

Alternatively, instead of fixed-effect model, we define the individual effect as $\alpha_i = \gamma' x_i$, so we have the following specification model:

$$\dot{Y}_{ij} = \delta_j + \beta' D_{ij} + \gamma' x_i + \varepsilon_{ij} \tag{3.3}$$

where x_i is the vector of individual attributes, including individual ability, overconfidence, risk attitude, social preference, gender, only child and GPA score. All the explanatory variables regarding individual characteristics are described in Table 3.4.

The regression results are reported in Table 3.5. The key finding is that the individual performance is highest under the most competitive incentive, i.e. individual tournament. The regression coefficients in the first three columns for "No Feedback" treatment show a strong difference between the individual tournament and other two team-based incentives: as compared with the baseline performance under piece rate, only the incentive of individual tournament has a significantly positive effect to improve performance. Meanwhile, the coefficients in the last three columns indicate that providing more information about relative performance does not improve individual performance under team-based incentives. Instead, the feedback makes the performance differences between three team-based incentives close to zero, which means the significant advantage of individual tournament disappears once introducing the feedback information.

Table 3.4 Individual characteristics related to performance under different incentives

Variable	Range	Description	Mean
Individual ability	$\{5, 6, \ldots, 21\}$	Baseline performance in puzzle-solving task paid by piece rate.	12.36 (2.71)
Confidence bias	[-100%, 100%]	Difference between self-assessment and actual relative performance.	0.16 (0.27)
Risk aversion	$\{10\%, 20\%, \ldots, 100\%\}$	Minimum probability of winning \$40 that makes lottery is preferred to fixed \$20.	0.65 (0.15)
Inequality aversion	$\{0, 1\}$	=1 if the dictator equally shares payoff with the receiver in dictator game.	0.51 (0.50)
Female	$\{0, 1\}$	=1 if a female; and =0 if a male.	0.55 (0.50)
Only child	$\{0, 1\}$	=1 if an only child; and =0 if a child with siblings.	0.52 (0.50)
GPA	[0, 4]	Grade-Point-Average is the most common method to evaluate academic performance.	3.30 (0.35)

Note:
The first four variables are elicited by experimental methods during stage 1;
The rest three variables are elicited by self-report questionnaire during follow-up stage;
Standard deviations in parentheses.

The only difference between Model 1 and 2 is that we added the order effects of three tasks during Stage 2. Consistent with the findings in Table 3.3, the coefficients for those time dummy variables verifies an increase trend at beginning and then followed with a reduction on performance. However, the increase is significant in either "No Feedback" or "Feedback" case, while the reduction is only significant in the case with feedback information.

To check the robustness of the results in fixed-effect model(s), we also examine other specifications. Since our experiment has used a rich battery of diagnostic games to elicit several non-cognitive characteristics in stage 1 and also collected more individual information from the follow-up questionnaire, it is quite convenient for us to replace the unobservable value of individual-effect with a multi-dimensional vector of elicited individual attributes. When we run the GLS regression model by adding several individual characteristics, there is no impact on all other existing coefficients. So, compared with the outcome of the fixed-effect model, the incentive effects as well as the time effects remain the same. Besides, we find the performance differences due to heterogeneity in personal characteristics: in the absence of feedback information, those risk-loving and over-confident subjects tend to make less effort responding to team-based incentives, whereas the effects of both risk attitude and self-confidence on individual performance become insignificant after subjects receive feedback about their relative performance; on the other hand, thanks to the provision of feedback information, the female subjects are more likely to improve their performance than males under team-based incentives, although the performance difference between them is not significant without feedback.

Furthermore, in order to eliminate the time effects on task performance, we run the regressions by using the subsample of Task 5 only. Since Task 5 is the first task in Stage 2, and the payment scheme used in this task is possibly any one out of those three team-based incentives. The results are reported in Table 3.6. It turns out that all standard errors are increased because of the reduced sample size. But we can find that values of coefficients and their significance levels are quite similar to the regression in Table 3.5. We hence keep our previous conclusion through this robustness check of subsample regression.

At the beginning of Stage 3, the sign-up bonus policy was introduced into our experimental labor market. As a consequence, the outcomes in Table 2 and 3 have demonstrated the remark-

Table 3.5 Regression for performance improvement during three tasks of Stage 2

Performance improvement	No Feedback			Feedback		
($\dot{Y}_{ij} = Y_{ij} - Y_{i0}$)	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
RS (δ_1)	–	–	–	–	–	–
IT ($\delta_2 - \delta_1$)	0.368**	0.352**	0.352**	0.094	0.162	0.162
	(0.182)	(0.159)	(0.159)	(0.196)	(0.175)	(0.175)
TT ($\delta_3 - \delta_1$)	0.268	0.198	0.198	0.025	0.024	0.024
	(0.182)	(0.159)	(0.159)	(0.196)	(0.175)	(0.175)
Orderdummy1		–	–		–	–
Orderdummy2		1.491***	1.491***		1.277***	1.277***
		(0.159)	(0.159)		(0.175)	(0.175)
Orderdummy3		-0.164	-0.164		-0.446**	-0.446**
		(0.159)	(0.159)		(0.175)	(0.175)
Individual ability (Y_{i0})			-0.510***			-0.505***
			(0.066)			(0.055)
Confidence bias			-0.011*			0.001
			(0.006)			(0.006)
Risk aversion			0.014*			0.008
			(0.008)			(0.008)
Inequality aversion			0.144			-0.183
			(0.249)			(0.233)
Female			0.026			0.406*
			(0.255)			(0.245)
Only child			0.230			0.359
			(0.249)			(0.236)
GPA score			0.130			0.208
			(0.351)			(0.358)
Constant	0.086	-0.328**	4.702***	0.411***	0.111	4.769***
	(0.129)	(0.145)	(1.433)	(0.138)	(0.158)	(1.405)
Observations	627	627	627	606	606	606
R-squared	0.010	0.249	0.282	0.010	0.208	0.356

Note: Standard errors in parentheses; $***p < 0.01, **p < 0.05, *p < 0.1$.

Table 3.6 Regression for performance improvement in first task of Stage 2

Performance improvement ($\dot{Y}_{ij} = Y_{ij} - Y_{i0}$)	No Feedback		Feedback	
RS (δ_1)	–	–	–	–
IT ($\delta_2 - \delta_1$)	0.909**	1.055***	0.475	0.671
	(0.395)	(0.337)	(0.394)	(0.311)
TT ($\delta_3 - \delta_1$)	0.234	0.29	-0.299	-0.050
	(0.399)	(0.341)	(0.425)	(0.335)
Individual ability Y_{i0}		-0.605***		-0.568***
		(0.076)		(0.062)
Confidence bias		-0.013*		-0.002
		(0.007)		(0.006)
Risk aversion		0.014		0.009
		(0.009)		(0.009)
Inequality aversion		0.130		-0.137
		(0.285)		(0.266)
Female		0.154		0.584**
		(0.295)		(0.277)
Only child		0.157		0.504*
		(0.287)		(0.268)
GPA score		-0.177		-0.180
		(0.405)		(0.406)
Constant	-0.529*	6.601***	0.085	6.467***
	(0.280)	(1.632)	(0.282)	(1.600)
Observations	209	209	202	202
R-squared	0.027	0.342	0.018	0.432

Note: Standard errors in parentheses; $***p < 0.01, **p < 0.05, *p < 0.1$.

able increase in performance after receiving an extra compensation. Although the additional bonus varied across subjects because the sign-up bonus is a random integer between 0 and 10, subjects performed better on average with a bonus. Moreover, by the specification in the following model, we can clearly identify the marginal effect of one Chinese Yuan.

$$\ddot{Y}_{ij} = \theta_0 + \theta_y \cdot Y_{ij} + \theta_b \cdot b_i + \epsilon_{ij} \tag{3.4}$$

where $\ddot{Y}_{ij} = Y'_{ij} - Y_{ij}$ represents the change in individual i's performance from first-round output Y_{ij} paid by incentive j without extra bonus to the second-round output of Y'_{ij} in the same incentive j but with a signup bonus b_i; ϵ_{ij} is the error term.

Table 3.7 shows the estimated regression coefficient and standard errors form this analysis on sign-up bonus. Irrespective of a particular incentive and/or feedback information, the sign-up bonus has a positive and highly significant effect on performance improvement in the final task. More specifically, an extra compensation of 10 Chinese Yuan could stimulate subjects to complete 0.8 puzzles more in "No Feedback" treatment, and 1 puzzle more in "Feedback" treatment. Furthermore, to probe the robustness of this result, we added incentive dummies to the regression. The impact of sign-up bonus is still positive in either "No Feedback" case or "Feedback" case, although the coefficient becomes smaller when controlling each specific incentive effect on performance.

Although the neoclassical economics suggests that the prepaid compensation in terms of sign-up bonus should have no effect on individual performance afterwards, recent discoveries in behavioral economics have led scholars to question this theoretical prediction. We use insights gained from one of the most influential lines of behavioral research, i.e. "gift exchange"[5], in an attempt to explain the performance improvement in response to extra sign-up bonus. In fact, a vast body of literature has stressed the importance of gift exchange for mitigating moral-hazard problems of incomplete contracts: since employees repay a gift in the form of higher wages by providing higher efforts, employers can expect the productivity-enhancing outcome by adopting the "gift" treatment. For instance, in Gneezy and List (2006), the authors conducted two field

[5] The notion of gift exchange was first proposed by Adams (1963), who posited that in social exchange between two agents the ratio of the perceived value of the inputs (e.g., wage) to the perceived value of outputs (e.g., resulting from the employee's effort) would be equivalent.

Table 3.7 Regression for performance improvement in final task with sign-up bonus

Performance Improvement ($\ddot{Y}_{ij} = Y'_{ij} - Y_{ij}$)	No Feedback		Feedback	
Sign-up bonus (b_i)	0.080*	0.079*	0.104**	0.083*
	(0.047)	(0.047)	(0.050)	(0.050)
Baseline performance (Y_{ij})	-0.468***	-0.471***	-0.412***	-0.449***
	(0.053)	(0.054)	(0.057)	(0.058)
PR dummy		–		–
RS dummy		-1.074**		-0.538
		(0.467)		(0.413)
IT dummy		-0.44		0.912**
		(0.379)		(0.425)
TT dummy		-0.253		-0.064
		(0.407)		(0.448)
Constant	7.350***	7.744***	6.431***	6.929***
	(0.712)	(0.729)	(0.779)	(0.782)
Observations	209	209	202	202
R-squared	0.279	0.298	0.216	0.258

Note: Standard errors in parentheses; $***p < 0.01, **p < 0.05, *p < 0.1$.

experiments to examine the effects of gift exchange. As a result, their field evidence suggests that worker effort in the first few hours on the job is considerably higher in the "gift" treatment than in the "nongift" treatment. After the initial few hours, however, no difference in outcomes is observed, and overall the gift treatment yielded inferior aggregate outcomes for the employer.

In our experiment, compared with "nongift" treatment without extra sign-up bonus, the individual perforamance in the "gift" treatment with prepaid bonus successfully induce employees to work harder. However, the limitation of this experimental design is we can only prove the effectiveness of sign-up bonus policy in the short run; after an immediate positive reaction to this incentive, the higher productivity in response to a gift of sign-up bonus might diminish over time. Hence, in the future work, it is important for us to verify the duration of sign-up bonus effect on individual performance to generalize the results of this study in field experiments.

3.5 Conclusions

So far, from our experimental labor market in the well-controlled laboratory environment, we have found the answers for three questions addressed at the very beginning.

Consistent with the findings in previous related studies, we also see that competitive incentive is more effective than corporative incentive to stimulate workers' performance. Meanwhile, the provision of feedback about relative performance reduces the performance differences between different team-based incentives, especially lessens the advantage of competitiveness on performance improvement.

More importantly, we have a surprising result from a quite simple compensation method, that is, an additional prepaid sign-up bonus can raise workers' productivity remarkably. Our data also suggest that the effect of sign-up bonus dominates both incentive effect and information effect. In other word, the negative influence on productivity from a certain incentive and/or feedback information can be easily offset by a higher sign-up bonus.

In practice, companies always strive for the efficient way to encourage their employees work hard. Because we have shown that some non-cognitive characteristics such as risk attitude and self-confidence influence individual performance in a systematic way, then subjects of similar ability might respond quite differently to the same incentive and feedback information. Therefore, compared with designing a distinctive incentive scheme and feedback treatment for each heterogeneous worker, it's much more feasible and convenient for companies to set a flexible sign-up bonus policy while holding other elements of workplace environment fixed.

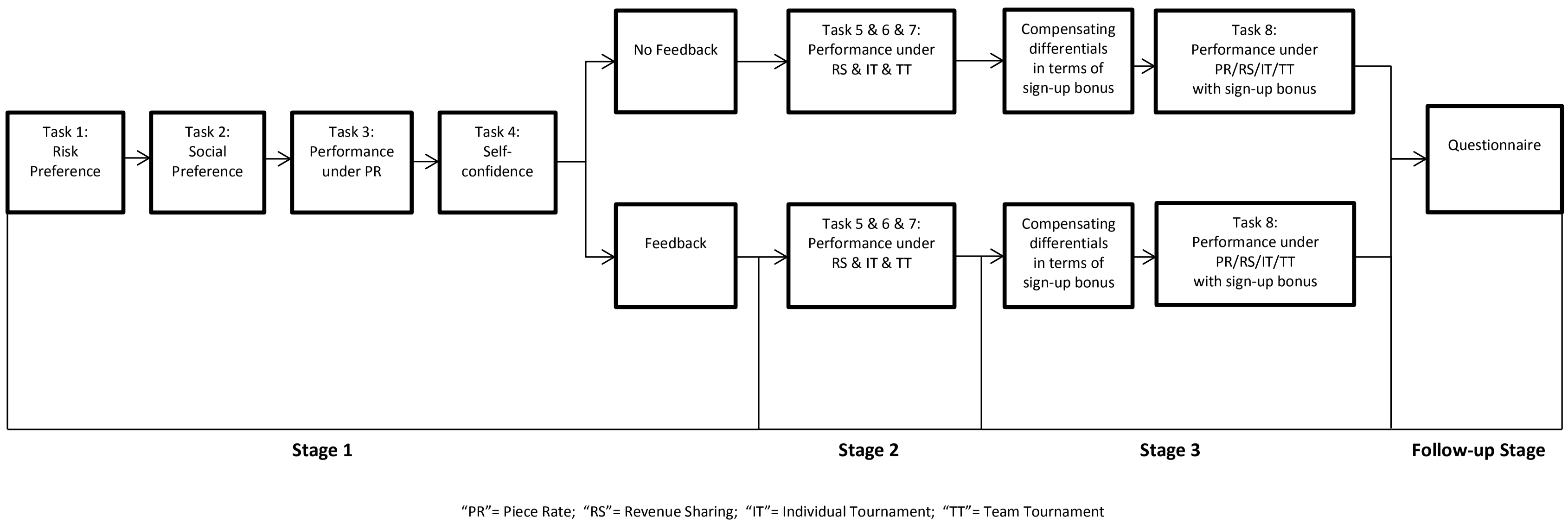


Figure 3.2 Procedure of the laboratory labor market experiment

CHAPTER 4. HOW TO ATTRACT TALENT? STATED PREFERENCES FOR DIFFERENT INCENTIVES FROM RANKING AND COMPENSATING DIFFERENTIAL

Abstract

In order to study whether and how people with heterogonous individual characteristics self-select into various incentives differently, we design a lab experiment where "workers" can join "companies" that pay them by different incentive schemes: (a) piece rate; (b) revenue sharing; (c) individual tournament; and (d) team tournament. An innovation in this experiment is that we ask subjects to not only rank those four incentives, but also elicit compensating differentials between any two consecutive incentives, which allows us to price the cost of a mismatch between a worker and a company. Lab data from our sample of 411 Chinese university students shows the following: (1) Subjects with high relative performance always prefer individual tournament to other two team-based incentives. (2) Risk-averse subjects are less likely to choose individual tournament if knowing the information about their relative performance. (3) Cooperative incentives attract more women than men, which is partially explained by gender-specific social preferences. (4) Compared to children with siblings, only children are less cooperative but more competitive. (5) In the absence of feedback, overconfident subjects are more likely to enter into individual tournament than those under-confident subjects with the same ability. Interestingly, the provision of information about their relative performance eliminates the impact of biased self-assessment. As a result, the feedback helps reduce the gender gap in competition as well as the difference between only child and child with siblings.

4.1 Introduction

Now, more than ever, companies depend on talent. But talented employees are becoming a rare commodity in the labor market, so company's success depends crucially on attracting and retaining the right people in place by proper incentives. In fact, numerous companies seek to win the so-called "war for talent" by continually trying more effective compensation schemes. For example, in January 2011, Dan Finnigan, who is the CEO of Jobvite,[1] wrote an article entitled "how to win the war for the most talented employees", in which he stated flatly: "Nowhere is this clearer than in the San Francisco Bay and other pockets of innovation across the country like Austin, Texas. Last month, the news was dominated by the story that Google offered 10% pay increases and holiday bonuses of $1,000 for every employee. Google's arming up for this targeted' war for talent as new startups continually try to poach talent from top companies– taking a page out of Google's playbook from five years ago" (http://mashable.com/2011/01/03/find-talented-employees/). Hence, it is worthwhile for us to explore which incentive is more attractive to talent in our experimental labor market.

The modern workplace puts an ever greater emphasis on teamwork. Therefore, in addition to the baseline incentive, i.e. piece-rate compensation scheme for individual-based working environment, the following three representative team-based incentives are designed to highlight the tradeoff between cooperation and competition: (1) revenue sharing where all earnings are equally shared among team members; (2) individual tournament where only the best performer in the team can earn a prize; and (3) team tournament where teams should compete first, and then the reward is shared within the winning team. In particular, the last team-based incentive, i.e. "team tournament", is defined as a hybrid of "individual tournament" and "team revenue-sharing", which are two contrasting "corporate cultures" one that focuses on competition and one that emphasizes collaboration.

According to our design, we elicit individual preferences for different incentives by a sorting procedure. An innovation in our design is that we ask subjects to not only rank those four payment schemes, that is: (1) piece rate, (2) revenue sharing, (3) individual tournament, and

[1] Jobvite is the recruiting platform for the social web that companies use to find and hire people.

(4) team tournament, from most preferred to least preferred, but also elicit their compensating differentials between any two consecutive incentives. To the best of our knowledge, this paper is the first one to implement such a mechanism in a laboratory experiment to price the cost of a mismatch between a worker and a company.

Furthermore, since we allow the self-selection across various compensation schemes, it's possible that workers with different individual characteristics feel attracted by different incentives. In fact, recent labor market research has suggested that workers of similar ability might behave quite differently in the workplace. For example, some workers with high ability might not be willing to work in a team because of the "free-riding" problem. Some might dislike competition and relative performance pay because they find it stressful. Although many authors (Niederle and Vesterlund 2007; Wozniak, Harbaugh and Mayr 2010) have stressed the importance of non-cognitive characteristics such as risk preference and self-confidence on labor market outcomes, only very few studies (e.g. Dohmen and Falk 2011) use incentivized elicitation methods to measure them. Therefore, in our experiment, we use a rich battery of diagnostic games to elicit individual characteristics including risk attitude, social preference and self-confidence.

As a result, we clearly show the relationship between the self-selection and individual characteristics, and the main findings include: (1) More able subjects are more willing to accept individual tournament because they have higher probability of winning in the competition. (2) Risk-averse subjects are less likely to choose individual tournament than other incentives after they receive the information about their relative performance. (3) Subjects who strongly prefer fairness are more willing to accept cooperative incentives including team revenue-sharing and team-tournament. As a consequence, cooperation incentives attract more women than men, which is partially driven by gender-specific social preferences. (4) Only children participants of our sample in China behave significantly more competitive than children with siblings. By contrast, in the latest paper on China's One-Child policy (Cameron et al. 2013), the authors conclude that only children are less competitive than the counterpart. (5) We verify that the provision of feedback about relative performance can adjust individuals' biased self-beliefs and then influence their sorting in the end. Consistent with the findings in Wozniak, Harbaugh and Mayr (2010), the feedback about relative performance encourages under-confident subjects with high ability to enter into individual tournament, and moves overconfident subjects with low ability towards less competitive incentive schemes. In one word, the feedback information leads all subjects to more optimal choices and helps reduce the gender gap as well as the difference between only child and child with siblings in competitiveness.

More importantly, thanks to our design for the compensating differentials, we are able to answer how much the additional compensation is needed to motivate any particular type of worker to accept a less preferred incentive, rather than her/his favorite one. Compared with the contingent choice/ranking models, which use best choice or ranking data to predict the probability that subjects choose a particular incentive among various options, our data on compensating differentials are more informative so that it allows us to estimate personal evaluations for different incentives in dollars.

The rest of this chapter is organized as follows. Section 2 presents a literature review. Section 3 describes the experimental design. The results are given in section 4. Finally, in section 5, further discussions and some concluding remarks are provided.

4.2 Literature Review

This research is related to several strands of literature. First, our experiment is a study of incentives in a laboratory experiment, especially for the interaction between different compensation schemes and heterogeneous individuals. Due to the labor mobility, each worker has freedom to move from one company to another, and such a movement is referred to the consequence of self-sorting, i.e., the possibility that workers of similar ability feel attracted by different incentive schemes (Niederle and Vesterlund 2007; Wozniak et al. 2010; Dohmen and Falk 2011). For example, the experimental findings in Dohmen and Falk (2011) indicate that, in addition to observable personal productivity and gender, some unobservable individual characteristics such as relative self-assessment, willingness to take risks affect the sorting decision in a systematic way. Our design is somewhat similar to the approach used by Dohmen and Falk (2011), but also differentiates in several aspects. In particular, Dohmen and Falk (2011) didn't test the performance before self-sorting but only have the results of individual productivity when all subjects work under their most preferred incentives. On the contrary, in order to disentangle the exogenous incentive effect and the endogenous self-sorting effect on individual performance, we add a stage before self-selection that forces all subjects to initially complete several real-effort tasks paid by four different incentives one by one. As a result, we identify the subjects' preferences by controlling their productivity variances under different incentives.

Second, our experiment ties into the literature on social preferences. Many experimental economists have claimed that a satisfactory understanding of workings under competition and cooperation is impeded by self-interest hypothesis. They have gathered overwhelming evidence that systematically refutes the self-interest hypothesis and suggests that a substantial fraction of the people exhibit social preferences (Fehr and Falk 2002; Fehr et al. 2007; Englmaier and Wambach 2010). While extensive findings indicate that individuals take account of the effect of their actions on others during laboratory games, but whether and how individuals exhibit social preferences in the workplace is largely unknown. So, in this paper, we investigate how the social preferences shape the individual behavior under certain team-based incentives. In the first step, we apply a diagnostic task, i.e. a dictator game to elicit subjects' social preferences. Next, we observe individual preferences under three distinct team-based incentives as well as the baseline payment of piece rate. In light of our findings, the heterogeneous preference for different incentives is closely related to individual social preference. For instance, in our experiment, the fact of cooperation incentives attracting more women than men is partially explained by gender-specific social preferences, because majority of female subjects are fair type who like equal pay within the team.

Third, our experiment is related to the literature on mechanism design for reducing gender difference in competitiveness. With very few exceptions[2], many experiments report a significant gender difference in performance under individual tournament (Gneezy et al. 2003; Gneezy and Rustichini 2004; Paserman 2007), and also in the willingness to enter a tournament (Niederle and Vesterlund 2007; Gneezy et al. 2009). For the sake of equality in the labor market, it's worthwhile to explore approaches that narrow the gender gap. On one hand, by introducing the element of cooperation into competition, Dargnies (2009) as well as Healy and Pate (2011)

[2] In Zhang (2011), by using a sample of ethnically diverse middle and high school students in China, the Han Chinese exhibited no gender differences in competitive inclination.

show that the gender gap in competitive inclination is reduced a lot when the subjects compete in teams. So, in order to investigate the gender-specific preferences for competition and cooperation, we classify three team-based incentives as follows: the "team revenue-sharing" which denotes the wholly cooperative incentive; the "individual tournament" which implies the pure competition in the workplace; and the combination of cooperation and competition i.e. the "team tournament". Our findings verify that, by controlling the performance, females are more likely to choose two cooperative incentives rather than the individual tournament. Moreover, according to diagnostic results for non-cognitive characteristics, it suggests that such a gender difference in willingness to accept cooperation is partially driven by their different social preferences but not the similar risk preferences between males and females. On the other hand, Wozniak et al. (2010) find that, with lack of actual feedback information, high ability subjects are more reluctant to enter tournament than expected value maximization would require and this effect is larger for high ability females. By contrast, too many low ability types enter competitive environments, and this effect is larger for males. The main reason for such a gender gap is that males are substantially more overconfident about their relative performance than females. Therefore, the provision of relative performance feedback helps to eliminate the gender gap by correcting individuals' wrong beliefs. In this paper, we also examine the effect of feedback by providing a detailed distribution of all subjects' performance before the stage of self-selection. However, different from the method used in Wozniak et al. (2010), we do not apply a within-subjects comparison that identifies the treatment effect by eliciting their preference twice before and after feedback disclosure; instead, we design a between subjects comparison, that is, one half subjects receive feedback after they finish their productivity test, and the other half never receive any feedback about their relative performance during the experiment.

4.3 Experimental Design

4.3.1 Procedure of the Lab Experiment

In order to study how individual characteristics and performance feedback affect their preferences for different compensation schemes, we implemented an experiment which consists of 8 tasks and is divided into 3 stages: Stage 1 includes the first four tasks that are designed to elicit each subject's risk preference, social preference, productivity in a puzzle-solving game and self-assessment about the relative performance, respectively. After the first four diagnostic tasks complete, subjects were randomly assigned into two different treatment groups: half subjects were given detailed feedback about their relative performance in Task 3 of puzzle-solving game, and the other half never knew any relative performance feedback during the whole experiment. Next, the following Stage 2 consists of Task 5, 6 and 7: by those three tasks, we observed individual performances in similar puzzle-solving games but paid by three different team-based incentives, i.e. revenue sharing, individual tournament and team tournament. And then, in Stage 3, based on individual stated-preferences for different incentives, all subjects completed the last puzzle-solving game i.e. Task 8 which was paid by a certain incentive plus an additional compensation in terms of prepaid sign-up bonus. More specifically, before starting Task 8, we elicited participants' preferences for given incentives through a self-sorting process. In particular, we asked subjects to not only rank those four payment schemes, that is: (1) piece rate, (2) revenue sharing, (3) individual tournament, and (4) team tournament, from most preferred to least preferred, but also state their compensating differentials between any two consecutive incentives. And then, the system would automatically determine an incentive scheme as well as an contingent compensation of extra sign-up bonus for each subject.

Figure 4.1 illustrates the sorting method we used to elicit subjects' preferences for different incentives. "team tournament company" adopts the incentive of "team tournament", "single contractor" adopts the incentive of "piece rate", "tournament company" adopts the incentive of "individual tournament"; and "team payment company" adopts the incentive of "revenue sharing".

If someone chose to be paid by piece rate scheme, then s/he would work as "single contractor" instead of working as an employee of a company. On the other hand, if someone chose to work for one of those three companies adopting team-based incentives described in Stage 2, her/his co-workers would be randomly selected among the players in the lab room; however, the performances of other three co-workers in the same company were their past performance when they worked for this company in a previous round of Stage 2. Using others' past performances

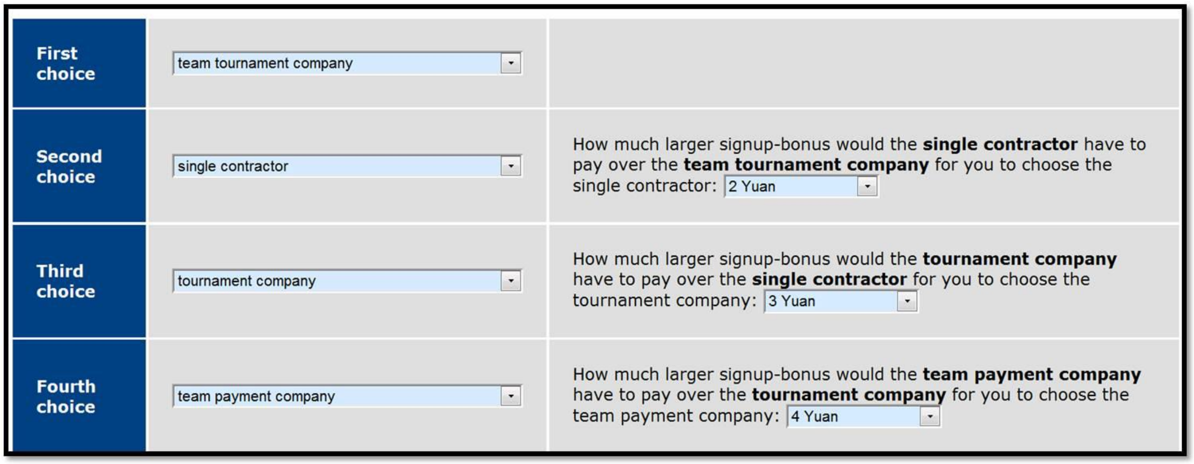


Figure 4.1 Sorting procedure to elicit individual preferences for different incentives

has several advantages: it avoids a potential source of error through biased beliefs about other participants' choices; furthermore, the choice of team pay does not affect the payment of any other participant, so it can rule out the possibility that someone may reluctant to choose a team-based incentive which imposes an externality on others.

In addition to asking subjects rank four incentives from most preferred to least preferred, we also elicited compensating differentials between any two consecutive options. We designed an incentive-compatible mechanism to elicit individuals' compensation demand to accept a less preferred option, and then we were able to use those valuations to price the cost of a mismatch between a company and a worker.

A variant of the Becker-DeGroot-Marschak (BDM) mechanism was used to encourage subjects tell truth when they answered those compensating differential questions. In fact, we introduced a "sign-up bonus" scheme as follows: The system automatically generates four random integers between 0 and 10, and then assigns each incentive a number as its prepaid signup-bonus. If the worker is matched with a particular incentive for Task 8, then her/his payoff in this final task will be the payment calculated by the incentive scheme plus the corresponding sign-up bonus.

Based on subjects' ranking and compensating differentials, our computer system applies the following rules to determine the matched incentive and sign-up bonus for each worker through three rounds of pairwise comparison between consecutively ranked incentives: In the first round of pairwise comparison, the system compares the two signup-bonuses of worker's first best choice and the second best, if the difference is not less than the minimum of the individual required compensating differential, then the system will remain the second best choice and delete the first one; otherwise, the system will keep the first but delete the second choice. Next, in the second round of pairwise comparison, the system continues to compare the signup-bonuses of the third best choice and the "survival" incentive from the first round comparison. Again, if the difference is not less than the minimum of the required compensating differential between those two incentives, the system will keep the third best choice rather than the other one; otherwise, the system will remain the "survival" incentive from the first round comparison. Finally, the system ends with the third round of pairwise comparison. Similar to the previous rounds, compared with the sign-up bonus associated with the dominant incentive in the second round of comparison, if the bonus difference between the fourth choice and its counterpart is not less than the minimum of individual required compensating differential, then the system will send worker to her/his fourth choice; otherwise, the system will send worker to the "survival" incentive from the second round of comparison. Therefore, by a sequence of three rounds pairwise comparison, according to each worker's stated compensating differentials for different incentives, the system is able to determine the best "compensation package" i.e. a certain incentive scheme with a prepaid sign-up bonus for her/him. All in all, this rule can be regarded as a variant of BDM method, in which the dominant strategy of each subject is "telling truth", and people will suffer a loss if they wrongly self-report their compensating differentials.

We then take an example to explain how this mechanism works. Given someone's rankings and compensating differentials as follows: team tournament $\succ$ piece rate (2 Chinese Yuan) $\succ$ individual tournament (3 Chinese Yuan) $\succ$ revenue sharing (4 Chinese Yuan), the numbers in parentheses denote extra bonus required to switch from more preferred option to less preferred option. Meanwhile, suppose the system has randomly assigned four integers between 0 and 10 to those four incentives, respectively: team tournament with 4 Chinese Yuan, piece rate with 3 Chinese Yuan, individual tournament with 9 Chinese Yuan, and revenue sharing with 5 Chinese

Yuan. So, in the first round of pairwise comparison, the difference of sign-up bonus between first best choice "team tournament" and second choice "piece rate" is equal to -1 ($= 3 - 4$), which is smaller than the minimum of required additional compensation (i.e. 2 Chinese Yuan). As a result, "team tournament" survives in the first round. And then, the system continues to compare "team tournament" with the third choice "individual tournament". The bonus difference between those two incentives is 5 ($= 9 - 4$), and it just achieves the minimum of required compensating differential that makes the subject would like to accept the less preferred incentive of individual tournament rather than the more preferred choice of team tournament (i.e. $2 + 3 = 5$ Chinese Yuan). Therefore, the system will keep "individual tournament" and delete "team tournament" in the second round of comparison. In the end, comparing the sign-up bonus of "individual tournament" and the fourth choice of "revenue sharing", we have the difference that equals -4 ($= 5 - 9$), and this value is definitely less than the minimum of compensating differentials between those two incentives (i.e. 4 Chinese Yuan). So, the incentive of "individual tournament" survives in the last round. As a consequence, the system will inform this subject that s/he will be paid by "individual tournament" with a sign-up bonus of 9 Chinese Yuan before s/he starts the last puzzle-solving game in Task 8.

Finally, when all eight tasks ends, a simple questionnaire was used to gather personal demographic data including gender, age, marital status, major fields of study, grades, etc.

4.3.2 Subject Pool

The experiment was implemented at Central University of Finance and Economics in Beijing, China. We had 411 complete observations, including 184 males and 227 females, and all of them were university students. In addition, there were 213 only children (and 198 children with siblings) in our subject pool. Because majority of the participants i.e. 70% were graduate students, the average college year was above 4 and the average age was 23.03 correspondingly. Only 2% of the participants were married. Finally, the average level of GPA was 3.3 (on a 4.0 scale) with a standard deviation of 0.35. However, since subjects' grades were self-reported, we cannot guarantee the accuracy of this information and treat it as a quite limited index to personal academic performance.

A total of 21 experiment sessions took place in the same computer lab at the university, and the average size of those sessions was 19.57 participants (with a standard deviation of 6.62). Each session lasted, on average about 50 minutes. Final payments for subjects were based on one randomly selected task out of eight tasks, plus a fixed show-up fee of 10 Yuan. The average payment was 30.85 Yuan.

4.4 Results

4.4.1 Ranking for Different Incentives

According to the results in self-sorting, the rank distribution across four incentives is summarized in Table4.1. From the outcomes in this table, it's clear that the scheme of piece rate is the most popular incentive by majority. Actually, it's rational for people to select piece rate that avoids the uncertainty and risk of team pay. Meanwhile, the provision of feedback has little effect on individual preference for the piece rate because it is an individual-based incentive which is irrelevant to any relative performance. However, once the information about the relative performance is released, the individual stated preferences for other three team-based incentives did change significantly. In particular, there is a remarkable increase in the percentage of subjects chose revenue sharing as the favorite while a dramatic reduction in the percentage chose individual tournament. One reasonable explanation for this change is that providing feedback about relative performance encourages under-confident subjects with high ability to enter into individual tournament, and moves overconfident subjects with low ability towards less competitive incentive scheme.

Model Specification of Ranking Data

We use regression analysis to probe the ranking diversity under different incentives driven by individual heterogeneity, so we develop a rank-ordered logit model with full ranking data to derive the specification as follows:

$$U_{ij} = \delta_j + \beta \cdot Y_{ij} + \gamma_j' x_i + \varepsilon_{ij} \tag{4.1}$$

An individual, labeled i, faces a choice among four alternative incentives. The decision maker would obtain a certain level of utility from each incentive, so the utility that individual i

Table 4.1 Ranking for four incentives

	No Feedback				Feedback			
	Rank 1	Rank 2	Rank 3	Rank 4	Rank 1	Rank 2	Rank 3	Rank 4
Piece Rate	33%	36%	22%	9%	38%	29%	28%	5%
Revenue Sharing	11%	22%	27%	40%	20%	29%	19%	32%
Individual Tournament	36%	20%	19%	25%	25%	18%	16%	41%
Team Tournament	20%	22%	32%	26%	17%	24%	37%	22%

Note: "Rank 1" is the most preferred choice, and "Rank 4" is the least preferred one. The value in each cell denotes the proportion of subjects.

obtains from incentive j is U_{ij}, where $j = 0$ denotes piece rate, $j = 1$ denotes revenue sharing, $j = 2$ denotes individual tournament, and $j = 3$ denotes team tournament. In addition, δ_j denotes the incentive-specific constant, Y_{ij} represents individual i's performance paid by incentive j, x_i is the vector of individual attributes including overconfidence, risk attitude, social preference, etc. In the end, ε_{ij} denotes the error term.

After knowing individuals' ranking results, we are able to investigate the relative effect of differing explanatory variables (including incentive and individual attributes) on the different ranking outcomes.

In particular, the probability that individual i prefer to choose incentive j rather than incentive k is:

$$Pr(j \succ k) = Pr(U_{ij} > U_{ik}) = Pr(\varepsilon_{ik} - \varepsilon_{ij} < (\delta_j - \delta_k) + \beta \cdot (Y_{ij} - Y_{ik}) + (\gamma_j - \gamma_k)' x_i) \quad (4.2)$$

Since there are at most three incentive-specific values can be identified with four alternatives in total, here we treat $j = 0$ i.e. piece rate as the baseline and then make the baseline values normalized to zero. Hence,

$$Pr(j \succ k) = Pr(\tilde{\varepsilon}_{ik} - \tilde{\varepsilon}_{ij} < (\tilde{\delta}_j - \tilde{\delta}_k) + \beta \cdot (\tilde{Y}_{ij} - \tilde{Y}_{ik}) + (\tilde{\gamma}_j - \tilde{\gamma}_k)' x_i \quad (4.3)$$

where $\tilde{\varepsilon}_{ij} = \varepsilon_{ij} - \varepsilon_{i0}; \tilde{\delta}_j = \delta_j - \delta_0; \tilde{Y}_{ij} = Y_{ij} - Y_{i0}$ and $\tilde{\gamma}_j = \gamma_j - \gamma_j$.

Furthermore, we use $r_i = (r_i^1, r_i^2, r_i^3, r_i^4)$ to index the individual i's ranking. For example, if individual i ranks the four incentive from most preferred to least preferred as follows: team tournament $\succ$ piece rate $\succ$ individual tournament $\succ$ revenue sharing, then we have $r_i^1 = 3, r_i^2 = 0, r_i^3 = 2, r_i^4 = 1$.

Hence, given a certain ranking vector r_i, the probability that individual i prefers incentive r_i^n to incentive $r_i^n, r_i^{n+1}, \ldots, r_i^4$ is equal to:

$$F_{r_i^n} = Pr(r_i^n \succ r_i^{n+1}, r_i^n \succ r_i^{n+2}, \ldots, r_i^n \succ r_i^4) = Pr(U_{ir_i^n} > U_{ir_i^{n+1}}, U_{ir_i^n} > U_{ir_i^{n+2}}, \ldots, U_{ir_i^n} > U_{ir_i^4}) \tag{4.4}$$

Moreover, we define the probability of observing i's ranking r_i as follows:

$$F_{r_i} = Pr(r_i^1 \succ r_i^2 \succ r_i^3 \succ r_i^4) = Pr(U_{ir_i^1} > U_{ir_i^2} > U_{ir_i^3} > U_{ir_i^4}) \tag{4.5}$$

Suppose individual preferences satisfy the independence of irrelevant alternative (IIA) property, and it implies that the conditional distribution of the utility of each choice is independent of the ranking of other choices. As a result, the probability assigned to a specific ranking of choices is equal to the product of standard logit probabilities, i.e.

$$F_{r_i} = \prod_{n=1}^{3} F_{r_i^n} \tag{4.6}$$

According to the specification model of rank-ordered logit model, we run the regression. First, all explanatory variables included in the vector of individual attributes x_i are described in Table 4.2. And then, the estimated coefficients and standard errors are reported in Table 4.3.

The coefficients returned from a logistic regression model in Table 4.3 are log-odds ratios. Since we have treated the incentive of piece rate as our baseline, the values of those coefficients tell us, with a one-unit change in the independent variable, how the log-odds of a certain team-based incentive being preferred to piece rate will change in response. Increasing the log-odds means increasing the probability that someone would like to choose the team-based incentive rather than the baseline of piece rate, and vice-versa decreasing the log-odds means decreasing such a probability. Therefore, the sign of a coefficient indicates the direction of its relationship between an independent variable and the probability of the certain team-based incentive.

Firstly, according to the regression results in the first column, in this simple analysis which only focuses on the incentive-specific effect but ignores the influences of heterogeneous individual attributes on self-sorting, we find that the probability of a team-based incentive being

Table 4.2 Individual characteristics related to preferences for different incentives

Variable	Range	Description	Mean
Relative performance	[0%, 100%]	Percentile rank for individual performance in puzzle-solving task paid by piece rate.	12.36 (2.71)
Confidence bias	[-100%, 100%]	Difference between self-assessment and actual relative performance.	0.16 (0.27)
Risk aversion	$\{10\%, 20\%, \ldots, 100\%\}$	Minimum probability of winning $40 that makes lottery is preferred to fixed $20.	0.65 (0.15)
Inequality aversion	$\{0, 1\}$	=1 if the dictator equally shares payoff with the receiver in dictator game.	0.51 (0.50)
Female	$\{0, 1\}$	=1 if a female; and =0 if a male.	0.55 (0.50)
Only child	$\{0, 1\}$	=1 if an only child; and =0 if a child with siblings.	0.52 (0.50)
GPA	[0, 4]	Grade-Point-Average is the most common method to evaluate academic performance.	3.30 (0.35)

Note:
The first four variables are elicited by experimental methods during stage 1;
The rest three variables are elicited by self-report questionnaire during follow-up stage;
Standard deviations in parentheses.

preferred to the piece rate scheme will decrease if the incentive is revenue sharing because it has a negative coefficient -0.883. Moreover, as compared with the revenue sharing, the probability that someone prefers individual tournament to piece rate will be higher because the differential intercept coefficient for the incentive "individual tournament" is positive (0.529). Similarly, the positive of differential intercept coefficient for the incentive "team tournament" (i.e. 0.261) also implies team tournament is more likely to be preferred than the revenue sharing during the self-sorting process, but it's still less preferred than the incentive of individual tournament since its coefficient is significantly lower at $p=$0.035. On the other hand, the provision of feedback information reduce the differences between those three team-based incentives because the two differential intercept coefficient become smaller as well as statistically insignificant different from zero. Meanwhile, compared with the baseline option of piece rate, any team-based incentive is less likely to be chosen because of the negative coefficient -0.684. Hence, the regression results are consistent with the ranking data shown in Table 4.1.

Secondly, we add more individual characteristics into the regression model. As a result, we verified that some non-cognitive characteristics such as overconfidence, risk aversion and social preference affect individual self-sorting in a systematic way. Overall, most of the coefficients shown in Table 4.3 have the expected signs: (1) Regardless of receiving the feedback information or not, able subjects with high relative performance are always more likely to choose individual tournament than other two less competitive incentives of team pay. (2) Over-confident subjects would more likely to select individual tournament because they wrongly believe that they have high probability to win in the competition, but the provision of feedback can effectively correct their biased beliefs so that both overconfident and under-confident people will select the proper incentive irrespective of their initial self-confidence levels. (3) Providing the information about relative performance helps reduce the uncertainty in the team work, so the differences of final payoff between those three team-based incentives can be clarified by workers. As a result, individual risk attitude as well as social preference has a significant impact on self-sorting. The more risk averse the worker is, the less likely s/he would prefer individual tournament and/or team tournament due to the risk of failure in competition. Furthermore, because both individual and team tournaments generate unequal outcomes, then subjects having inequality aversion derive disutility from unequal outcomes so that they are definitely less likely to accept those competitive incentives; on the other hand, the revenue sharing scheme is more attractive to the inequality averse subjects than to those who don't care about the fairness. (4) Regarding the gender difference, compared with men, women is more likely to accept the cooperation in team revenue whereas shy away from individual competition without feedback. However, more information about relative performance successfully reduces the gender gap. (5) Only children and children with siblings state quite different preferences for those team-based incentives. In the absence of feedback, only children are less likely to choose cooperation in revenue sharing but more likely to prefer competition under individual tournament. In China, only children and children with siblings grow up in considerably different social and economic environments, so they might have distinctive personalities that differentiate their behaviors in the labor market. Although we have already controlled the risk aversion and social preference, there still exists some unobservable characteristics make those two groups different. Fortunately, the feedback mechanism allows us to lessen the gap between only children and children with siblings in our experiment. So, it indicates the difference should be caused by the uncertainty of team pay, and then the provision of symmetric information can lead two groups behave the same in the end.

4.4.2 Compensating Differentials Between Team-based Incentives and Piece Rate

As we have mentioned, the coefficients from ranking model are not informative enough, because the scale of the utility function cannot be identified. Hence, by ranking data alone, it is impossible to estimate the individual valuations for different incentives in terms of money. And then, thanks to a novel contingent evaluation method used in our experiment, it is convenient for us to measure the utility difference between any two options in a monetary unit. According to the survey design, by eliciting the differentials of sign-up bonus between each two consecutively ranked incentives, we are able to figure out the compensating differential between any two incentives in pair. For example, if someone stated her/his rankings and additional compensations as follows: team tournament $\succ$ piece rate (5 Chinese Yuan) $\succ$ individual tournament (2 Chinese Yuan) $\succ$ revenue sharing (3 Chinese Yuan), the numbers in parentheses denote extra

Table 4.3 Rank-ordered logit model

	No Feedback		Feedback	
Revenue Sharing	-0.883***	-1.603	-0.684***	-2.644
($\tilde{\delta}_1$)	(0.126)	(1.410)	(0.126)	(1.536)
Individual Tournament	0.529***	-1.794	-0.198	-0.292
($\tilde{\delta}_2 - \tilde{\delta}_1$)	(0.130)	(1.440)	(0.134)	(1.612)
Team Tournament	0.261**	-0.808	0.036	1.249
($\tilde{\delta}_3 - \tilde{\delta}_1$)	(0.124)	(1.359)	(0.124)	(1.502)
Performance difference		0.068**		0.041
(β)		(0.032)		(0.031)
Relative performance		0.007		-0.016**
($\tilde{\gamma}_1^{relative}$)		(0.007)		(0.007)
Relative performance*IT		0.029***		0.034***
($\tilde{\gamma}_2^{relative} - \tilde{\gamma}_1^{relative}$)		(0.008)		(0.008)
Relative performance*TT		-0.001		0.003
($\tilde{\gamma}_3^{relative} - \tilde{\gamma}_1^{relative}$)		-0.007		-0.007
Confidence bias		0.006		-0.004
($\tilde{\gamma}_1^{confidence}$)		(0.008)		(0.008)
Confidence bias*IT		0.024***		-0.006
($\tilde{\gamma}_2^{confidence} - \tilde{\gamma}_1^{confidence}$)		(0.009)		(0.009)
Confidence bias*TT		-0.007		-0.005
($\tilde{\gamma}_3^{confidence} - \tilde{\gamma}_1^{confidence}$)		(0.008)		(0.008)
Risk aversion		-0.013		-0.004
($\tilde{\gamma}_1^{risk}$)		(0.009)		(0.009)
Risk aversion*IT		-0.007		-0.019*
($\tilde{\gamma}_2^{risk} - \tilde{\gamma}_1^{risk}$)		(0.009)		(0.010)
Risk aversion*TT		0.002		-0.016*
($\tilde{\gamma}_3^{risk} - \tilde{\gamma}_1^{risk}$)		(0.008)		(0.009)
Inequality Aversion		0.181		0.690**
($\tilde{\gamma}_1^{inequal}$)		(0.271)		(0.271)
Inequality Aversion*IT		-0.545*		-1.128***
($\tilde{\gamma}_2^{inequal} - \tilde{\gamma}_1^{inequal}$)		(0.285)		(0.297)
Inequality Aversion*TT		-0.31		-0.970***
($\tilde{\gamma}_3^{inequal} - \tilde{\gamma}_1^{inequal}$)		(0.263)		(0.268)

Table 4.3 (Continued)

Rank-ordered logit model

	No Feedback		Feedback	
Female		0.526*		0.300
$(\tilde{\gamma}_1^{female})$		(0.281)		(0.287)
Female*IT		-1.007***		-0.565*
$(\tilde{\gamma}_2^{female} - \tilde{\gamma}_1^{female})$		(0.294)		(0.311)
Female*TT		-0.345		0.124
$(\tilde{\gamma}_3^{female} - \tilde{\gamma}_1^{female})$		(0.274)		(0.281)
Only child		-0.604**		0.244
$(\tilde{\gamma}_1^{onlychild})$		(0.27)		(0.273)
Only child*IT		0.448*		0.045
$(\tilde{\gamma}_2^{onlychild} - \tilde{\gamma}_1^{onlychild})$		(0.279)		(0.297)
Only child*TT		-0.128		0.416
$(\tilde{\gamma}_3^{onlychild} - \tilde{\gamma}_1^{onlychild})$		(0.262)		(0.271)
GPA		0.304		0.689
$(\tilde{\gamma}_1^{gpa})$		(0.395)		(0.425)
GPA*IT		0.525		0.25
$(\tilde{\gamma}_2^{gpa} - \tilde{\gamma}_1^{gpa})$		(0.392)		(0.458)
GPA*TT		0.449		-0.001
$(\tilde{\gamma}_3^{gpa} - \tilde{\gamma}_1^{gpa})$		(0.379)		(0.421)
Observations	836	836	808	808
Number of groups	209	209	202	202
Log likelihood	-636.821	-604.315	-615.126	-557.175

Note: Standard errors in parentheses; $***p < 0.01, **p < 0.05, *p < 0.1$.

bonus required to switch from more preferred option to less preferred option. In this case, all compensating differentials between any two incentives can be obtained by simple calculation of addition and/or subtraction. Suppose we take piece rate as the baseline, then the differentials required to accept other three incentives instead of piece rate payment are: team tournament requires $0 - 5 = -5$ Chinese Yuan, individual tournament requires 2 Chinese Yuan, and revenue sharing needs $2 + 3 = 5$ Chinese Yuan. Notice that a negative differential means the given incentive "team tournament" is preferred to piece rate by this subject, so s/he is willing to give up some money to ensure her/his preferred incentive instead of the less preferred option of piece rate; in contrast, a positive differential means the incentive "individual tournament/revenue sharing" is less preferred to the piece rate scheme, so the subject require more sign-up bonus as the additional compensation for accepting her/his less preferred incentive.

In Table 4.4, we report the summary statistics of compensating differentials between the baseline incentive (i.e. piece rate) and other three team-based incentives, respectively. The fact of all compensating differentials having positive means indicates the baseline of piece rate should the most popular so that subjects require higher sign-up bonus for accept those team-based incentives.

Without feedback information, the contingent evaluations for three team-based incentives are significantly different from each other. Firstly, among three team-based incentive, individual tournament is most popular with the least compensating differential; Secondly, it is followed by the incentive of team tournament, and the within-pair difference between team and individual tournament equals 1.94 (with std. dev. of 0.79), which is significantly larger than zero at p =0.007; Thirdly, as the least preferred incentive, revenue sharing has the highest compensating differentials, and then the within-pair difference between revenue sharing and team tournament equals 1.01 (with std. dev. of 0.65), which is significantly positive at p =0.059.

In addition, the feedback information has remarkable impact on the evaluations for team-based incentives. The significant change happens on the compensating differential for individual tournament, a dramatic increase in the compensating differential after receiving feedback, which also leads individual tournament becomes the least preferred option among three team-based incentives. The difference of compensating differential for individual tournament between two treatment groups equals 2.55 (with std. dev. 1.04), which is significantly positive at p =0.007. But the difference for other two incentives between "No feedback" treatment and "Feedback" treatment are insignificant. Meanwhile, the provision of information also reduces the difference between any two team-based incentives so that all within-pair differences are not significantly different from zero.

Table 4.4 Summary statistics of compensating differentials

	No Feedback					Feedback				
	Obs.	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max
Revenue Sharing	209	4.35	7.87	-19	24	202	3.39	8.99	-25	30
Individual Tournament	209	1.40	10.11	-29	30	202	3.95	10.93	-30	30
Team Tournament	209	3.34	9.62	-23	25	202	3.26	9.91	-30	30

Model Specification of Compensating Differentials

We then specify a regression model to probe the individual heterogeneity in compensating differentials. Given that $U_{ij} = \delta_j + \beta \cdot Y_{ij} + \gamma_j' x_i + \varepsilon_{ij}$, and take $j = 0$ piece rate scheme as the baseline, then for $\forall j \neq 0$, the compensating differential CD_{ij} that make individual i is willing to accept incentive j instead of piece rate scheme is equal to:

$$CD_{ij} = U_{i0} - U_{ij} = -(\tilde{\delta}_j + \beta \cdot \tilde{Y}_{ij} + \tilde{\gamma}_j' x_i + \tilde{\varepsilon}_{ij}) \tag{4.7}$$

The regression results are reported in Table 4.5. If different methods (i.e. ranking and compensating differential) are used to quantify the same thing, they should yield the same outcome, which means there exists equality between preferences derived from rank-ordered logit model and the model on compensating differentials. It is worth noting that an increase in the probability of being preferred implies a decrease in the demand for additional compensation, hence the sign of coefficients in these two models should be exactly opposite with each other. As a result, we do find the consistency between two specification models.

In the first column i.e. regression model (i), when we only discuss the incentive-specific effect while neglecting the influences of heterogeneous individual attributes on compensating differentials, we can conclude that any team-based incentive is less attractive than the baseline

Table 4.5 GLS model of compensating differentials

	No Feedback			Feedback		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Revenue sharing	4.354***	10.080*	8.117	3.391***	7.008	5.624
($\tilde{\delta}_1$)	(0.640)	(6.118)	(6.683)	(0.702)	(6.868)	(7.527)
Individual tournament	-2.957***	4.024	9.998	0.554	19.190**	10.16
($\tilde{\delta}_2 - \tilde{\delta}_1$)	(0.760)	(7.230)	(7.851)	(0.803)	(7.543)	(8.286)
Team tournament	-1.014	4.226	2.712	-0.134	7.555	-0.816
($\tilde{\delta}_3 - \tilde{\delta}_1$)	(0.760)	(7.232)	(7.849)	(0.803)	(7.543)	(8.286)
Performance difference		-0.208	-0.277*		-0.208	-0.203
(β)		(0.162)	(0.161)		(0.164)	(0.163)
Relative performance		0.001	-0.016		0.045*	0.057*
($\tilde{\gamma}_1^{relative}$)		(0.023)	(0.037)		(0.024)	(0.037)
Relative performance*IT		-0.056**	-0.165***		-0.192***	-0.159***
($\tilde{\gamma}_2^{relative} - \tilde{\gamma}_1^{relative}$)		(0.026)	(0.042)		(0.025)	(0.040)
Relative performance*TT		-0.036	-0.001		-0.033	-0.015
($\tilde{\gamma}_3^{relative} - \tilde{\gamma}_1^{relative}$)		(0.026)	(0.042)		(0.025)	(0.040)
Confidence bias			-0.018			0.014
($\tilde{\gamma}_1^{confidence}$)			(0.039)			(0.041)
Confidence bias*IT			-0.150***			0.041
($\tilde{\gamma}_2^{confidence} - \tilde{\gamma}_1^{confidence}$)			(0.046)			(0.045)
Confidence bias*TT			0.049			0.02
($\tilde{\gamma}_3^{confidence} - \tilde{\gamma}_1^{confidence}$)			(0.046)			(0.045)
Risk aversion			0.053			0.034
($\tilde{\gamma}_1^{risk}$)			(0.041)			(0.045)
Risk aversion*IT			0.033			0.088*
($\tilde{\gamma}_2^{risk} - \tilde{\gamma}_1^{risk}$)			(0.049)			(0.050)
Risk aversion*TT			-0.01			0.091*
($\tilde{\gamma}_3^{risk} - \tilde{\gamma}_1^{risk}$)			(0.049)			(0.050)
Inequality aversion			-0.695			-2.458*
($\tilde{\gamma}_1^{inequal}$)			(1.306)			(1.349)
Inequality aversion*IT			0.939			4.116***
($\tilde{\gamma}_2^{inequal} - \tilde{\gamma}_1^{inequal}$)			(1.533)			(1.488)
Inequality aversion*TT			0.778			3.728**
($\tilde{\gamma}_3^{inequal} - \tilde{\gamma}_1^{inequal}$)			(1.533)			(1.488)

Table 4.5 (Continued)

GLS model of compensating differentials

	No Feedback			Feedback		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Female		-0.694	-0.884		-1.87	-1.583
$(\tilde{\gamma}_1^{female})$		(1.330)	(1.345)		(1.402)	(1.415)
Female*IT		4.675***	3.598**		4.537***	3.642**
$(\tilde{\gamma}_2^{female} - \tilde{\gamma}_1^{female})$		(1.571)	(1.580)		(1.543)	(1.561)
Female*TT		0.369	0.568		-0.114	-1.008
$(\tilde{\gamma}_3^{female} - \tilde{\gamma}_1^{female})$		(1.572)	(1.580)		(1.544)	(1.562)
Only child		2.186*	2.289*		-1.354	-1.18
$(\tilde{\gamma}_1^{onlychild})$		(1.309)	(1.300)		(1.375)	(1.364)
Only child*IT		-2.718*	-2.422*		0.189	0.262
$(\tilde{\gamma}_2^{onlychild} - \tilde{\gamma}_1^{onlychild})$		(1.547)	(1.527)		(1.512)	(1.502)
Only child*TT		0.492	0.532		-1.497	-1.443
$(\tilde{\gamma}_3^{onlychild} - \tilde{\gamma}_1^{onlychild})$		(1.549)	(1.530)		(1.517)	(1.508)
GPA		-1.987	-1.97		-1.167	-1.364
$(\tilde{\gamma}_1^{gpa})$		(1.842)	(1.846)		(2.082)	(2.067)
GPA*IT		-1.651	-1.869		-3.921*	-4.045*
$(\tilde{\gamma}_2^{gpa} - \tilde{\gamma}_1^{gpa})$		(2.176)	(2.170)		(2.293)	(2.282)
GPA*TT		-1.21	-1.451		-1.676	-1.716
$(\tilde{\gamma}_3^{gpa} - \tilde{\gamma}_1^{gpa})$		(2.176)	(2.169)		(2.293)	(2.283)
Observations	627	627	627	606	606	606
R-squared	0.017	0.062	0.101	0.001	0.102	0.137

Note: Standard errors in parentheses; $***p < 0.01, **p < 0.05, *p < 0.1$.

of piece rate scheme. In the absence of feedback information, on average, the "revenue sharing" company should pay 4.354 Chinese Yuan to attract one "single contractor" to join the company. By contrast, the two companies with competitive incentives pay less for additional compensation: the "individual tournament" company should pay $4.354 - 2.957 = 1.397$ Chinese Yuan to attract a "single contractor", and the "team tournament" needs to pay $4.354 - 1.014 = 3.340$ Chinese Yuan as the extra sign-up bonus. The compensating differentials between team-based incentives and piece rate are significantly positive. Meanwhile, among three incentives, the difference between the most popular incentive "individual tournament" and the least popular incentive "revenue sharing" is also statistically significant. On the contrary, providing workers with feedback about their relative performance change the compensating differentials. Although each team-based incentive still has to pay a positive compensation to the "single contractor" (i.e. on average, the company "revenue sharing" needs to pay 3.391 Chinese Yuan; the company "individual tournament" should pay $3.391 + 0.554 = 3.945$ Chinese Yuan; and the company "team tournament" should pay $3.391 - 0.134 = 3.257$ Chinese Yuan), the differences between any two team-based incentives shrink to insignificant levels.

Next, in model (ii) and (iii), we probe the effects of some important individual attributes on the stated-preferences for incentives. The key difference between those two models is: besides control variables (i.e. performance difference, relative performance, gender, only child, GPA) included in model (ii), we add more non-cognitive characters such as overconfidence, risk aversion and social preference in the regression model (iii). As a result, we investigate how those individual attributes affect the compensating differentials in a systematic way.

Effect of Performance Difference

It is rational for the subject to require less compensation for a certain team-based incentive if s/he performs quite well under that incentive. Since s/he can get a higher expected payoff with the performance increasing, the compensating differential between a team-base incentive and the baseline of piece rate is negatively related to the performance difference between those two incentives. Nevertheless, the marginal effect of one unit change in the performance difference is quiet small and insignificant except for the case in regression (iii) with "No Feedback" treatment.

Effect of Relative Performance

Different from the individual-based incentive "piece rate", those three team-based incentives determine the final payment based on workers' relative performance within the team. Hence, in order to maximize the expected payoffs, subjects with high relative performance are more likely to accept the individual tournament due to a high winning probability in the competition. On the contrary, able subjects probably more reluctant to accept the revenue sharing because of the free riding problem. From the results in model (iii), under "Feedback" treatment, 1% increase in the relative performance indicates an additional of 0.057 Chinese Yuan required by the "single contractor" to enter into the company "revenue sharing"; whereas the same 1% increase in the relative performance brings a decrease in compensating differential between individual tournament and the baseline of piece rate, that is, $-0.159 + 0.057 = -0.102$ Chinese Yuan.

Effect of Overconfidence

Consistent with the findings in Wozniak et al. (2010), when no feedback is provided, overconfident subjects are more likely to choose individual tournament than other less competitive incentives. From our regression results on compensating differentials, in the case without feedback, the marginal effect of the overconfidence is significantly higher in the individual tournament than in the revenue sharing. Hence, as compared with the compensating differential for revenue sharing, the individual tournament can save 0.150 Chinese Yuan for every 1% increase in the level of over-confidence. In fact, the majority of subjects are overconfident (see Table 3), so the individual tournament becomes much more attractive than revenue sharing if people make decision based on their biased beliefs. However, the provision of feedback helps all subjects to update their biased beliefs so that they state their contingent evaluations based on their actual relative only, and then the impact of self-assessment (i.e. overconfidence) become very small and insignificant.

Effect of Risk Aversion

Overall, the risk aversion has positive coefficients for all three team-based incentives, which means the more risk-averse subjects are less likely to accept any team-based incentive instead of the baseline of piece rate. Thanks to the uncertainty from the teammate and/or opponent's performance, people might consider those team-based schemes to be more risker options. Therefore, as the degree of risk-aversion increases, the attractiveness of team-based incentives falls down so that a higher compensating differential is required. Moreover, compared with the coefficients without feedback, the marginal effect of risk aversion on compensating differentials is magnified with the provision of feedback. In particular, we have a clear evidence to show that more risk-averse subjects are less likely to choose competitive incentive schemes (i.e. individual tournament and team tournament) as compared with revenue sharing. For instance, after receiving feedback information, given the degree of risk aversion increases by 10%, then the "single contractor" would require $0.034 * 10 = 0.34$ Chinese Yuan as the additional compensation to join the company "revenue sharing". Meanwhile, in addition to the extra bonus of 0.34 Chinese Yuan, the company "individual tournament" needs to pay more $0.088 * 10 = 0.88$ Chinese Yuan to poach this employee from the company "revenue sharing", whereas "team tournament" should increase its compensation by $0.091 * 10 = 0.91$ Chinese Yuan to poach the same person from "revenue sharing".

Effect of Social Preference

One of the most exciting findings in this experiment is that we verified the impact of "inequality aversion" on individual preferences for different team-based incentives. Since both individual and team tournaments generate unequal outcomes while revenue sharing guarantee a fair distribution, the subjects of inequality averse type require more compensating differential to accept the tournaments but ask for less to enter into the "revenue sharing" company. Similar to the changes in coefficients for risk aversion, the marginal effects of inequality preference are also amplified by providing the feedback information. Compared with the results in "No Feedback" treatment, the coefficients keep the same signs but enlarge the size at significant level. More specifically, as compared with a subject who doesn't care the equality of team pay, a "single contractor" who prefers fairness would like to require 2.458 Chinese Yuan less to join the company "revenue sharing". Conversely, the company "individual tournament" needs to pay more 4.116 Chinese Yuan to hire an inequality-averse employee rather than someone doesn't care about the equality distribution. Similarly, the company "team tournament" should also increase its compensation by 3.728 Chinese Yuan to attract an inequality-averse worker.

Effect of Gender

Consistent with the previous studies on gender difference in competition, our results also verified that females require more compensating differentials for accepting a competitive incentive. Moreover, during the first stage of diagnostic games for individual characteristics, it is shown that women are more risk-averse but less confident than men, so women should be reluctant to enter into competition. Also, the unequal outcome in tournament brings a disutility to each inequality-averse subject, so women are more likely to accept the revenue sharing incentive rather than the individual tournament due to their gender-specific preference for fairness.

In the absence of feedback, when we run a regression without controlling those non-cognitive characteristics in model (ii), female require 0.694 Chinese Yuan less than male to accept the cooperation in "revenue sharing" company while ask for a higher compensating differential ($-0.694+4.675 = 3.981$ Chinese Yuan) to enter into the competition of "individual tournament" company. And then, after we control three non-cognitive characteristics (i.e. overconfidence, risk attitude and social preference) in the regression model (iii), the difference between only children and children with siblings still exist, although the marginal effect of gender dummy become smaller. So, it indicates that the gender gaps in cooperation and competition should be partially driven by the different risk attitudes, social preferences and self-confidences between female and male. However, besides those three non-cognitive characteristics we elicited in our, the two groups probably have other intrinsic difference related to their heterogeneous preferences for team-based incentives.

In addition, we also find that providing more information doesn't help to reduce the gender gap. After controlling three non-cognitive characteristics (i.e. overconfidence, risk attitude and social preference) in the regression model (iii), females require 1.583 Chinese Yuan less than males to accept the cooperation in "revenue sharing" whereas ask for a higher compensating differential ($-1.583 + 3.642 = 2.059$ Yuan) to enter into the competition of "individual tournament".

Effect of Only Child

During the first stage of diagnostic games for individual characteristics, we have some significant differences between only children and children with siblings: only children are less risk-averse but also less over-confident than the children with siblings; however, there is no significant difference in social preference between two groups, which provides evidence against the stereotype that only children are always more selfish than children with siblings. Therefore, we expected the two groups would have quite different preferences for those team-based incentives due to the heterogeneity in their non-cognitive characteristics.

In the absence of feedback, when we run a regression without controlling those non-cognitive characteristics in model (ii), compared to children with siblings, only children require a higher compensating differentials (i.e. 2.186 Chinese Yuan) to accept the cooperation in "revenue sharing" company while a smaller compensating differential ($2.186 - 2.718 = -0.532$ Chinese Yuan) to enter into the competition of "individual tournament" company. On one hand, only children are often viewed as disadvantaged as a result of sibling deprivation, which may lead to their being less cooperative, and less likely to get along with peers (Blake 1981). On the other hand, because only children are more risk-loving but less confident at the same time, their preference for the individual tournament would take account of the positive effect from their risk attitude as well as the negative effect from their lower confident in the competition.

Therefore, in the following step, we control three non-cognitive characteristics (i.e. over-confidence, risk attitude and social preference). As a result, the difference between only children and children with siblings still exist and the relevant coefficients have little changes. It implies there should be some unobservable difference between the two groups except three non-cognitive characteristics we elicited in our experiment that influence their preferences for team-based incentives. However, it is worth noting that the introduction of feedback about relative performance effectively reduces the difference between only children and others. So it suggests the hidden factor in the black box should be some non-cognitive characteristics related to information processing in uncertain environments.

Price the Cost of Mismatch

Without the extra sign-up bonus, everyone will definitely self-sort into her/his best choice. However, the compensating differentials generates the possibility that someone may accept a certain less preferred choice with a relatively high sign-up bonus and even be better off than taking her/his most preferred incentive but with a low sign-up bonus. In other words, since all subjects evaluate their utility differences between incentives in money, the cost of any mismatch between a company and a worker can be priced based on their contingent evaluation in terms of compensating differentials. Now, suppose those three team-based companies want to hire a talented employee who has the following characteristics: female, only child, the performance is quite stable and unchanged across different incentives (i.e. $\tilde{Y}_{ij} = 0$), relative performance is at least above 80%, with a little overconfidence around 20%, she is risk-neutral and inequality averse, GPA achieves 3.5 points. From the regression results in model (iii), without feedback information, the company of "revenue sharing" should pay an amount of 2.942 Chinese Yuan (i.e. $8.117 + (-0.277) * 0 + (-0.016) * 80 + (-0.018) * 20 + (0.053) * 50 + (-0.695) * 1 + (-0.884) * 1 + 2.289 * 1 + (-1.970) * 3.5$) as the compensating differential to attract this "single contractor" join the company. And then, the difference of compensating differential between "revenue sharing" and "individual tournament" is equal to -4.306 Chinese Yuan (i.e. $9.998 + (-0.277) * 0 + (-0.165) * 80 + (-0.150) * 20 + (0.033) * 50 + (0.939) * 1 + (3.598) * 1 + (-2.422) * 1 + (-1.869) * 1$, which means the individual tournament is very attractive to this employee so that she would like to give up 4.306 Chinese Yuan to switch from the company of "revenue sharing".

4.5 Conclusions

The main purpose of this study is to build a mapping from the vector of individual characteristics to their stated preferences for different team-based incentives. As a consequence, our findings from the lab data have several important implications in the real world.

First of all, when designing incentive system, organizations and companies should take into account of the interaction between incentives and individual self-sorting. We have verified that workers with different characteristics do not only make different productivity response to various compensation schemes, but also have heterogeneous preferences for those incentives. Hence, in order to achieve particular targeted composition of the workforce, different companies can offer different incentives to get the right people on the right job. Besides, since it's very difficult to elicit those non-cognitive characteristics during the process of recruitment in practice, the incentive design should be very useful to serve as a screening device.

In addition, without enough information, there are numerous subjects wrongly believe their relative performance and then make improper decision depending on their biased self-assessment. By contrast, providing feedback about relative performance moves those subjects towards better choices of incentive. It hence suggests the feedback mechanism plays an important role in the work organization under team-based incentives.

More importantly, as an innovation in our experiment, the compensating differential provides an effective way to price the cost of a mismatch between a worker and a company. Therefore, compared with designing a distinctive incentive scheme and feedback treatment for each heterogeneous worker, it's much more feasible and convenient for companies to set a flexible sign-up bonus policy to attract any talent they want. Nevertheless, we have to admit that the contingent evaluations in terms of sign-up bonus might be restricted to our particular experiment setting and subject pool, but it is still worthwhile to conduct a pilot test that can obtain useful information for the future generalized study.

Besides, our findings about gender effect on stated-preferences shed light on the explanation for gender difference in career choice, and ultimately for the existence of the gender wage gap. In the previous literature, it suggests that the gender gap in competition is driven by gender-specific risk preferences and/or different self-confidence levels. However, in our study, we further verify that the social preference of "inequality aversion" also plays an important role in shaping gender-specific preferences for different team-based incentives.

Finally, between the groups of only children and children with siblings, we observe a substantial divergence in their preferences for different team-based incentives and then a dramatic change after knowing feedback about relative performance. It's well-known that the Chinese government launched One-Child policy in 1979. Many people criticize this policy because it may negatively influence children personalities and ultimately shape their preferences when they grow up and enter into the labor market. Now, according to our results in the experiment, it suggests that only children and children with siblings indeed behave quite differently due to their distinct personalities. However, it is shown that the symmetric information helps reduce the gap between two groups, but the causal factors are still hidden in the black box. Therefore, the further investigation on the issue of only children should be a challenging work for improving the labor market in China.

CHAPTER 5. TRUST BETWEEN STRANGERS: BELIEF UPDATING FROM NOISY FEEDBACK

Abstract

We study how trust between strangers evolves in a setting with noisy feedback about the trustworthiness of others. We first design a laboratory experiment using a two-player sequential trust game where each trustor receives a sequence of noisy binary signals that reveal the trustworthiness type of the trustee. As a result, we track the evolution of trustors' beliefs about trustworthiness types of trustees to document that subjects process information in an asymmetric way compared to a perfect Bayesian: they react more to negative signals rather than positive. Trust is therefore much easier to lose than to gain. Next, we show that our empirical results arise naturally in a theoretical model where individuals optimally manage their trust in others with noisy feedback. We discuss whether trustors should use Bayesian updating to calculate their trust levels or whether they should process information in a biased manner. The conclusion suggests that there exists a complementary relationship between initial trust and optimally biased Bayesian information processing. Hence, we theoretically predict that greater initial trust must be counter-balanced by more asymmetric belief updating. We then use a novel method to demonstrate this hypothesis in the following-up experiment. We match participants from two different universities (in Hong Kong and Beijing, respectively) and prime them on the social identity of their counterparts. Consequently, by the introduction of social identity, we find that both initial trust level and asymmetry of belief updating are stronger for in-group matches than out-group matches, which is consistent with our theoretical prediction.

5.1 Introduction

It is the mutual interest that activates human associations together, and then it is the mutual trust that holds human cooperation successfully. Hence, trusting others and reciprocating that trust with trustworthy actions are the essential aspects of our everyday lives. However, the trust should be specified by particular situations, for instance, the trust between strangers, or trust between friends. In Porta et al. (1997), the authors argued that "trust should be more essential for ensuring cooperation between strangers", "than for supporting cooperation among people who interact frequently and repeatedly".[1] The problem of how impersonal trust can build and persist is intertwined with the transient and anonymous social exchanges which are the foundations of free market economies and successful political systems (Aghion et al. 2010; Cook 2001; Johnson and Mislin 2009).

To investigate the trust between strangers, we select the "trust game" (Guth el al. 1997; Bohnet and Huck 2004; Cox 2004; Malhotra 2004) to conduct our experimental study in a game theoretic context. The trust game is an interaction between two players, a "trustor" and a "trustee".[2] For the trustor, choosing to cooperate indicates her/him to incur a potential cost to benefit the trustee, thus s/he will do so only if s/he believes the opportunity of being repaid by the trustee exceeds the risk of being betrayed. In such a case, the trust can be defined as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another (Rousseau et al. 1998)". Specifically, the positive expectations might vary across individuals with different trust levels. So, in order to measure the differences among individual trust levels accurately, we define the degree of trust as trustors' beliefs about trustworthiness types of trustees.[3] In our design, we use an incentive-

[1] Consider the case of interaction between friends, the cooperation could be persisted by the reputation and ample opportunities for future punishment, instead of being supported by the pure trust alone. In contrast, when we enter into a relation (commercial, political, etc.) with some person, group or organization with which we do not have personal experience before, it requires impersonal trust, that is, trust built without personal history and repeated interactions.

[2] As shown is Figure 5.1, the trustor (i.e. Player A) starts by deciding between non-cooperation ("A1") and cooperation ("A2"), whereas the trustee (i.e. Player B) can only react in case of cooperation by either dividing the rewards equally ("B1") or exploiting the trustor ("B2").

[3] We simply identity types of trustees by their actions in the trust game: a trustee is a "trustworthy" type if s/he responds to trustor's trust with a strategy of "B1"; on the other hand, a trustee is a "untrustworthy" type if s/he reacts with a strategy of "B2".

compatible way to elicit subjects' trust beliefs on others in a carefully controlled experimental environment, and then the experimental data exhibit a range of distinctive trust levels between 0 and 1 across subjects.[4]

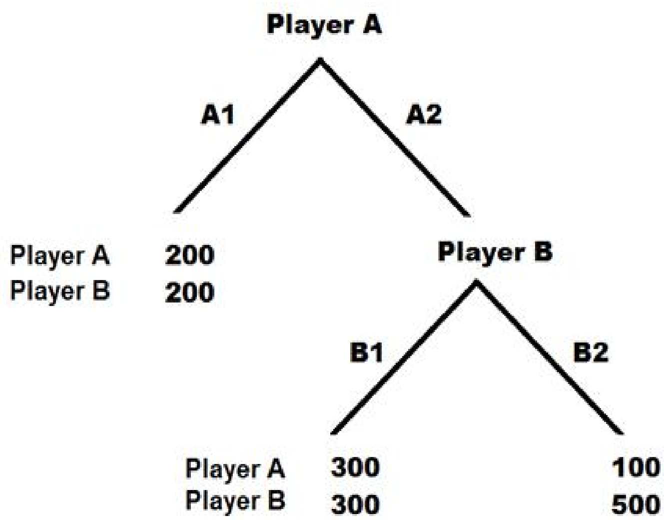


Figure 5.1 Two-player sequential trust game

In addition, we propose that adding information to this trust game should fundamentally alter the trust and cooperation between players. However, the assumption that there is no noise at all in the feedback seems to be too strong in most real world applications. In this paper we therefore investigate the information processing when the assumption of no noise is dropped. We designed a noisy feedback mechanism and then allowed trustors to obtain signals which could reveal trustees' actual trustworthiness types to some extent. Our findings show how trustors update their trust beliefs by noisy feedback about trustees to make better decisions, both in terms of experimental outcomes and in terms of theoretical predictions.

The first contribution of this study is that we do not only reject the perfect Bayesian updating of individual trust on others[5], but also show exactly in which ways it fails using experimental data on changes in trustors' beliefs. According to our experimental design, the

[4] Actually, in this two-player sequential trust game, we apply the strategy method so that all subjects need to state contingent choices for every decision node they may face before being matched in pair. Hence, by eliciting every subject's individual choices as both a trustor and a trustee, we find that individual own trustworthiness is positively correlated to her/his initial trust belief about the trustworthiness of others.

[5] There is a large and growing body of psychological research suggests that the way people process information often departs systematically from perfect Bayesian updating.

trustees are defined by two types: the trustworthiness equals 1 if the trustee repays trustor's trust in the trust game; whereas the trustworthiness equals 0 if the trustee betrays trustor's trust. And then, during the feedback stage, we send trustors a sequence of four noisy binary signals to reveal the trustworthiness of the trustee with a correct probability of 2/3. Hence, no matter the trustee is actually trustworthy or not, it's possible to return either a positive signal that s/he is trustworthy or a negative signal that s/he is untrustworthy. The main finding is that trustors respond different feedback in an asymmetric manner: they respond less to positive signals than to negative ones, which implies that trust is hard to gain but easy to lose.

The second contribution is that we develop a theoretical model to explain the asymmetry evolution of trust which we observed in our laboratory experiment. We show that our empirical results arise naturally in a simple model where individuals optimally manage their trust in others with noisy feedback. First of all, we claim an important assumption for our theory that everyone starts with cooperation and trusting others in the absence of feedback. To support this assumption, we find convincing evidence from the evolutionary biology research (Nowak 2006; Manapat et al. 2012; Rand et al. 2012) as well as our experimental data in the trust game. Those biologists verified that cooperation with trust is our instinct as human-beings through natural selection. On the other hand, in our trust game, the unique subgame perfect equilibrium predicts non-cooperation ("A1", "B2") although this is payoff-dominated by fair cooperation ("A2", "B1"). But different from the rationally self-interested behaviors, there still exists a substantial proportion i.e. 27% of trustors who decided to cooperate based on the initial trust belief of 0.55 on average, which suggests that humans are indeed hard-wired to trust and cooperate to some extent. Next, given that Nature pushes us towards trusting each other at the very beginning, by adding more noisy information about others, we theoretically analyze whether and how people favor "rational" withdrawal from trust by reflection and prospective reasoning. In other words, people intuitively tend to trust with high initial beliefs, but in order to improve the efficiency of "correct" cooperation with trustworthy partners only, the asymmetric non-Bayesian updating is demonstrated as a counterbalance to generate optimal trust levels in the end.

Finally, in recognition of the complementary relationship between initial beliefs and updating process, we theoretically predict that the higher the initial trust belief is, the more asymmetric updating occurs with noisy feedback. And then, we use a novel way to examine this hypothesis in our following-up experiment, which also contributes to research on how social identity affects individual trust updating. We first match participants from two different universities in Hong Kong and Beijing respectively, and then prime them on the social identity of their counterpart.[6] Thus, by the introduction of social identity information, we should be able to artificially sort trustors' initial beliefs into two different ranges. That is, trustors are expected to show higher initial trust for an in-group partner than an out-group partner (Tajfel and Turner, 1979; Tajfel and Turner, 1986[7]; Bernhard et al. 2006; Goette et al. 2006[8]; Eckel and Grossman 2005; Chen and Li 2009[9]). In fact, we confirmed such in-group favoritism from our experimental data: compared to out-group matches, subjects exhibit higher initial trust in in-group members. Meanwhile, consistent with our theoretical prediction, the asymmetry of belief updating for in-group matches is stronger than out-group matches: to a greater extent, subjects overreact to negative signals as well as underreact to positive signals about in-group members during the feedback stage.

The remainder of this chapter is organized as follows. Section 2 provides the experimental design for empirical baseline with anonymous matching in the trust game. Section 3 reports main findings from the experiment. Section 4 analyzes trust in a theoretical model. Section 5 reports the following-up experiment on social identity. Section 6 concludes.

[6]Based on social identity theory, in our experiment, we define the matching between subjects at the same university as the in-group relationship, whereas the matching between subjects from different universities is the out-group relationship.

[7]In Tajfel and Turner (1979, 1986), the social identity theory is developed to understand the psychological basis for intergroup discrimination.

[8]There are numerous studies using natural groups to find effects of social identity on behavior. For instance, Bernhard et al. (2006) uses a dictator game experiment with third-party punishment in two distinct, native social groups in Papua New Guinea. The authors find that third parties show stronger altruism toward ingroup victims. Similar to this finding, in the experiment of Goette et al. (2006), by using natural groups (platoons) in the Swiss Army, they also find more cooperation when subjects interact with ingroup members in a prisoner's dilemma game.

[9]In comparison with natural group identities, Eckel and Grossman (2005) use induced team identity to study the effects of varying identity strength on cooperative behavior in a repeated-play public goods game in the laboratory. Their finding suggests that high degrees of team identification may enhance individual contribution and team cooperation in environments with a public good. Likewise, in Chen and Li (2009), the authors design a laboratory experiment that measures the effects of induced group identity on social preferences. As a result, they show that participants are more altruistic toward an ingroup match.

5.2 Experiment Design and Methodology

The experiment consists of four stages, and we will explain each stage in detail below. During the first stage, given the trust game tree shown in Figure 1, we applied the strategy method and elicited each subject's decisions as both Player A ("trustor") and Player B ("trustee") before being assigned to a specific role in an actual matching. And then, in the end of Stage 1, we matched subjects into pairs. During the second stage, we applied an incentive-compatible way to elicit Player A's initial belief about the trustworthiness of her/his matched partner i.e. Player B. Thirdly, during the feedback stage we repeated the following procedure four times. First, each Player A received a binary signal that indicated whether Player B was trustworthy with a correct probability of 2/3. We then measured each Player A's updated belief from the noisy feedback. Overall, Player A received four independent signals, and stated their updated beliefs after each signal. During the last stage of information purchasing, we allowed Player A to bid for precise information about the trustworthiness of Player B.

5.2.1 Elicit Player A and B's Strategy in a Trust Game

In this stage, we explained the two-player sequential game tree in Figure 5.1 to all subjects by our clear instruction: there are two roles in this game, Player A and Player B. And, the game is played sequentially, i.e. Player A moves first. If s/he chooses A1, the game ends and each receives 200 tokens. However, if Player A chooses A2, then the game allows Player B to make a choice between two options. As a consequence, when Player B chooses B1, each receives 300 tokens; when Player B chooses B2, then Player A receives 100 tokens whereas Player B receives 500 tokens.

In fact, the problem faced by Player B is very simple: the only thing s/he needs to consider is the comparison between choosing "B1" to repay the trust and choosing "B2" to betray the trust on the assumption that Player A decided to trust in the first step. Specifically, Player B will select "B1" if and only if s/he is a trustworthy type; otherwise, Player B will selfishly grasp the maximum profits by selecting "B2". On the contrary, for Player A, the decision-making problem becomes more complicated. S/he has to do a tradeoff between a safe option "A1" which brings a certain outcome for sure and a risky option "A2" which brings either a gain from a trustworthy Player B or a loss from an unreliable Player B. Nevertheless, as long as Player A believes that Player B must be a trustworthy type in the second step, s/he is willing to show her/his trust by selecting "A2" in the first step.

According to our design, during the first stage, each subject was required to elicit not only her/his strategy as Player A, but also the strategy as Player B. From the strategy as Player B, each subject was diagnosed whether to be a trustworthy type or not. On the other hand, from the strategy as Player A, the impersonal trust level could be revealed.

Moreover, at the end of Stage 1, we matched subjects into pairs. Because we primarily aimed to study Player A's belief about the trustworthiness of Player B in this trust game, we then treated all subjects as Player A from the second stage. Meanwhile, all the feedback about Player B and final payoff received by each Player A were associated with the actual action that her/his matched partner had chosen as Player B in Stage 1.

5.2.2 Elicit Player A's Initial Belief about the Trustworthiness of Player B

From this stage, each subject has been already fixed in an assigned matching with someone else. Thus, as Player A, each subject was required to estimate the probability that her/his partner – Player B had chosen "B1" in the trust game during the first stage. That is, they need to estimate the trustworthiness of their partners.

We applied an incentive compatible way[10] that encouraged subjects to make good estimations. They were presented with two following options, 1) Receive 100 tokens if his/her partner – Player B had chosen "B1"; 2) Receive 100 tokens with probability $x \in \{0, 0.01, 0.02, \ldots, 1\}$.

And we asked subjects for what value of x they would be indifferent between two options. Then a random number $y \in \{0, 0.01, 0.02, \ldots, 1\}$ was drawn to compare with the threshold value of x. As a result, subjects would be paid 100 tokens with probability y when $y > x$ and otherwise receive 100 tokens if their partner had chosen "B1" in Stage 1.

In order to present this crossover mechanism in a simple narrative form, we told the story as follows: Now the computer system generates 100 different robots, and each robot is programmed

[10]The mechanism was described by Mobius et. al. (2011).

to choose "B1" with a certain probability. For instance, "Robot 1" chooses strategy "B1" with 1% probability, "Robot 2" chooses "B1" with 2% probability... and "Robot 100" chooses "B1" with 100% probability. One of those robots has been randomly assigned to be the potential substitute for your actual matched partner.

We will pay you an extra 100 tokens if either your actual partner – Player B or the substitute "Robot X" chose strategy "B1". However, you have to tell us when we should use your actual partner's strategy and when we should use the substitute Robot's strategy. As a result, if and only if "Robot X" has a better chance of choosing "B1", i.e. the assigned probability $X\%$ is larger than your estimate of the probability that your actual partner has chosen "B1", we will use the Robot's strategy instead of your actual partner's; Otherwise, we will use your partner – Player B's strategy as usual.

5.2.3 Repeated Elicit Player A's Updated Beliefs with a Sequence of Four Independent Noisy Signals about Player B

During the feedback stage, as Player A, each subject received four independent noisy signals about her/his matched partner – Player B's strategy one by one. In each period after receiving a new noisy signal, each Player A needed to submit her/his updated estimate of the probability that Player B had chosen "B1".

In fact, the noisy feedback was a binary signal of $\{B1, B2\}$ with a correct probability of 2/3 to report Player B's actual strategy during Stage 1. We used a dice rolling mechanism to explain the signal accuracy in the experiment instruction. Suppose that Player B's actual strategy was written on six faces of a dice. But only four of the faces recorded the Player B's strategy correctly, whereas the remaining two faces showed the wrong record. For example, if Player B had chosen "B1" indeed, the dice would have four faces with "B1" and two faces with "B2". And then, the computer rolled the dice and produced one result to report. Since each of those six faces was equally likely to get, the report of Player B's strategy by such a dice rolling mechanism should be correct with a probability of 2/3. Hence, after receiving such a noisy signal about Player B's strategy, each Player A should update her/his belief and submit a better estimation about the trustworthiness of the partner.

Overall, this dice rolling mechanism worked repeatedly for four times, and we tracked subjects' belief updating followed by those four independent noisy signals. Furthermore, after receiving a sequence of four signals, each Player A was allowed to make her/his decision in the trust game again. According to the last updated belief about the trustworthiness of Player B, s/he could restate the same strategy or revise the previous decision in the first stage.

5.2.4 Elicit Player A's Willingness to Pay for Noiseless Feedback about Player B

During the final stage of purchasing perfect information, as Player A, all subjects stated their willingness to pay for the accurate feedback whether their partners had chosen "B1" or not. The bid price was bounded between 0 and 100 tokens. And the bidding process was a standard application of the Becker-DeGroot-Marschak method (BDM): The bid was compared to a price determined by a random number generator. If Player A's bid was greater than the price, s/he needed to pay the price and then received the accurate feedback about Player B's strategy. If Player A's bid was lower than (or equal to) the price, s/he paid nothing but also received nothing.

5.3 Main Results

In May 2012, we conducted the trust experiments at Hong Kong Baptist University, and Central University of Finance and Economics, Beijing. We recruited students to participate in our study by email, and 168 students completed the experiment online. In order to make subjects understand the instruction better, we required them to attend the experiment in the laboratory staffed by our experimenter. The final sample was 168 students, including 88 students from Hong Kong Baptist University, and 80 students from the university in Beijing. Table 5.1 provides the summary statistics of the subject pool: 94% of participants in Hong Kong were undergraduate students, but 74% of participants in Beijing were graduate. The gender ratio of the subject pool was around 0.6, i.e. 103 females and 65 males.

Table 5.1 Summary statistics of subject pool in the trust experiment

	Hong Kong					Beijing				
Variable	Obs.	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max
Age	88	20.95	1.41	18	24	80	23.28	1.73	20	31
Undergraduate	88	0.94	0.23	0	1	80	0.26	0.44	0	1
Female	88	0.63	0.49	0	1	80	0.6	0.49	0	1
Only Child	88	0.33	0.47	0	1	80	0.56	0.5	0	1

5.3.1 Trust and Trustworthiness

During Stage 1 and 2, the heterogeneities of trustworthiness and trust beliefs are observed in the population. In Figure 5.2, the subjects' initial trust levels are divided into five intervals from low to high, and each bar represents the frequency of subjects' initial belief included in that divided interval. So from this distribution of initial trust, it is shown that 45% of subjects elicited their initial belief less than 0.2, while only 13% of subjects have an initial trust belief higher than 0.6. Moreover, consider the ratio of cooperation strategy("A2") and noncooperation strategy("A1") within each interval, the occurrence of choosing cooperation("A2") increases with the initial belief rises. Hence, it indicates that for each individual, there should exist a cutoff point of trust belief, over which s/he will prefer to cooperate with her/his partner.

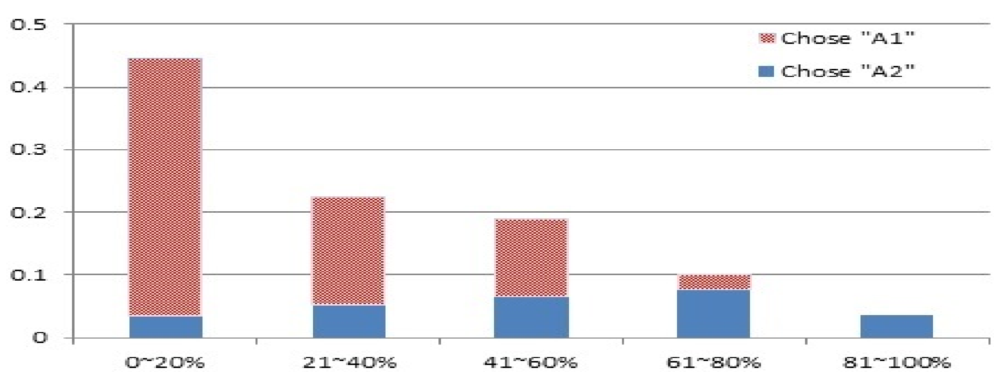


Figure 5.2 Distribution of Player A's initial trust and strategy

The regression result in Table 5.2 verifies a positive relationship between the selection of cooperation strategy and the elicited degree of initial trust. By using a simple Probit model, the estimates in Table 5.2 confirm that, the higher initial belief is, the more likely Player A chose "A2" in the trust game.

Table 5.2 Probit model for the occurrence of trust strategy "A2"

Regressor	I	II
initial belief (μ^0)	0.031***	0.031***
	(0.005)	(0.005)
Beijing		-0.669*
		(0.378)
Female		-0.060
		(0.255)
Undergraduate		-0.439
		(0.378)
Only child		0.136
		(0.256)
Constant	-1.754***	-1.197**
	(0.223)	(0.469)
Observations	168	168

Standard errors in parentheses;
$***p < 0.01, **p < 0.05, *p < 0.1$.

And, we also find that subjects' own trustworthiness is correlated to their trust beliefs about others. From the experimental data in Stage 2, the correlation coefficient between those two factors was equal to 0.20. As a consequence, untrustworthy individuals generally had quite low beliefs on other's trustworthiness (with mean of 27.11, and std.dev. of 26.78), so majority of them (i.e. 97 over 120) showed their noncooperation in the trust game. By contrast, trustworthy type was more likely to trust others, and almost half of them (22 over 48) chose to cooperate with their partners in the game during the first stage. Besides, the mean of their initial trust levels equaled 39.08 (with std. dev. of 26.75), which was significantly higher than the mean of the counterpart with low trustworthiness. This result is consistent with the finding presented in Butler, Giuliano and Guiso (2009).[11]

Besides, individual trustworthiness as well as trust levels were quite different between two

[11] In Butler, Giuliano & Guiso (2009), the authors complement the survey evidence with experimental evidence showing that own trustworthiness and expectations of others' trustworthiness in a trust game are strongly correlated.

universities. For Players B in the trust game during Stage 1, there were 29 out of 88 (i.e. 32.95%) students in Hong Kong Baptist University chose "B1" to show their trustworthiness; however, in Beijing, the trustworthy type of Player B dropped to 19 out of 80 (i.e. 23.75%). Meanwhile, when subjects' roles were switched to Player A, Hong Kong students placed more trust on strangers than Beijing students did. The mean of elicited initial trust belief was 35.80 (with s.d. of 3.00) in Hong Kong, which was significantly greater than the mean of initial trust in Beijing, i.e. 24.74 (with s.d. of 2.81). As a consequence, there were 31 out of 88 (i.e. 35.2%) Players A in Hong Kong Baptist University chose "A2" to cooperate with someone else; but only 14 out of 80 (i.e. 17.5%) Players A in Beijing chose cooperation with trust.

5.3.2 The Evolution of Trust with Noisy Feedback

During the feedback stage, we focused on how subjects' trust beliefs evolve subsequently with four rounds of noisy signals about the trustworthiness of their partners. As our benchmark, we first analyze a specification model of belief updating under the assumption that Player A is a "perfect Bayesian" who uses the correct signal distribution when applying Bayes' rule to form a posterior. We then compare subjects' observed belief updating to the Bayesian benchmark.

Secondly, we discuss the information processing of a "biased Bayesian" who also uses Bayes' rules but can choose at the very beginning how to interpret the informativeness of positive ("B1") and negative("B2") signals about the trustworthiness of others. And then, a comparison between these two information processings will be made.

5.3.2.1 Information Processing of a "Perfect Bayesian"

There are $T = 4$ discrete time periods and in each period $t = 1, \ldots, 4$ Player A receives a binary signal $s_t \in \{B1, B2\}$. The signals are noisy as well as independent identically distributed: a trustworthy Player B who chose "B1" in the trust game returns a "B1" signal to Player A with probability p; and an untrustworthy Player B who chose "B2" might also return a "B1" signal to Player A with probability $q < p$.[12]

[12] In our experiment, $p = 2/3$ and $q = 1/3$.

Suppose in each period of receiving a new noisy signal, $t = 1, \ldots, 4$, Player A has a prior belief of μ^{t-1} and then receives a signal $s_t \in \{B1, B2\}$ to update her/his belief about the trustworthiness of Player B. The perfect Bayesian principal should derive her/his posterior μ^t using Bayes' rule, so when s/he receives a positive signal of "B1",

$\Pr(B\ chose\ B1|s_t = B1) = \frac{\Pr(s_t = B1|B\ chose\ B1)\cdot\Pr(B\ chose\ B1)}{\Pr(s_t=B1|B\ chose\ B1)\cdot\Pr(B\ chose\ B1)+\Pr(s_t=B1|B\ chose\ B2)\cdot\Pr(B\ chose\ B2)}$.

Hence,

$$(\mu^t|s_t = B1) = \frac{p \cdot \mu^{t-1}}{p \cdot \mu^{t-1} + q \cdot (1 - \mu^{t-1})} \tag{5.1}$$

Similarly, if Player A receives a negative signal of "B2" at period t, then the posterior belief becomes

$$(\mu^t|s_t = B2) = \frac{(1-p) \cdot \mu^{t-1}}{(1-p) \cdot \mu^{t-1} + (1-q) \cdot (1 - \mu^{t-1})} \tag{5.2}$$

And then, we graph Figure 5.3 to compare the actual belief updates of our subjects (in the restricted sample)[13] to perfect Bayesian updates. Categorized by five intervals of Player A's prior belief (i.e. μ^{t-1} at each time $t = 1, 2, 3, 4$), the mean belief revision $(\mu^t - \mu^{t-1})$ in response to a "B1" as well as a "B2" signal is plotted for each of the four observations of the overall subjects. First of all, except for the case starting with very low prior beliefs, subjects were more conservative and updated less than the perfect Bayesian when they receive a positive signal of "B1". By contrast, when receiving a negative signal of "B2", subjects over-reacted to the information than the perfect Bayesian benchmark. We will study these phenomena using a linear regression approach next and will confirm the pattern revealed by this figure.

First of all, in order to simplify the information processing model into a linear form, we apply a monotonic transformation on both sides of Equation 5.1 by the logistic function $logit(x) = ln(\frac{x}{1-x})$, then we have

$$logit(\mu^t|s_t = B1) = logit(\mu^{t-1}) + \ln(\frac{p}{q}) \tag{5.3}$$

[13] An important concern in the stage of belief updating is whether subjects understood the description of the signal accuracy and then submitted their true beliefs correctly. During the experiment, they were free to report beliefs inconsistent with perfect Bayesian updating, such as updates in the wrong direction and neutral updates (i.e. reporting the same belief as in the previous round). Except for a high proportion of 17% subjects updated their beliefs in the wrong direction in the first round, the number of subjects who updated in the wrong direction declines dramatically and maintains at a low proportion during the following three rounds (6% in Round 2, 7% in Round 3, 6% in Round 4). Consequently, in order to exclude subjects who misunderstood or ignored the experimental instructions from the data of belief updating, for most of our analysis we use a restricted sample of subjects who didn't update the beliefs in the wrong direction during all four rounds, so this leaves us with 57 Hong Kong students and 61 Beijing students.

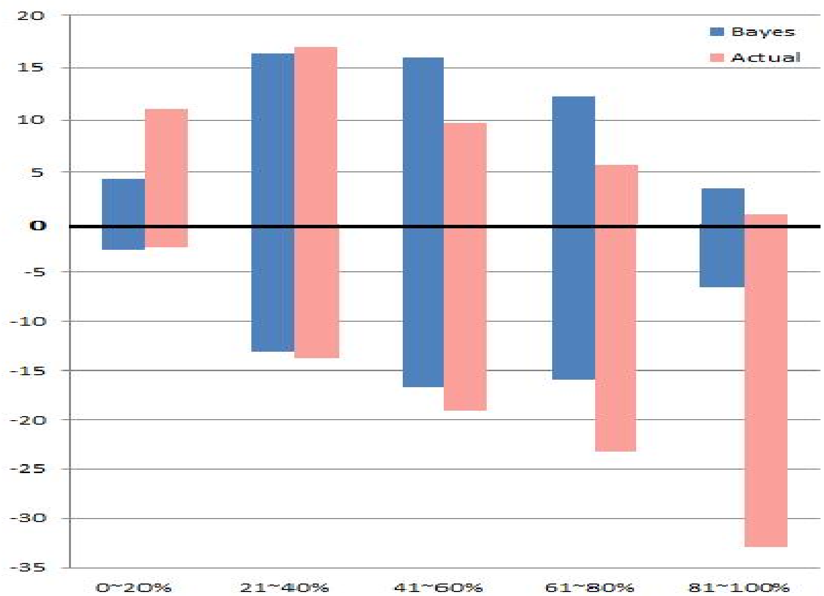


Figure 5.3 Perfect Bayesian updating vs. actual belief updating

Likewise, after taking the monotonic transformation by the same logistic function $logit(x) = ln(\frac{x}{1-x})$, Equation 5.2 is rewritten as

$$logit(\mu^t|s_t = B2) = logit(\mu^{t-1}) + \ln(\frac{1-p}{1-q}) \tag{5.4}$$

Let $\lambda_P = \ln(\frac{p}{q})$ and $\lambda_N = \ln(\frac{1-p}{1-q})$ denote the log-likelihood ratios, i.e. informativeness of positive and negative signals, then the vector $\overrightarrow{\lambda} = (\lambda_P, \lambda_N)$ summarizes the signal structure.

Therefore, we can use the following equation to express a perfect Bayesian belief updating:

$$logit(\mu^t) = logit(\mu^{t-1}) + I(s_t = B1) \cdot \lambda_P + I(s_t = B2) \cdot \lambda_N \tag{5.5}$$

5.3.2.2 Information Processing of a "Biased Bayesian"

Compared to a perfect Bayesian, a biased Bayesian determined by her/himself how to interpret the informativeness of positive("B1") and negative("B2") signals during the feedback stage. Formally, assume the biased Bayesian chose to believe that the log-likelihood ratio of a positive signal is $\widehat{\lambda}_P > 0$ and that of a negative signal $\widehat{\lambda}_N < 0$, along with an initial belief

$\widehat{\mu}^0$. The vector $\overrightarrow{\widehat{\lambda}} = (\widehat{\lambda}_P, \widehat{\lambda}_N)$ thus summarizes how the biased Bayesian interprets the signal structure. Hence, as a biased Bayesian, Player A's posterior belief $\widehat{\mu}^t$ evolves according to Bayes' rule but using her/his distinctively chosen interpretations:

$$logit(\widehat{\mu}^t) = logit(\widehat{\mu}^{t-1}) + I(s_t = B1) \cdot \widehat{\lambda}_P + I(s_t = B2) \cdot \widehat{\lambda}_N \tag{5.6}$$

In addition, we also define the parameters $\beta_P = \frac{\widehat{\lambda}_P}{\lambda_P}$ and $\beta_N = \frac{\widehat{\lambda}_N}{\lambda_N}$ as the biased Bayesian's relative responsiveness to positive and negative signals, respectively; we will directly estimate these parameters in our experiment.

Now we can use the following linear empirical specification for a (possibly biased) Bayesian:

$$logit(\widehat{\mu}_i{}^t) = \delta \cdot logit(\widehat{\mu}_i{}^{t-1}) + \beta_P \cdot \lambda_P \cdot I(s_{it} = B1) + \beta_N \cdot \lambda_N \cdot I(s_{it} = B2) + \epsilon_{it} \tag{5.7}$$

This empirical model allows us to test for the core properties of Bayesian updating as well as measure the biased responsiveness to positive and negative information. According to our design, given $p = 1 - q = 2/3$, we have $\lambda_P = -\lambda_N = \ln(2)$. The coefficient δ captures the persistence of prior information; and the coefficients β_P and β_N capture relative responsiveness to positive and negative information, which also help us to distinguish perfect and biased Bayesian. A perfect Bayesian is fully responsive to positive and negative information ($\beta_P = \beta_N = 1$). In contrast, for example, a biased Bayesian is less responsive to positive information if $\beta_P < 1$, and is more responsive to negative information if $\beta_N > 1$. Thus, it also leads an asymmetric respond to different types of feedback if $\beta_P < 1 < \beta_N$.

We present round-by-round and pooled OLS estimates in Table 5.3. First of all, although the coefficient δ on prior logit-beliefs was presumed to be 1 for both perfect and biased Bayesians, OLS estimate for δ is close to but significantly less than unity. Secondly, the regression results provide strong evidence to confirm that subjects respond differently to positive and negative information as suggested in Figure 5.3. The OLS estimate of β_P is 0.652, which is substantially and significantly less than unity; by contrast, the estimate of β_N is 1.425, which is substantially and significantly greater than unity. Moreover, the difference $\beta_P - \beta_N$ is consistently negative, and significantly different form zero across all rounds as well as for the pooled specification.

In addition, we have also considered the influence of university, gender, only child, and undergraduate status on individual belief updating. By applying the similar OLS model but

Table 5.3 Asymmetric belief updating

	Round 1	Round 2	Round 3	Round 4	All Rounds	Unrestricted
δ	0.748***	0.800***	0.854***	0.927***	0.829***	0.774***
	(0.049)	(0.053)	(0.042)	(0.031)	(0.022)	(0.022)
β_P	0.723***	0.926***	0.157	0.588***	0.652***	0.358***
	(0.251)	(0.253)	(0.220)	(0.167)	(0.115)	(0.108)
β_N	1.534***	1.318***	1.436***	1.272***	1.425***	1.148***
	(0.237)	(0.235)	(0.182)	(0.147)	(0.103)	(0.094)
$Pr(\delta = 1)$	0.000	0.000	0.001	0.022	0.000	0.000
$Pr(\beta_P = 1)$	0.271	0.770	0.000	0.015	0.003	0.000
$Pr(\beta_N = 1)$	0.026	0.179	0.018	0.067	0.000	0.117
$Pr(\beta_P = \beta_N)$	0.036	0.306	0.000	0.006	0.000	0.000
Observations	118	118	118	118	472	672
R-squared	0.805	0.798	0.88	0.946	0.859	0.770

Notes:
1. Each column is a separate regression.
2. The outcome in all regressions is the log posterior odds ratio.
δ is the coefficient on the log prior odds ratio;
β_P and β_N are coefficients for positive and negative signals, respectively.
Perfect Bayesian updating corresponds to $\beta_P = \beta_N = 1$.
3. Estimation samples are restricted to subjects whose beliefs were always within (0, 1).
4. Columns 1-5 restrict to subjects who never updated their beliefs in the wrong direction;
Column 6 includes subjects violating this condition.
5. Columns 1-4 examine updating in each round separately;
Column 5 and 6 pool the 4 rounds of updating.
6. Standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

adding the differential estimates for those factors (i.e. $\beta_{P/N}^{factor} - \beta_{P/N}$), it turns out none of them is significantly different from zero. Hence, those factors are not correlated with the asymmetry pattern of information processing.

5.3.3 The Willingness to Pay for Perfect Information

At the end of Stage 3, each Player A had a chance to rechoose her/his strategy based on her/his updated belief after receiving the overall four signals. So, this latest belief should be treated as the prior belief when subjects calculate the expected value of a noiseless signal about their partners. Based on different levels of this prior belief, we divide subjects into 5 ranges from low trust to high trust, and Figure 5.4 indicates the average bid price within those five ranges. Intuitively, the willingness to pay for perfect information should be equal to its expected benefit for Player A. Therefore, the value of the noiseless signal is quiet small in both extreme cases that Player A's belief is either too low or too high. Meanwhile, as Player A's belief is moving towards the cutoff point,[14] the value of perfect information is increasing. Thus, as illustrated in Figure 5.4, the mean of willingness to pay increases first, and then decreases with a peak around the prior belief interval between 40% and 60%. In one word, the willingness to pay for perfect information is determined by individual last updated belief after noisy information processing.

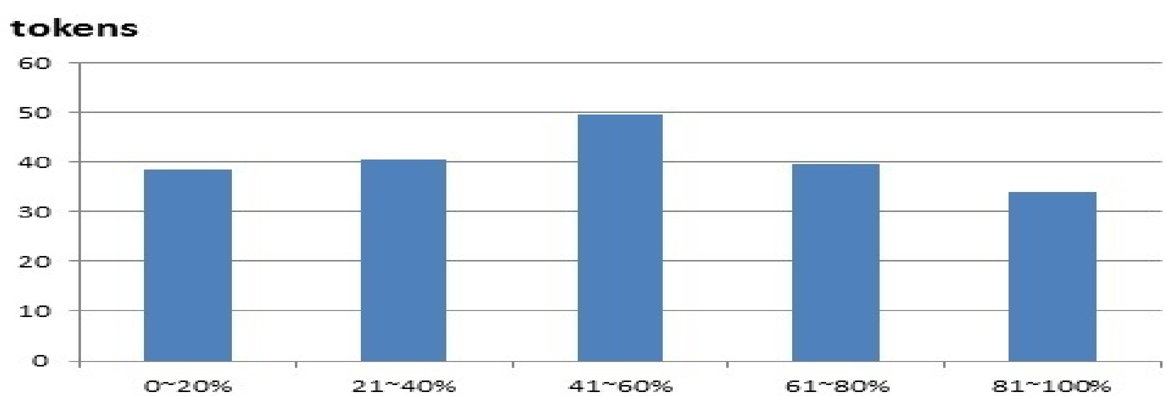


Figure 5.4 Distribution of willingness to pay for noiseless signal

[14]The cutoff point is a threshold of individual trust belief, over which Player A prefers to cooperate with trust.

Furthermore, according to the tendency of change in willingness to pay, it motivates us to generate a piecewise form of a empirical specification as follows:

If subject i ultimately decided noncooperation at time $T = 4$ (i.e. s/he rechose "A1" as the strategy after updating the belief by four noisy signals), we have

$$WTP(\widehat{\mu_i}^T) = \alpha_1 + \gamma_1 \cdot \widehat{\mu_i}^T + \varphi_i \tag{5.8}$$

On the contrary, if subject j decided to cooperate at time $T = 4$ (i.e. s/he rechose "A2" as the strategy after updating the belief by noisy feedback), then

$$WTP(\widehat{\mu_j}^T) = \alpha_2 + \gamma_2 \cdot \widehat{\mu_j}^T + \psi_j \tag{5.9}$$

Table 5.4 presents the OLS estimates for all coefficients. The regression results are consistent with our economic intuitions. The coefficient $\widehat{\gamma_1} > 0$ suggests that the bid price is positively related to the degree of trust for those players who prefer noncooperation; in contrast, the coefficient $\widehat{\gamma_2} < 0$ implies a negative correlation between the willingness to pay and prior belief when the degree of trust is above the cutoff value. Moreover, everyone would like to pay a positive amount for the noiseless feedback regardless of her/his prior belief. Even for those subjects with extremely low or high beliefs, rather than paying nothing, they would like to bid a positive price for the perfect information. So, this amount can be regarded as their evaluation for the truth, which is quite irrelevant to the expected loss or gain from the game per se.

5.4 Theoretical Model

In our experiment on trust belief evolving with noisy feedback, we have found that the subjects updated their beliefs as biased Bayesians and always weighed negative signals much more than positive ones. In this section we show that such asymmetry arises naturally in a theoretical model. We first develop a model to organize our experimental findings in a unified manner and then provide more discussions on optimally biased trust updating as well as some refutable theoretical predictions.

Consider a trustor-trustee relationship, assume each trustor is paired with a trustee whose

Table 5.4 Willingness to Pay for a noiseless signal

	Chose "A1"		Chose "A2"	
	I	II	I	II
updated belief (μ^4)	0.275**	0.345**	-0.256*	-0.342*
	(0.124)	(0.166)	(0.134)	(0.198)
Beijng		0.845		-9.764
		(7.119)		(16.43)
$\mu^4 \times Beijing$		-0.184		0.172
		(0.255)		(0.300)
constant	34.000***	33.750***	55.810***	61.500***
	(3.524)	(5.212)	(7.752)	(12.730)
Observations	123	123	45	45
R-squared	0.039	0.045	0.078	0.086

Standard errors in parentheses; $***p < 0.01, **p < 0.05, *p < 0.1$.

trustworthiness type is the private information for the trustee but unknown for the trustor.[15] At time $t = 0$, each trustor has a subjective prior belief $\mu^0 \in [0, 1]$ that s/he is paired with a "trustworthy" partner.

Next, there are T discrete time periods and in each period $t = 1, \ldots, T$ the trustor can receive a binary signal $s_t \in \{P, N\}$.[16] The signals are noisy as well as independent identically distributed: a "trustworthy" trustee returns a positive signal "P" to the trustor with probability p; and a "untrustworthy" trustee can return a positive signal "P" with probability $q < p$.

Finally, in the end of last period $t = T$, the trustor has to make a decision between cooperation with the trustee and non-cooperation. Non-cooperation gives payoff 0. On the other hand, if the trustor chooses cooperation, s/he can gain 1 as the benefit from working with a "trustworthy" partner but will lose $c > 0$ if s/he wrongly cooperates with an "untrustworthy" partner.[17] The timeline of this model is described in Figure 5.5.

[15] In general, assume there are only two types of trustees: "trustworthy" and "untrustworthy". Take our experiment as an example, in the trust game, Player B who chose B1 is diagnosed as a "trustworthy" type, and Player B who chose "B2" is diagnosed as a "untrustworthy" type.

[16] In our experimental design, a positive signal "P" is denoted by strategy "B1", and a negative signal "N" is denoted by strategy "B2".

[17] Compared to the the trust game we used in our experiment, here we build a trustor-trustee model in a more generalized setting. Actually, we can take our experimental game as a specific example for this theoretical model. Therefore, for the trustor (i.e. Player A), let the non-cooperation payoff i.e. 200 tokens be the reference point, then the marginal payoff from cooperating with a trustworthy parter equals $(300 - 200) = 100$ tokens, whereas the marginal loss of a wrong cooperation is $(100 - 200) = -100$ tokens. So, if we normalize the marginal benefit of 100 tokens to 1, then the cost after normalization also equals 1. That is, we have $c = 1$ in this case.

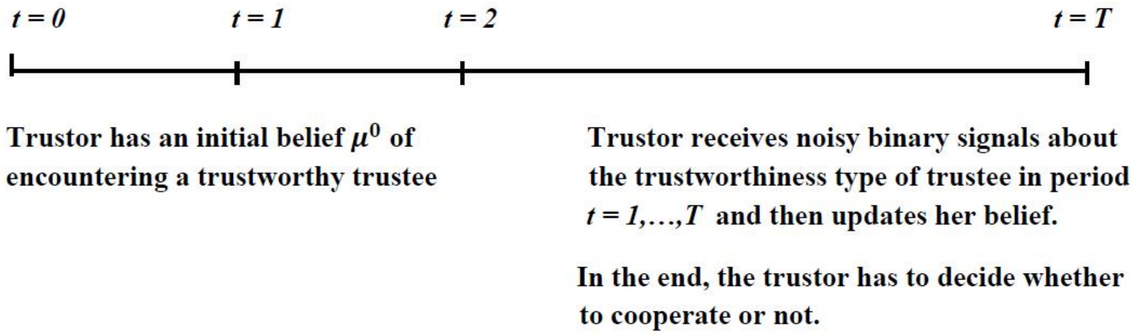


Figure 5.5 Timeline of the trust model

Apparently, the trustor will cooperate if and only if the expected profit of cooperation is positive, which means $\mu^{T} \cdot 1+(1-\mu^{T})\cdot(-c) = \mu^{T} \cdot(1+c)-c > 0$. In other words, the trustor only cooperates when being sufficiently sure that the trustee is trustworthy (i.e. $\mu^{T} > \frac{c}{1+c} = \widetilde{C}$). As a result, the ultimate success of cooperation crucially depends on how the trustor updates her/his posterior beliefs from the noisy feedback.

Based on the specification model of belief updating for a possibly biased Bayesian, i.e. Equation 5.7, we have derived in the previous section of experiment, so here in our theoretical model, we can formulate the similar equation to summarize the information processing of a possibly biased Bayesian as follows:

$$logit(\mu^{t}) = logit(\mu^{t-1}) + I(s_t = P) \cdot \beta_P \lambda_P + I(s_t = N) \cdot \beta_N \lambda_N \tag{5.10}$$

where $\lambda_P = \ln(\frac{p}{q})$ and $\lambda_N = \ln(\frac{1-p}{1-q})$ denote the informativeness of positive and negative signals by Bayes rule. Meanwhile, β_P and β_N are individual relative responsiveness to signals when using Bayesian updating as the benchmark. So, a perfect Bayesian always chooses $\beta_P = \beta_N = 1$.

In addition, denote by $S_P^t(S_N^t)$ the number of $P(N)$ signals the trustor has received by time t. We can easily deduce that a possibly biased Bayesian posterior belief μ^t satisfies

$$logit(\mu^t) = logit(\mu^0) + S_P^t \cdot \beta_P \lambda_P + S_N^t \cdot \beta_N \lambda_N \tag{5.11}$$

The number of $S_P^t(S_N^t)$ received by the trustor is a random variable that depends on the actual trustworthiness type of the trustee. According to the property of noisy signals drawn from a well-defined distribution, the ex-ante expectation of trustor's logit-belief μ^t if s/he was actually paired with a "trustworthy" partner is:

$$E[logit(\mu^t)|trustee\ is\ trustworthy] = logit(\mu^0) + t \cdot [p \cdot \beta_P \lambda_P + (1-p) \cdot \beta_N \lambda_N] \quad (5.12)$$

Similarly, the ex-ante expectation of trustor's logit-belief μ^t if being paired with a "untrustworthy" partner is:

$$E[logit(\mu^t)|trustee\ is\ untrustworthy] = logit(\mu^0) + t \cdot [q \cdot \beta_P \lambda_P + (1-q) \cdot \beta_N \lambda_N] \quad (5.13)$$

Furthermore, the variances of trustor's logit-beliefs when encountering two types of trustees are derived, respectively:

$$Var[logit(\mu^t)|trustee\ is\ trustworthy] = t \cdot p(1-p) \cdot [\beta_P \lambda_P - \beta_N \lambda_N]^2 \quad (5.14)$$

$$Var[logit(\mu^t)|trustee\ is\ untrustworthy] = t \cdot q(1-q) \cdot [\beta_P \lambda_P - \beta_N \lambda_N]^2 \quad (5.15)$$

Figure 5.6(a) and (b) graph the trustors' information processing with different initial beliefs, in which the evolutions of logit-beliefs are illustrated as function of time. In particular, we consider the benchmark of perfect Bayesian updating in two graphs. Equation 5.12 indicates that the ex-ante expectation of $logit(\mu^t)$ increases over time if the trustor is actually paired with a "trustworthy" partner, whereas Equation 5.13 implies that the expectation decreases over time if the trustee is "untrustworthy" type. Hence, no matter where the initial logit-belief $logit(\mu^0)$ is located[18], the red line show the evolving mean logit-belief when the trustee type is "trustworthy" and the blue line show the evolving mean logit belief when the trustee type is "untrustworthy". In fact, the trustor will decide to cooperate with the trustee at time T if and only if her/his logit-belief is higher than the cutoff value of $logit(\widetilde{C})$.[19] As more and more information received by the trustor, the mean logit-beliefs of "trustworthy" and

[18]In Figure 5.6(a), the trustor has a relatively high initial logit-belief that is above the cutoff point for cooperation; In Figure 5.6(b), the trustor has a quite low initial logit-belief that is less than the threshold for cooperation.

[19]Note that $\widetilde{C} = \frac{c}{1+c} \in (0,1)$, so $logit(\widetilde{C}) = \ln(\frac{\widetilde{C}}{1-\widetilde{C}}) = \ln(c)$

"untrustworthy" partners converge to $+\infty$ and $-\infty$ at rate t while the standard deviation increases only at rate $\sqrt{t}$. Therefore, the trustor will gradually reduce the probability to make a wrong cooperation decision as $T \to \infty$, because the probability that her/his logit-belief is on the correct side of the cutoff converges to 1 in each state.

On the other hand, in Figure 5.6, we also discuss the effect of initial beliefs on the final cooperation decision by the comparison between (a) and (b). As shown in subfigure (a), a trustor with high initial trust level is very likely to always trust and cooperate with her/his trustee if s/he doesn't have enough feedback periods to update her/his belief. For example, given that the trustor has been paired with an "untrustworthy" partner, and if the noisy signalling process stops at a certain period before τ_a, then the probable outcome for the trustor is a loss from trusting and cooperating with that untrustworthy person. By contrast, subfigure (b) implies that a trustor with low initial belief might always refuse to cooperate with the trustee if the number of feedback periods is less than τ_b. As a result, the trustor will miss out the benefit from cooperating with a "trustworthy" partner. To sum up, a trustor with high initial trust always needs more time to learn from negative signals about an "untrstworthy" type while a trustor with low initial trust needs more feedback to identify a "trustworthy" partner. But the influence of different initial beliefs on trustors' final cooperation decisions will diminish as $T \to \infty$.

More importantly, rather than assuming the initial belief μ^0 can be any real number between 0 and 1, in our model, trustors' initial beliefs are restricted to be higher than the threshold for cooperation i.e. $\widetilde{C}$. In other words, there is an important assumption for our theory that everyone starts with cooperation in the absence of feedback. Actually, this assumption comes from the compelling findings in evolutionary biology research on human cooperation and mutual trust. For example, in Rand et al. (2012), the authors argued that cooperating with trust is human instinct through natural selection, because cooperative heuristics are developed in daily life where cooperation is typically advantageous. And by providing numerous convincing evidence, they finally concluded that people's first intuitive response is always to select cooperation and that too much thinks encourage selfishness which may undermine their cooperative impulses. Besides, consider the simple "trust game" we used in our experiment, while

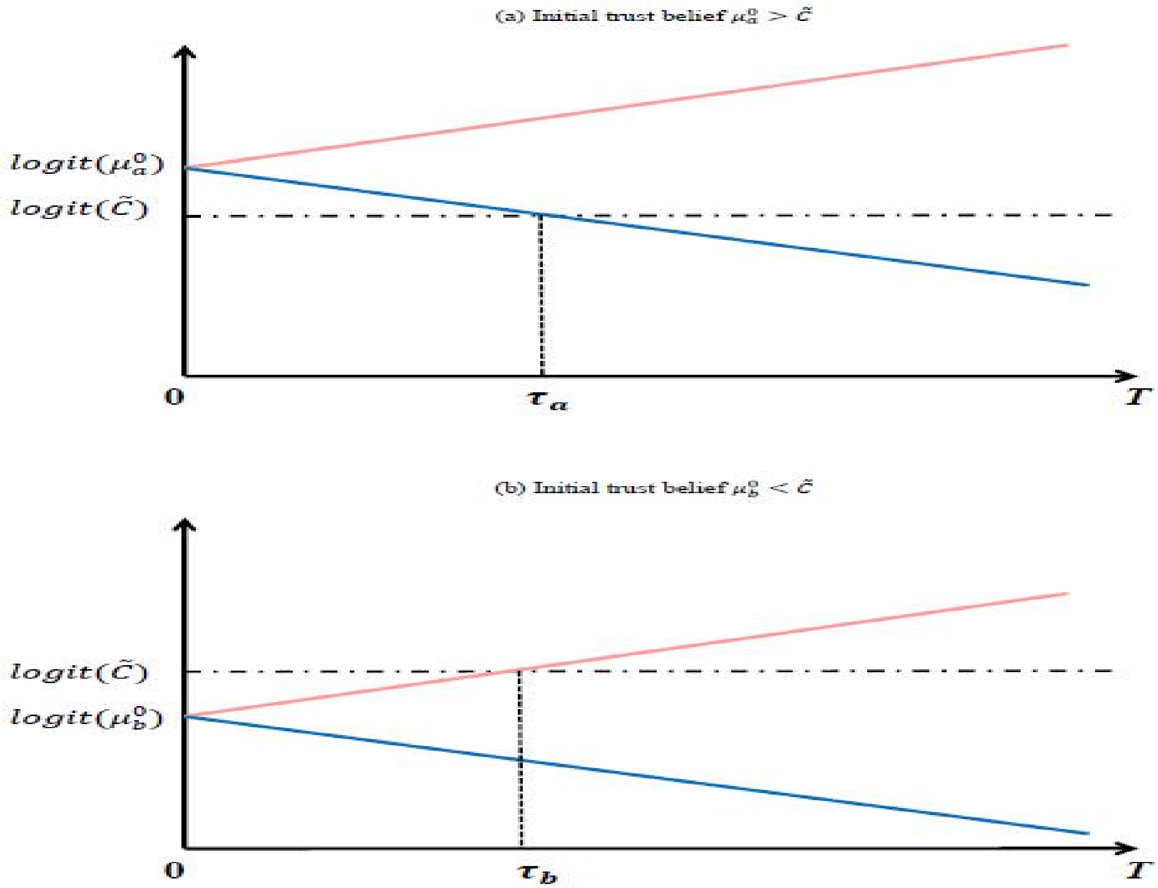


Figure 5.6 Evolution of perfect Bayesian logit-beliefs

rational self-interest trustor should not trust and trustee should not be trustworthy in one-shot anonymous interactions, but we found both trusting in others and reciprocating that trust with trustworthy actions in our data. Thanks to the study of Manapat et al. (2012), this seemingly irrational altruistic behavior has been explained to be a reasonable outcome of social learning in an evolutionary framework. Therefore, it suggests that humans are hard-wired to trust and cooperate to some extent. As a consequence, the following question we are interested in is: if Nature pushes us towards trusting others at the beginning, how do we optimally update our beliefs in the trustworthiness of others by learning from noisy feedback?

The ultimate objective of belief updating is to distinguish the untrustworthy partners from trustworthy ones. Since people intrinsically tend to more cooperation with trust, an optimal information processing must help trustors to overcome defects of "naive" trust in strangers and especially withdraw their trusts from untrustworthy partners with rational reflection and prospective reasoning. Hence, if learning from noisy feedbacks efficiently, trustors should only cooperate with trustworthy partners and also prevent any costly cooperation with untrustworthy type.

Proposition 5.1: Given an initial belief $\mu^0 > \widetilde{C}$, the optimal belief updating process should be a biased Bayesian with $\beta_P < 1 < \beta_N$.

The intuition for this result can be illustrated graphically by the comparison between the benchmark of a perfect Bayesian and a biased Bayesian belief updating in Figure 5.7. Specifically, the solid lines denote the evolving mean logit-beliefs for a perfect Bayesian, and the dotted lines graph the mean logit-beliefs evolve in an asymmetric manner that we have observed in our trust experiment, i.e. with $\beta_P < 1 < \beta_N$. Hence, compared with a perfect Bayesian, who responds to positive and negative signals in a symmetric way[20], a biased Bayesian processes information asymmetrically: s/he overreacts to negative signals but underreacts to positive signals at the same time. When such a biased Bayesian is actually paired with a "trustworthy" partner, then her/his logit-belief will stay in a neighborhood around the initial level for quite a long time regardless of how many positive signals s/he receives. So it's very hard for a trustee to gain more trust from this trustor even s/he is a trustworthy type indeed. However, since the initial belief is high enough to achieve a cooperation unconditionally, the trustor will keep a relatively high trust belief and maintain the cooperation with a trustworthy partner after receiving more feedback. On the contrary, if the trustee is "untrustworthy", the trustor's belief will decrease dramatically as the response for negative signals from that trustee. Thus, the trustor will lose her/his trust in an untrustworthy partner very quickly even though s/he has a high initial belief as well as a strong tendency to cooperation. Undoubtedly, once the trustor's belief falls down to the threshold, s/he will wisely choose non-cooperation as the final decision. To sum up, the underreaction to positive signals hardly affect the trustor's successful cooperation with a trustworthy partner, but the overreaction to negative feedback can definitely

[20]In fact, the perfect Bayesian reacts two types of signals symmetrically if and only if the noisy signalling structure satisfies $p + q = 1$. For example, in our experimental design, $p = 2/3$, and $q = 1/3$.

improve her/his efficiency of identifying the untrustworthy type to avoid a wrong cooperation.

Although people intuitively tend to trust with high initial beliefs, for the sake of judgment precision on the trustworthiness of others, the biased Bayesian updating is demonstrated as a counterbalance to generate relatively reasonable trust levels in the end. Hence, based on the complementary relationship between the initial trust and the information processing in the optimally biased belief updating, the increase in initial belief should be offset by the more asymmetric updating process. So, we have the following prediction.

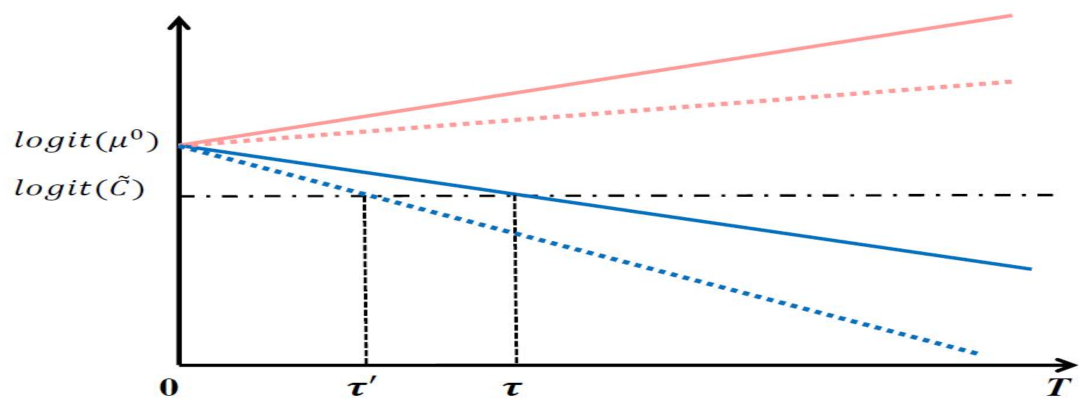


Figure 5.7 Perfect Bayesian updating vs. optimally biased Bayesian updating with $\beta_P < 1 < \beta_N$

Proposition 5.2: Given initial beliefs $\mu_2^0 > \mu_1^0 > \widetilde{C}$, then the optimal belief updating processes should satisfy $\beta_{P2} < \beta_{P1} < 1 < \beta_{N1} < \beta_{N2}$.

Figure 5.8 illustrates how the optimally biased Bayesian adjusts the belief updating in response to a change in the initial belief μ^0. First of all, as the baseline, given that individual trust belief starts with $\mu_1^0 > \widetilde{C}$, we have verified in Proposition 1, that instead of a perfect Bayesian, an optimally biased Bayesian should overweigh negative signals relative to positive ones. This is the best way to counter-balanced the disadvantage of "naive" trust at the very beginning. Next, suppose the initial belief rises up to a higher level μ_2^0, then we need to figure out how the optimally biased Bayesian updating changes. According to the comparative analysis in the graph, there is no doubt that trustors should choose a more asymmetric manner to update their beliefs: the degree of overreaction to negative feedback as well as the extent of underreaction to positive feedback must be strengthened at the same time. Otherwise, if a trustor still keeps the same information processing, because of a greater initial trust belief, the possibility of this naive trustor trusting an untrustworthy partner is increased significantly, and then it is more likely for her/him to suffer a loss from a wrong cooperation in the end.

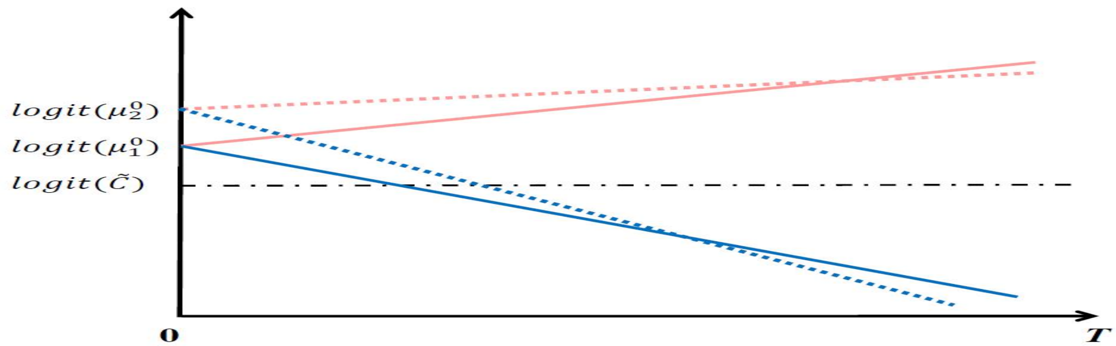


Figure 5.8 Higher initial belief but more asymmetric updating

5.4.1 Value of Perfect Information

Now we analyze how trustors value noiseless information about the trustworthiness of their trustees. Suppose in the end, just after receiving the last signal in period T but before making the final cooperation decision, a trustor with subjective belief μ^T has an opportunity to purchase a perfectly informative signal of the trustee's type. To derive the trustor's willingness to pay, $WTP(\mu^T, c)$, the discussion should be divided into two cases:

5.4.1.1 $WTP(\widehat{\mu}^T, c)$ of Trustors with High Trust

In the absence of perfect information, after receiving the last noisy feedback in period T, trustor i decides to cooperate if and only if her/his expected profit of cooperation is positive, i.e. $\mu_i^T \cdot 1 + (1 - \mu_i^T) \cdot (-c) > 0 \Rightarrow \mu_i^T > \frac{c}{1+c} = \widetilde{C} \in (0, 1)$

Next, if trustor i is allowed to purchase a perfect signal before making the final decision, her/his expected payoff becomes: $\mu_i^T \cdot 1 + (1 - \mu_i^T) \cdot 0$, where the first term represents the expected payoff when s/he receives a positive signal and decides to cooperate in the end, while the second term denotes the expected payoff when s/he receives a negative signal and decides not to cooperate.

Hence, we can now define trustor i's willingness to pay for the perfect information as follows:

$$WTP(\mu_i^T, c) = [\mu_i^T \cdot 1 + (1-\mu_i^T) \cdot 0] - [\mu_i^T \cdot 1 + (1-\mu_i^T) \cdot (-c)] = (1-\mu_i^T) \cdot c \tag{5.16}$$

Moreover, since $\frac{\partial WTP(\mu_i^T, c)}{\partial \mu_i^T} = -c < 0$, the higher the trustor's belief is, the less s/he concerns about the extra information. On the other hand, $\frac{\partial WTP(\mu_i^T, c)}{\partial c} = 1 - \mu_i^T > 0$, so the trustor would like to pay more if the opportunity cost of wrong cooperation is relatively high.

5.4.1.2 $WTP(\widehat{\mu}^T, c)$ of Trustors with Low Trust

Similar to the first case we have discussed above, at the end of period T before purchasing an extra noiseless signal, trustor j prefers non-cooperation if and only if her/his expected profit of cooperation is non-positive, i.e. $\mu_j^T \cdot 1 + (1-\mu_j^T) \cdot (-c) \leq 0 \Rightarrow \mu_j^T \leq \frac{c}{1+c} = \widetilde{C}$. As a result, her/his expected profit equals 0 because s/he will not cooperate at all.

Next, if trustor j can purchase a perfect signal before making the final decision, her/his expected profit becomes: $\mu_j^T \cdot 1 + (1-\mu_j^T) \cdot 0$.

Hence, trustor j's willingness to pay for the perfect information can be defined as follows:

$$WTP(\mu_j^T, c) = [\mu_j^T \cdot 1 + (1-\mu_j^T) \cdot 0] - 0 = \mu_j^T \tag{5.17}$$

Opposite to the first case, $\frac{\partial WTP(\mu_j^T, c)}{\partial \mu_j^T} = 1 - c > 0$ indicates that the lower the trustor's belief is, the less s/he concerns about the extra information. In addition, we have $\frac{\partial WTP(\mu_j^T, c)}{\partial c} = 0$, the willingness to pay is not affected by the opportunity cost of cooperation because the trustor has already preferred non-cooperation to cooperation.

To sum up, the value of willingness to pay can be defined in the piecewise function as follows:

$$WTP(\mu^T, c) = \begin{cases} \mu^T \;\; if \; \mu^T \leq \frac{c}{1+c} \\ (1-\mu^T) \cdot c \;\; if \; \mu^T > \frac{c}{1+c} \end{cases} \tag{5.18}$$

Since the reference point of $\widetilde{C}$ is monotonically increasing with cost value c, given $\widetilde{C} = \frac{c}{1+c}$, then we have $c = \frac{\widetilde{C}}{1-\widetilde{C}}$. So, the piecewise function above can be rewritten as:

$$WTP(\mu^T, \widetilde{C}) = \begin{cases} \mu^T \quad if\ \mu^T \leq \widetilde{C} \\ (1-\mu^T) \cdot \frac{\widetilde{C}}{1-\widetilde{C}} \quad if\ \mu^T > \widetilde{C} \end{cases} \tag{5.19}$$

We illustrate this piecewise function in Figure 5.9. As shown in the graph, all trustors with beliefs between 0 and 1 would like to pay a positive amount for the noiseless information. Then, given a fixed cost of $c > 0$, the willingness to pay for a noiseless signal is increasing with the low belief approaching to the reference point $\widetilde{C}$ on left, and then the demand for perfect information is decreasing with the trust belief rising up continually. Besides, as the opportunity cost of c rises up, the maximum of willingness to pay (i.e. $\widetilde{C}$) does increase correspondingly, because it's worthwhile to pay more for avoiding an "expensive" mistake on cooperation decision. Last but not least, compared with the empirical results in previous section of experiment, subjects' willingness to pay varied in the same manner as we describe in this theoretical model.

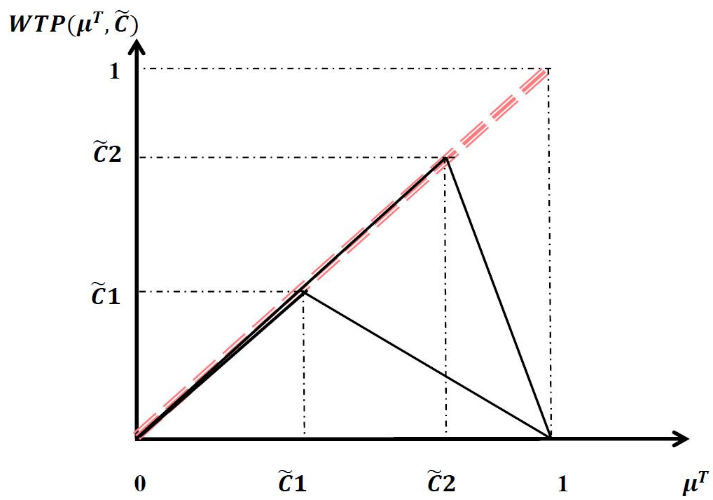


Figure 5.9 Willingness to pay for noiseless feedback

5.5 Follow-up Experiment on Social Identity

In the theoretical model of trust belief updating from noisy feedback, we have claimed in Proposition 5.2, that greater initial trust must be counter-balanced by more asymmetric updating in terms of optimally biased Bayesian information processing. As the supplement for the graphic interpretation in Figure 5.8, we design a follow-up experiment to examine this theoretical hypothesis by actual experimental data.

5.5.1 Experimental Design: In-group vs. Out-group

In order to test the effect of an increase in initial belief on the asymmetry of belief updating, we utilize the influence of social identity information on individual behaviors to artificially sort trustors' initial beliefs into two ranges. During the follow-up experiment, all participants were randomly matched again to play a new trust game. The game structure was still the same, but different from the previous game with completely anonymous matching, each subject could know some identity information about her/his partner before playing the new trust game, that is, everyone was informed of which university her/his partner came from. In fact, we recruited the students from two universities, i.e. Hong Kong Baptist University and Central University of Finance and Economics, Beijing. By randomly assigning those students into two different treatment groups, half of the subjects were informed that they were matched with someone who came from a different university, while the other half were told that they were paired with students from the same university as themselves. And then, the whole procedure from Stage 1 to 4 was repeated again.[21]

According to the social identity theory, the matches between two subjects from the same university can be defined as the in-group relations, whereas the matches between two subjects from different universities are regarded as out-group interactions. We first randomly matched the subjects from two different groups, and then prime them on the social identity of their counterpart. Hence, according to the theory of in-group favoritism, subjects are expected to show higher initial trust for an in-group member than an out-group member. Furthermore, we tracked how subjects' update their beliefs in response to a sequence of noisy signals about the trustworthiness of their partners, so the comparison between in-group and out-group belief updating documents the influence of social identity on the trust evolution in a setting with

[21] Stage 1. Elicit Player A and B's strategy in a trust game; 2. Elicit Player A's initial belief; 3. Elicit Player A's belief updating from a sequence of four noisy signals; 4. Elicit Player A's willingness to pay for noiseless feedback.

noisy feedback.

5.5.2 In-group Trust vs. Out-group Trust

First of all, compared with subjects' strategies in the previous trust game with completely anonymous matching, most of Player B (i.e. 85%) didn't change their trustworthiness type after receiving the social identity information in this new trust game.[22] In general, the individual trustworthiness was not significantly influenced by the social identity information, so Player B treated the trustor in the same way no matter the trustor was an in-group member or out-group member.

However, both Player A's initial belief and the belief updating had remarkable changes due to the intervention of social identity. Since we randomly assigned the subjects into two different treatments at the beginning of the new game, then 82 subjects were matched with an in-group member, while the rest 86 subjects were matched with an out-group partner. Figure 5.10(a) shows the distribution of initial beliefs in an ingroup member and an outgroup member, respectively. The initial trust level are substantially different between two treatments. For instance, there were only 26% of subjects placed an initial belief that is greater than 40(%) in an outgroup counterpart; but, for an ingroup matching, the percentage was doubled so that 51% of subjects placed an initial trust belief that is greater than 40(%). In addition, if we consider the observations separately in two universities as shown in Figure 5.10(b) and (c), the difference between ingroup trust and outgroup trust was still significant. However, the gap shrank in Beijing. More precisely, the mean of initial belief in an outgroup partner was 31.5 in Hong Kong and 26.1 in Beijing. On the other hand, the mean initial trust in an ingroup member was 44.2 in Hong Kong and 33.3 in Beijing.

Next, we run a simple OLS regression to measure the effects of social identity information: the dependent variable is the change in Player A's initial belief compared to her/his initial belief in the anonymous matching baseline, then independent variables include the ingroup and

[22] For those Player B who matched with an in-group member, 9 subjects out of 82 switched their strategy from "B2" to "B1" to increase their trustworthiness for an in-group trustor. By contrast, for those with an out-group matching treatment, there were 5 subjects out of 86 switched the strategy from "B1" to "B2", which reduced their trustworthiness in the new game.

outgroup treatment dummies and we also test the effect of university dummy. The results in Table 5.5 verifies the in-group favoritism by having a significantly positive coefficient for ingroup treatment dummy. However, although the coefficient for the outgroup treatment dummy is negative, but it is not significantly different from zero. Moreover, compared to Hong Kong students, the coefficients of interaction variables indicate that Beijing students would like to place more trust in an out-group member while slightly reduce the in-group trust, but both tendencies are not significant.

Table 5.5 In-group trust vs. Out-group trust

Regressor	I	II
Ingroup	8.098***	8.238**
	(2.666)	(3.732)
Outgroup	-1.256	-4.13
	(2.603)	(3.566)
Ingroup×Bejing		-0.288
		(5.343)
Outgroup×Bejing		6.18
		(5.228)
Observations	168	168
R-squared	0.054	0.062

Standard errors in parentheses;
$* * *p < 0.01, * * p < 0.05, *p < 0.1$.

Furthermore, when all the subjects updated their trust beliefs from a sequence of noisy signals, we then found that Players A were more asymmetric in reacting to feedback about in-group trustees if compared to anonymous matching baseline. In Figure 5.11(a), we make a comparison between the belief updates of subjects with ingroup treatment in the second trust game and their previous belief updates in the first trust game. Categorized by five intervals of individual prior beliefs about the the trustworthiness of Player B, the mean belief revisions in response to a "B1" and "B2" signal are plotted for each of the four observations of the overall subjects. It is obvious that the subjects overreact to negative feedback and underreact to positive feedback about in-group members. On the contrary, also take the belief updating in the anonymous trust as the baseline, Players A seemed to increase their responsiveness to positive signals if they were informed of being matched with an out-group member, which is illustrated in Figure 5.11(b).

This within-subject comparison motivates the following empirical specification for a biased Bayesian:

$$\begin{aligned}
&logit(\widehat{\mu}_i{}^t) = \delta \cdot logit(\widehat{\mu}_i{}^{t-1}) + \beta_P \cdot \lambda_P I(s_{it} = B1) + \beta_N \cdot \lambda_N I(s_{it} = B2) + \\
&(\delta^{Ingroup} - \delta) \cdot logit(\widehat{\mu}_i{}^{t-1}) \cdot I(i\ receives\ ingroup\ treatment) + \\
&(\beta_P^{Ingroup} - \beta_P) \cdot \lambda_P I(s_{it} = B1) \cdot I(i\ receives\ ingroup\ treatment) + \\
&(\beta_N^{Ingroup} - \beta_N) \cdot \lambda_N I(s_{it} = B2) \cdot I(i\ receives\ ingroup\ treatment) + \\
&(\delta^{Outgroup} - \delta) \cdot logit(\widehat{\mu}_i{}^{t-1}) \cdot I(i\ receives\ outgroup\ treatment) + \\
&(\beta_P^{Outgroup} - \beta_P) \cdot \lambda_P I(s_{it} = B1) \cdot I(i\ receives\ outgroup\ treatment) + \\
&(\beta_N^{Outoup} - \beta_N) \cdot \lambda_N I(s_{it} = B2) \cdot I(i\ receives\ outgroup\ treatment) + \epsilon_{it}
\end{aligned} \tag{5.20}$$

Here, we pool the updating data from both feedback stages of anonymous matching baseline and matches with social identity information. So $(\delta^{In/Outgroup} - \delta)$, $(\beta_P^{In/Outgroup} - \beta_P)$ and $(\beta_N^{In/Outgroup} - \beta_N)$ are the differential responses attributable to receiving ingroup treatment or outgroup treatment. These coefficients tell us whether subjects process information differently between anonymous matching and treatments with social identity information. Table 5.6 reports the regression results.

Firstly, the baseline coefficients are similar to the estimates in Table 5.3, which restricts to the belief updating data in the first game only.

Secondly, because we cannot reject $\delta^{In/Outgroup} - \delta = 0$ (Column I), it indicates subjects don't change the way to weigh prior information after receiving social identity information. However, their relative responsiveness to signals are influenced by the social identity: $\beta_P^{Ingroup} - \beta_P$ is both negative and significant, which means the subjects become less responsive to positive feedback when it concerns the trustworthiness of an ingroup member; by contrast, when they receive negative signals from the ingroup member, a significantly positive $\beta_N^{Ingroup} - \beta_N$ verifies the overreaction becomes even worse than the asymmetric updating in the baseline. On the other hand, however, the effects of out-group identity on β_P and β_N are not significant. Therefore, in a word, the degree of trust between ingroup members are very likely to be undermined by negative signals, while this strong effect is hardly to be offset by

Table 5.6 Ingroup vs. Outgroup belief updating

	I	II	III
δ	0.875***	0.873***	0.874***
	(0.021)	(0.015)	(0.020)
β_P	0.560***	0.555***	0.702***
	(0.108)	(0.101)	(0.125)
β_N	1.272***	1.276***	1.406***
	(0.099)	(0.092)	(0.114)
$\delta^{Ingroup} - \delta$	-0.002		
	(0.036)		
$\beta_P^{Ingroup} - \beta_P$	-0.307*	-0.303*	-0.404*
	(0.178)	(0.161)	(0.211)
$\beta_N^{Ingroup} - \beta_N$	0.311*	0.307**	0.346*
	(0.165)	(0.149)	(0.187)
$\delta^{Outgroup} - \delta$	-0.006		
	(0.037)		
$\beta_P^{Outgroup} - \beta_P$	0.021	0.032	-0.225
	(0.175)	(0.158)	(0.196)
$\beta_N^{Outgroup} - \beta_N$	-0.06	-0.074	-0.271
	(0.179)	(0.152)	(0.184)
Observations	768	768	496
R-squared	0.902	0.902	0.888

Notes:
1. Each column is a separate regression.
2. Estimation samples are restricted to subjects whose beliefs were always within (0, 1); and never updated in the wrong directions during the experiment.
3. Columns 3 restricts to subjects who received the same sequence of signals in two games.
4. Standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$.

plenty of positive feedback. So compared to the trust between absolute strangers, the trust is easier to build up with the help of ingroup association, but it's relatively difficult to maintain the trust when encountering negative shocks. Consistent with our theoretical prediction, for in-group matches, subjects choose to process noisy information in a more asymmetric manner during the feedback stage to counter-balance their greater initial beliefs at the beginning.

Lastly, Figure 5.12 graphs subjects' demand for perfect information either from ingroup partner or outgroup partner. Similar to the distribution of willingness to pay in Figure 9, when we divide subjects' updated beliefs at last feedback period into five ranges from low to high, for both treatment groups, the mean of willingness to pay is increasing first and then decreasing, with a peak around the trust belief interval between 40(%) and 60(%). Hence, the distribution of willingness to pay is again consistent with our theoretical model.

5.6 Conclusions

This study aims to open the black box of trust updating between strangers in a setting with noisy feedback about the trustworthiness of others. We first examine how players manage their trust in others by a laboratory experiment. In a two-player sequential trust game, trustors' trust levels as well as the trustworthiness of trustees are measured. Then with four rounds of subsequent noisy feedback about trustworthiness types of trustees, we clearly separate the role of priors and signals in shaping trustors' posterior beliefs during the belief updating process. The experimental findings do not only soundly reject the hypothesis that subjects use perfect Bayesian updating, but also document that subjects process information in an asymmetric way: they react more to negative feedback rather than positive.

Next, we build a simple theory of optimally biased information processing to support our empirical results: given that people intuitively tend to trust others with high initial beliefs, the biased Bayesian updating acts as a counterbalance to adjust individual beliefs properly and then improve the efficiency of mutual beneficial cooperation in the end. So, we claim that there should be a complementary relationship between initial trust and optimal belief updating by noisy feedback.

Consequently, an important implication for our theory is that greater initial trust must be counter-balanced by more asymmetric belief updating. We test this hypothesis by designing a following-up experiment involved with social identity. We match participants from two different universities (in Hong Kong and Beijing, respectively) and prime them on the social identity of their counterpart. As we have expected, compared to anonymous matching baseline, subjects' average initial trust level and asymmetry of information processing are significantly stronger for in-group matches, which provides solid empirical evidence to support our theoretical model.

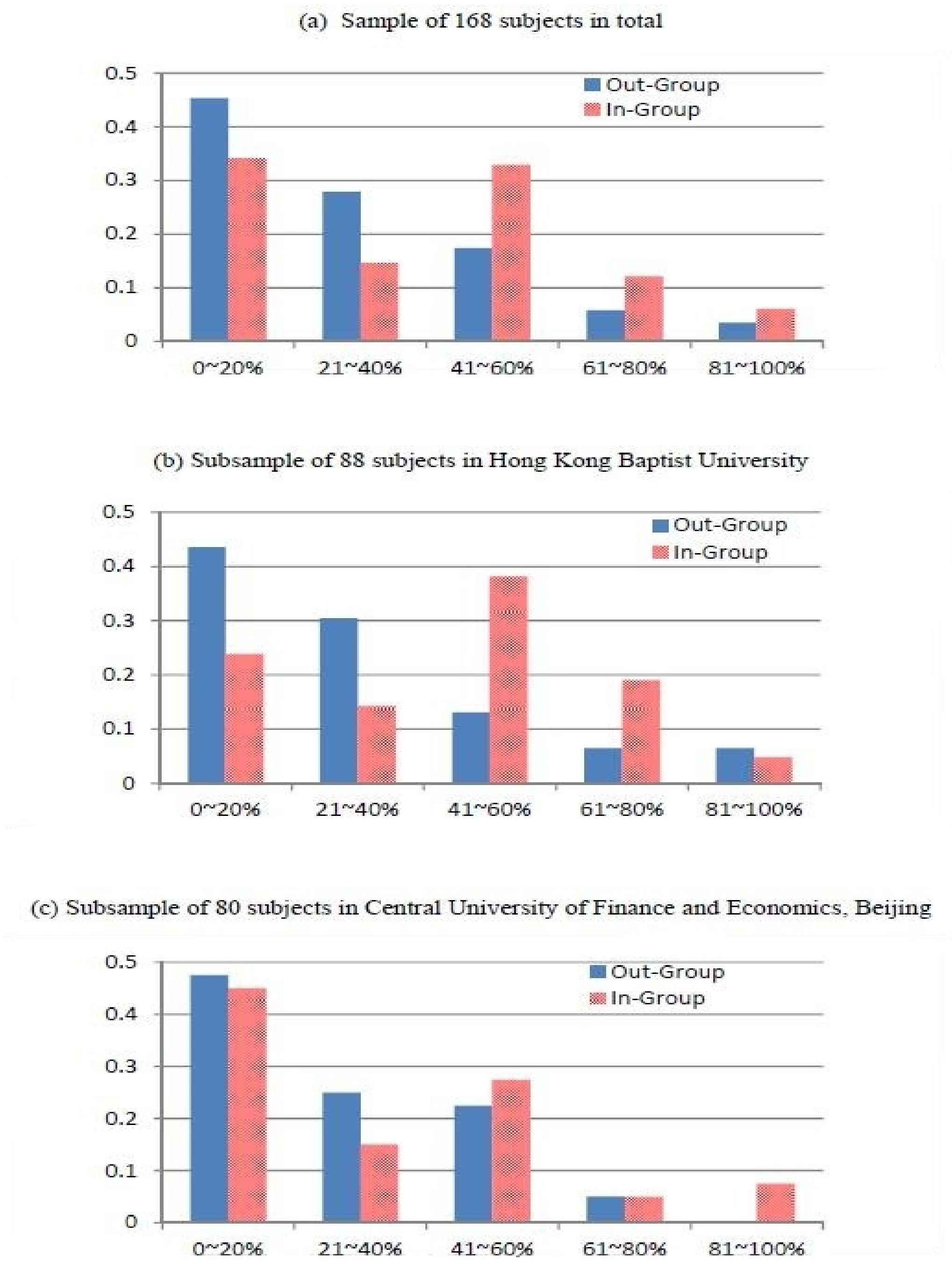


Figure 5.10 In-group initial trust vs. Out-group initial trust

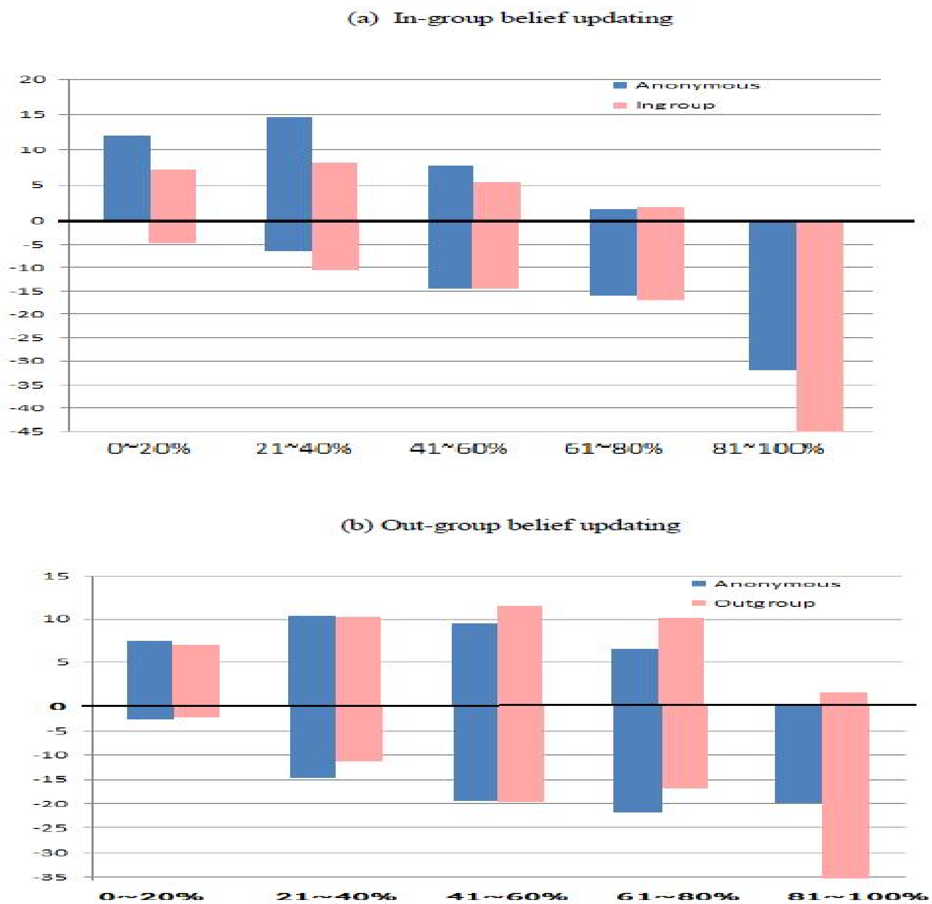


Figure 5.11 In-group trust updating vs. Out-group trust updating

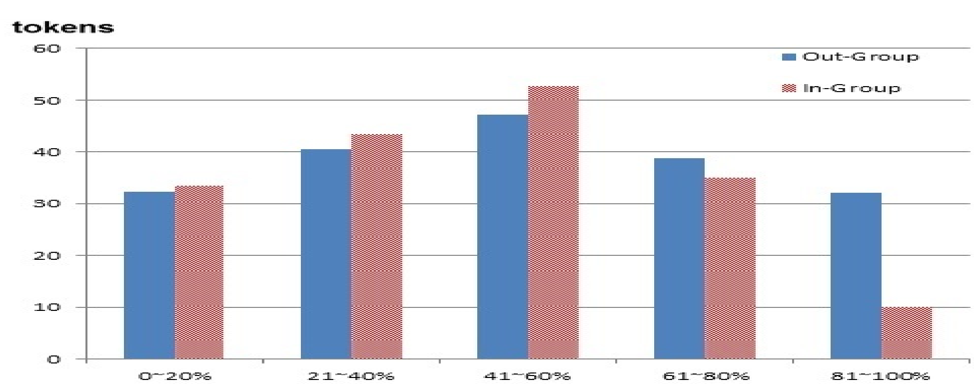


Figure 5.12 Distribution of willingness to pay for noiseless feedback: In-group vs. Out-group

CHAPTER 6. GENERAL CONCLUSIONS

In summary, although our research touches multiple topics in experimental economics, all of them revolve around individual decision-making under uncertainty, that is, the uncertainty of payoffs and choices made by others. Hence, we investigate how people make decisions based on their non-cognitive characteristics as well as their expectations about other's behavior in the interacted workplace. Moreover, by introducing the feedback mechanism, we clearly see the ways that people process the received information to update their beliefs and then change their actual decisions.

Our first labor market experiment uses various simple but efficient experimental methods to elicit individual productivity response to various team-based incentives as well as their preferences for different types of compensation. In order to probe the importance of non-cognitive characteristics on individual behavior, we also use a rich battery of diagnostic games to elicit those characteristics including risk attitude, social preference and self-confidence. As a result, we have a good understanding of the interaction between a worker and a company in our experimental labor market. It is worth noting that this design is the first one to implement the "compensating differential" in a laboratory experiment to price the cost of a mismatch between a worker and a company.

In addition, we obtain a lot of intriguing findings due to the uniqueness of our subject pool of Chinese university students. In particular, we document that China's One-Child Policy has produced significantly more risk-averse, more able, less self-confident, less cooperative, and more competitive individuals. Thus, we can conclude that the One-Child Policy has important ramifications for Chinese labor market and the society as a whole.

Our second study explores how impersonal trust ensures cooperation between strangers in a setting with noisy feedback about the trustworthiness of others. As a consequence, we reject the behavioral predictions based on standard Bayesian rule and verify that individuals follow an optimally biased information processing rule to update their trust beliefs. Moreover, this study also contributes to social identity research by highlighting the impact of social identity on individual initial trust as well as trust updating.

In fact, it still has several potential future extensions on the current work. First of all, compared with the trust between strangers, the trust between friends might be fundamentally different. As another crucial trust form of the basis for cooperation, it is worthwhile to apply the similar experimental approach and investigate how people process the positive as well as negative feedback about the trustworthiness from a friend instead of a stranger; Secondly, social identity should be a very promising topic to strengthen in our future work on trust, and we plan to enlarge our subject pool to make cross-country comparisons such as a US-China trust investigation. Although there are several papers made comparative analysis on the social trust within different countries, few of them concerned the impersonal trust between different societies and cultures. Therefore, we can apply experiments to measure the social trust between different countries, and then probe the determinants for the creation and development of multicultural cooperation.

APPENDIX A. EXPERIMENT INSTRUCTIONS FOR CHAPTER 2 – 4

Welcome to our experiment!

In this study, you will be asked to perform a sequence of tasks and answer questions. We expect that your participation in this study will take about 50 minutes. Your participation in this study is purely voluntary. Your data will be kept completely confidential. When the research is completed, the anonymous study results will be stored in a computer database.

You will receive 10 Yuan for showing up to this experiment. You will receive this show-up fee even if you do not complete the experiment. You can increase your earnings by following the experimental instructions correctly. Your earnings will be paid at the end of the experiment in an envelope. To protect your privacy, the envelope will only show the numeric ID that we have given to you.

If you have any questions about the study before indicating your voluntary consent to participate, please contact Fanzheng Yang (Email: yangfz@iastate.edu). If you would like to participate in the study, please proceed to the next page.

[New Screen]

You are about to participate in a decision making experiment. This experiment is part of a research project financed by Iowa State University.

The experiment consists of eight separate tasks. After you complete all eight tasks, **one of them** will be randomly selected for your payment.

Your total earnings from the experiment will be the show-up fee of 10 Yuan plus the earnings from this randomly selected task.

Although some tasks are similar each other, they are in fact all different. Different participants will perform different tasks in different orders. Therefore, please keep your eyes on your screen and do not disturb other participants. Thank you for your cooperation!

[New Screen]

Task 1

Your first task is to choose several times between a fixed payment and a lottery. The fixed payment on the left will always pay you 20 Yuan while the lottery on the right will pay you either 40 Yuan with some probability or 0 Yuan. For each binary choice, please select either the fixed payment or the lottery. One of your choices will be randomly selected for payment.

Table A.1 Ten binary choices between a fixed payoff and a lottery

Choice 1	Get 20 Yuan for sure	Get 40 Yuan with 10% probability and 0 Yuan with 90% probability
Choice 2	Get 20 Yuan for sure	Get 40 Yuan with 20% probability and 0 Yuan with 80% probability
Choice 3	Get 20 Yuan for sure	Get 40 Yuan with 30% probability and 0 Yuan with 70% probability
Choice 4	Get 20 Yuan for sure	Get 40 Yuan with 40% probability and 0 Yuan with 60% probability
Choice 5	Get 20 Yuan for sure	Get 40 Yuan with 50% probability and 0 Yuan with 50% probability
Choice 6	Get 20 Yuan for sure	Get 40 Yuan with 60% probability and 0 Yuan with 40% probability
Choice 7	Get 20 Yuan for sure	Get 40 Yuan with 70% probability and 0 Yuan with 30% probability
Choice 8	Get 20 Yuan for sure	Get 40 Yuan with 80% probability and 0 Yuan with 20% probability
Choice 9	Get 20 Yuan for sure	Get 40 Yuan with 90% probability and 0 Yuan with 10% probability
Choice 10	Get 20 Yuan for sure	Get 40 Yuan with 100% probability and 0 Yuan with 0% probability

[New Screen]

Task 2

In this task you are randomly matched with 1 other player in this room. One of you will play the role of the Player 1 while the other player will be Player 2. You are equally likely to be player 1 or player 2.

Player 1 in this game gets 30 Yuan. Player 1 can either keep all the money or send some amount to player 2. Whatever amount player 1 sends, is doubled.

For example, if player 1 sends 0 Yuan, then player 2 receives 2*0=0 Yuan and player 1 keeps 30 Yuan; If player 1 sends 10 Yuan, then player 2 receives 2*10=20 Yuan and player 1 keeps 20 Yuan; If player 1 sends 30 Yuan, then player 2 receives 2*30=60 Yuan and player 1 keeps 0 Yuan.

If the computer selects you to be player 1, how many Yuan do you want to send to player 2?

The amount of money I want to send to player 2 is [0,1,. . . ,30] Yuan.

[New Screen]

Task 3

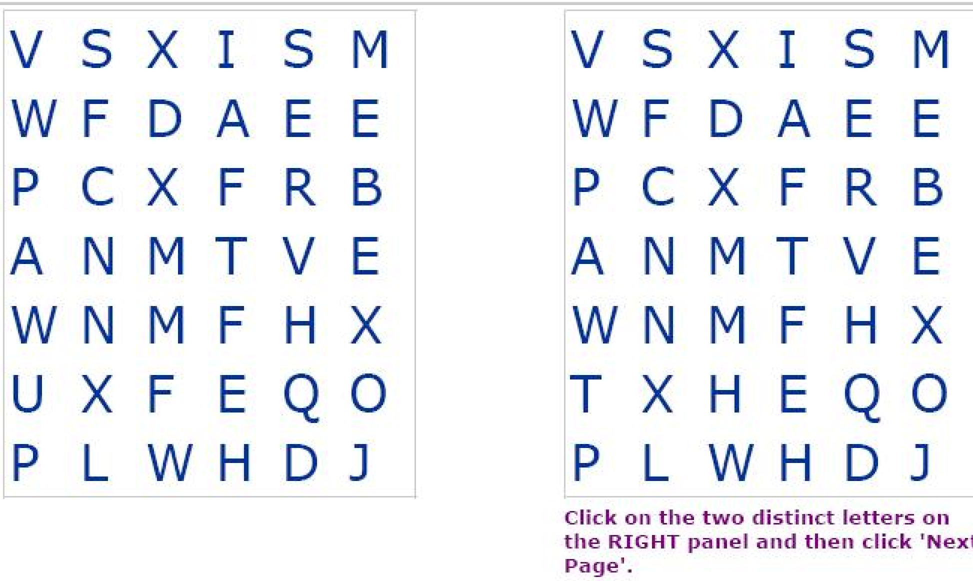


Figure A.1 Example of letter puzzle

In this task, you are asked to solve as many puzzles as you can in 3 minutes. Each solved puzzle will increase your earnings by 1 Yuan. On this page, you have the opportunity to solve three example puzzles to familiarize yourself with the task. They will not count towards your earnings.

The square with characters on the right differs from the square of characters on the left in two letters. You have to find those letters and click on them to solve the puzzle. Whenever you click a letter other than those two, your remaining time is reduced by 1 second: therefore, you should try to only click letters that differ!

Only go to the next page when you are ready. The game will start immediately!

[New Screen]

Task 4

On this page, we ask you to estimate how well you performed in the previous puzzle-solving task compared to a randomly selected other player in this room.

What is the probability that some randomly selected other player in this room is **worse than you**? You can increase your earnings by submitting a good estimate: the computer will give you a "helper robot" who has a certain chance of winning against a randomly selected player. The computer will also compare your performance to the performance of a randomly selected other player in this room.

If the helper robot is better than how you report your ability, then your helper's performance will be used against the other player. If the helper robot is worse than you, your own ability will be used. If you or your helper's performance is greater than the performance of the other player, 10 Yuan will be added to your earnings.

You will maximize the probability of earning an extra 10 Yuan by truthfully reporting your estimate that some other player is worse than you.

For example, if you think you were the best player, then you should report "I think there is a 100 percent chance that a randomly selected other player is worse than me"; If you think that half of all players were better than you, and half of them worse than you, your answer would be "I think there is a 50 percent chance that a randomly selected other player is worse than me"; If you think that all player were better than you, your answer would be "I think there is a 0 percent chance that a randomly selected other player is worse than me".

I think that the probability that a randomly selected other player is worse than me is: [0%, dots, 100%].

[New Screen]

In the next three tasks, you will be an employee of a **company** that specializes in solving puzzles. In each company, you will work together with 3 randomly selected people in this room. However, each company compensates its workers differently.

[New Screen]

Task 5

You are now joining the first company.

This company pays you through an **individual tournament**. The company will consist of two teams of two employees.

You will be a member of one of the teams. You will only be paid **if you solve more puzzles than your team mate** in 3 minutes. In this case you will receive 2 Yuan for every solved puzzle. If you solve in total fewer puzzles than your team mate you will get nothing.

For example, if you solve 15 puzzles and your team mate solves 10 puzzles, then you will earn 30 Yuan.

Only go to the next page when you are ready. Your company will start work immediately and all workers including you will work for 3 minutes.

[New Screen]

Task 6

You are now joining the second company.

This company pays you through a **team tournament**. The company will consist of two teams of two employees.

You will be a member of one of the teams. You will only be paid **if your team solves more puzzles than the other team**. In this case your team will receive 2 Yuan for every puzzle that you or your team mate solved in 3 minutes. These earning will be divided equally between yourself and your team mate.

For example, if you and your team mate solve together 30 puzzles and the other team solved only 20, then your team will receive 60 Yuan which is equally divided into 30 Yuan for yourself and 30 Yuan for your team mate.

Only go to the next page when you are ready. Your company will start work immediately and all workers including you will work for 3 minutes.

[New Screen]

Task 7

You are now joining the third company.

This company pays you through a **revenue sharing**. The company will consist of two teams of two employees.

You will be a member of one of the teams. Your team will receive 1 Yuan for every puzzle that you or your team mate solved in 3 minutes. These earnings will then be equally divided between you and your team mate.

For example, if you solve 15 puzzles and your team mate solves 10 puzzles, then you solved 25 puzzles together and each of you will get 12.5 Yuan.

Only go to the next page when you are ready. Your company will start work immediately and all workers including you will work for 3 minutes.

[New Screen]

Task 8

For your final task you solve again puzzles for 3 minutes. You have several options:

If you choose to work for one of the three companies, your co-workers will be randomly selected among the players in this room. The computer will look up these workers' performance when they worked for the same type of company in a previous round and will use these performances to determine your wage. If you prefer to work as a **single contractor**, then you will be paid by piece rate and receive 1 Yuan for each solved puzzle in this task.

[New Screen]

Please order the 4 employment opportunities from the most preferred to the least preferred. In each case, tell us how much greater of a signup bonus the less preferred company or customer would have to pay you to make you switch from your more preferred option.

For example, you might prefer the individual tournament company the most and working as a single contractor is your second choice. However, if the customer paid you 2 Yuan more in signup bonus, you would become a single contractor rather than work for the individual tournament company. For instance, if the individual tournament company pays you 3 Yuan signup bonus and the customer pays you 6 Yuan, you would become a single contractor. However, if the customer pays you only 4 Yuan, you would rather work for the tournament company.

Table A.2 Ranking and compensating differentials

First choice	[Piece Rate/ Revenue Sharing/ Individual Tournament/ Team Tournament]	
Second choice	[Piece Rate/ Revenue Sharing/ Individual Tournament/ Team Tournament]	How much larger signup bonus would the second choice have to pay over the first choice for you to choose the second choice: [0/ 1 / . . . / 10] Yuan.
Third choice	[Piece Rate/ Revenue Sharing/ Individual Tournament/ Team Tournament]	How much larger signup bonus would the third choice have to pay over the second choice for you to choose the third choice: [0/ 1 / . . . / 10] Yuan.
Fourth choice	[Piece Rate/ Revenue Sharing/ Individual Tournament/ Team Tournament]	How much larger signup bonus would the fourth choice have to pay over the third choice for you to choose the fourth choice: [0/ 1 / . . . / 10] Yuan.

[New Screen]

Your work environment in the final task is [piece rate/ revenue sharing/ individual tournament/ team tournament]. And, your signup bonus is [0 / 1 / . . . / 10] Yuan.

Only go to the next page when you are ready. The game will start immediately.

[New Screen]

Questionnaire

This is the final step to complete the whole experiment! The experiment is now almost over. You just have to answer a brief questionnaire of a few questions.

1. What is your gender? [Male / Female]

2. What is your age? [16 / 17 / . . . / 40]

3. What is your current marital status? [Single / Married]

4. What is your major in the university?

5. What is your college year now? [1st year undergraduate / 2nd year undergraduate / 3rd year undergraduate / 4th year undergraduate / 1st year master / 2nd year master / 3rd year master / 1st year PhD / 2nd year PhD]

6. Where was your Hukou before entering the university?

7. Are you the only child in your family? [Yes / No]

8. What was your score in the University Entrance Examination?

9. What was the type of your University Entrance Examination? [Arts / Science]

10. What is your GPA score?

11. It was easy for me to understand the experiment instruction. [Yes / I had some difficulties / No]

12. I was able to distinguish the differences among the different tasks clearly. [Yes / I had some difficulties / No]

13. I think the number of the tasks is: [OK / somewhat high / too high]

14. I think the length of the experiment is: [OK / somewhat long / too long]

15. I would consider participating again in the similar experiments if I am invited. [Yes / No]

[New Screen]

The experiment is over. Thank you very much for your participation! Your earnings in the experiment will be paid to you within one week. To protect your privacy, your earnings will be put in a sealed envelope with only your ID on it. Please make sure that you pick up your earnings with your ID.

APPENDIX B. EXPERIMENT INSTRUCTIONS FOR CHAPTER 5

Welcome to our experiment!

This experiment is part of a research project financed by Iowa State University. The experiment is by invitation only. All participants in this experiment are students from two universities, including Hong Kong Baptist University and Central University of Finance and Economics, Beijing.

In this study, you need to answer questions and make your decisions in different environment. We expect that your participation in this study will take about 50 minutes.

Your participation in this study is purely voluntary. Your data will be kept completely confidential. When the research is completed, the anonymous study results will be stored in a computer database.

You will receive $30 for participating in this experiment. In addition, you can increase your earnings by following the experiment instructions correctly. **This experiment has four games**, the amount of money you earn will depend upon the decisions you make and on the decisions other people make. Once you complete all games, **only one of them will be randomly selected for your final payment. So, your total earnings from the experiment will be the participation fee of $30 plus the earnings from the randomly selected task.**

Everyone will be paid in private as soon as the experiment is complete. Your earnings in the experiment are given in tokens. At the end of the experiment you will be paid IN CASH based on the exchange rate $1 = 25 tokens. According to your performance, you can earn up to maximum amount of $60.

If you have any questions about the study before indicating your voluntary consent to participate, please contact the experimenter – Fanzheng Yang (Email: yangfz@iastate.edu)

If you would like to participate in the study, please proceed to the next page.

[New Screen]

Now we start the experiment. You will make decisions in four games. Each decision and outcome is independent of each of your other decisions, so that your decisions and outcomes in one game will not affect your outcomes in any other game.

In every game, you will be anonymously matched with one other participant. Moreover, for each new game, you will be randomly matched with a different participant than in the previous one.

There are two roles in each game, A or B. The decisions will be made sequentially, in alphabetical order: Player A will make a decision first and, next, Player B will make a decision. Your decision may affect the payoffs of others, and the decisions of your match may also affect your payoffs. However, you will not be informed of the results of any previous period or game prior to making your decision.

Please proceed to the first game if everything is clear to you in this introduction.

[New Screen]

Game 1

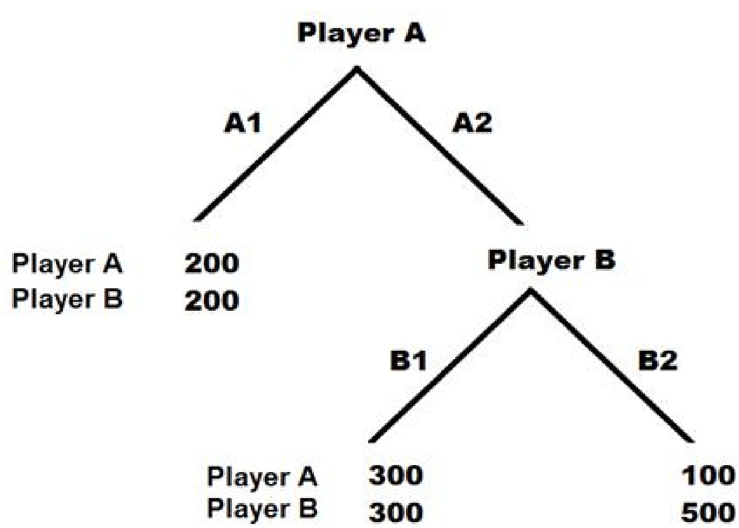


Figure B.1 Trust game tree

In first game, you are Player B.

You will be anonymously matched with one other participant. Player A could be anyone from the different or the same university as yours.

Suppose that Player A has already made a choice. If s/he has chosen A1, you will each receive 200. Your decision only affects the outcome if Player A has chosen A2. Thus, you should choose B1 or B2 on the assumption that Player A has chosen A2 over A1. If Player A has chosen A2 and you choose B1, you will each receive 300. If Player A has chosen A2 and you choose B2, then Player A will receive 100, and you will receive 500.

Decision: Assume Player A has chosen A2, as Player B, I choose [B1/ B2].

[New Screen]

Game 2

In this new game, you are Player B again.

The game structure is the same as the previous one, but you will be randomly matched with a different participant than in the previous decision. **At this time, in particular, Player A will be the student from [Hong Kong Baptist University, Hong Kong / Central University of Finance and Economics, Beijing].**

Suppose that Player A has already made a choice. If s/he has chosen A1, you will each receive 200. Your decision only affects the outcome if Player A has chosen A2. Thus, you should choose B1 or B2 on the assumption that Player A has chosen A2 over A1. If Player A has chosen A2 and you choose B1, you will each receive 300. If Player A has chosen A2 and you choose B2, then Player A will receive 100, and you will receive 500.

Decision: Assume Player A has chosen A2, as Player B, I choose [B1/ B2].

[New Screen]

Game 3

In this game, you are Player A.

You will be anonymously matched with one other participant. Player B could be anyone from the different university or the same university as yours.

You may choose A1 or A2.

If you choose A1, you will each receive 200. If you choose A2, then Player B's choice of B1 or B2 will determine the outcome. If you choose A2 and Player B chooses B1, you will each receive 300. If you choose A2 and Player B chooses B2, you will receive 100, and s/he will receive 500.

Please note that Player B has made a choice without being informed of your decision. Player B knew that his or her choice would only affect the outcome if you had chosen A2, so s/he chose B1 or B2 on the assumption that you had chosen A2 over A1.

Decision: As Player A, which strategy will you choose? [A1/A2]

[New Screen]

Game 3: Please estimate your partner's strategy

No matter which strategy you just chose in Game 3, in this period, we would like to know what you think about your matched partner's strategy. That is, what do you think is the probability that your partner – Player B has chosen B1 in that game?

To make it worthwhile for you to think a bit about this question, we give you a chance of winning additional reward by submitting a good estimate.

Now our system has generated 100 different robots, and each robot is programmed to choose strategy B1 with a certain probability: "Robot 1" chooses strategy B1 with 1% probability, "Robot 2" chooses B1 with 2% probability... and "Robot 100" chooses B1 with 100% probability. One of those robots has been randomly assigned to be the potential substitute for your actual matched partner.

We will pay you an extra 100 tokens if either your actual partner – Player B or the substitute "Robot X" chose strategy B1. However, you have to tell us when we should use your actual partner's strategy and when we should use the substitute Robot's strategy. As a result, if and only if "Robot X" has a better chance of choosing B1, i.e. the assigned probability X% is larger than your estimate of the probability that your actual partner has chosen B1, we will use the Robot's strategy instead of your actual partner's; Otherwise, we will use your partner – Player B's strategy as usual.

In one word, you are most likely to earn the additional 100 token if you simply report your true belief. So, what is your estimate of the probability that your partner – Player B has chosen B1 in Game 3? [0%, 1%,...,100%]

[New Screen]

Game 3: Please update your estimate after feedback 1

You have already made an initial estimation on the probability that your partner – Player B has chosen B1. In the following periods, we would like to you release more information about Player B's strategy to help you update your estimate.

Suppose that Player B's strategy has been written on six faces of a dice. However, only four of the faces record the Player B's strategy correctly, whereas other two faces show a wrong record. For example, if Player B has chosen B1, the dice would have four faces with B1 and two faces with B2.

And then, the computer will roll the dice and produce one result to tell you. Since each of those six faces being equally likely to get, the report of Player B1's strategy by such a dice rolling mechanism should be correct with a probability of 2/3.

Now, the dice is rolled once and it turns out to be (B1/B2).

According to this feedback, we ask you to estimate the probability that Player B has chosen B1; you can restate the same probability or revise your estimate.

Similarly, you can increase your earnings by submitting a good estimate. As before, a "Robert X" has been randomly assigned to be the potential substitute for your actual matched partner, and you have to tell us when we should use your actual partner's strategy and when we should use the substitute Robot's strategy. As a result, if and only if "Robot X" has a better chance of choosing B1, we will use the Robot's strategy instead of your actual partner's, and pay you an additional 100 tokens if the strategy is B1.

Therefore, you are again most likely to earn the additional 100 tokens if your estimate is as accurate as possible. So, what is your updated estimate of the probability that your partner – Player B has chosen B1 in Game 3? [0%, 1%,...,100%]

[New Screen]

Game 3: Please update your estimate after feedback 2

In this period, the dice rolling mechanism, which correctly reports Player B1's with a probability of 2/3, works for the second time. The dice is rolled again and it turns out to be (B1/B2).

According to this new feedback as well as the previous one, we ask you to update the probability that Player B has chosen B1.

As before, you can increase your earnings by submitting a good estimate. Therefore, you are most likely to earn the additional 100 tokens if your estimate is as accurate as possible. So, what is your updated estimate of the probability that your partner – Player B has chosen B1 in Game 3? [0%, 1%,. . . ,100%]

[New Screen]

Game 3: Please update your estimate after feedback 3

In this period, the dice rolling mechanism, which correctly reports Player B1's with a probability of 2/3, works for the third time. The dice is rolled again and it turns out to be (B1/B2).

According to this new feedback and the previous ones, we ask you to update the probability that Player B has chosen B1.

As before, you are most likely to earn an additional 100 tokens if your estimate is as accurate as possible. So, what is your updated estimate of the probability that your partner – Player B has chosen B1 in Game 3? [0%, 1%,. . . ,100%]

[New Screen]

Game 3: Please update your estimate after feedback 4

In this period, the dice rolling mechanism, which correctly reports Player B1's with a probability of 2/3, works for the last time. The dice is rolled again and it turns out to be (B1/B2).

According to this new feedback and the previous ones, we ask you to update the probability that Player B has chosen B1.

As before, you are most likely to earn an additional 100 tokens if your estimate is as accurate as possible. So, what is your updated estimate of the probability that your partner – Player B has chosen B1 in Game 3? [0%, 1%,. . . ,100%]

[New Screen]

Game 3: Please rechoose your strategy as Player A

So far, you have received four feedbacks, and already updated your estimate of the probability that your matched partner – Player B has chosen B1 in Game 3.

No matter which strategy you chose as Player A in Game 3, in this period, we give you chance of making your decision again. So, at this time, you can restate the same strategy or revise your decision in Game 3.

Since your matched partner– Player B's strategy has been fixed, any changes in outcomes and payoffs will be only affected by your new strategy as Player A. **However, at the end of the experiment, we will randomly select one strategy from your previous one and this new one, to calculate your payoff as well as your partner's payoff in Game 3.**

Decision: As Player A, which strategy will you rechoose? [A1 / A2]

[New Screen]

Game 3: Bid for the noiseless feedback on Player B's strategy

In this period, you have chance to purchase the noiseless feedback. Now we ask you to state your willingness to pay for the following two bundles:

(1) In order to receive a correct report about your partner Player B's strategy, how much is your bid price for this bundle? [0, 1, . . . , 100] tokens

(2) In order to receive a correct report about your partner Player B's strategy as well as all other participants' strategies as Player B in the game, how much is your bid price for this bundle? [0, 1, . . . , 100] tokens

Once you complete the bids for both bundles, one bundle out of those two will be randomly selected and you will be able to purchase the corresponding bundle if and only if your bid price exceeds the randomly generated price. If so, you can get that bundle by paying for the generated price.

[New Screen]

Game 4

In this new game, you are Player A again.

The game structure is still the same, but you will be randomly matched with a different participant than in the previous games. **Different from Game 3 in which Player B can be anyone from either your own university or the other university, at this time, your partner Player B will be the student from [Hong Kong Baptist University, Hong Kong / Central University of Finance and Economics, Beijing] in particular.**

You may choose A1 or A2.

If you choose A1, you will each receive 200.

If you choose A2, then Player B's choice of B1 or B2 will determine the outcome. If you choose A2 and Player B chooses B1, you will each receive 300. If you choose A2 and Player B chooses B2, you will receive 100, and s/he will receive 500.

Please note that Player B has made a choice without being informed of your decision except for knowing which university you come from. Player B knew that his or her choice would only affect the outcome if you had chosen A2, so s/he chose B1 or B2 on the assumption that you had chosen A2 over A1.

Decision: As Player A, which strategy will you choose? [A1 / A2]

[New Screen]

Game 4: Please estimate your partner's strategy

In this period, similar to the estimation tasks in Game 3, we would like to know what you think about your matched partner's strategy in Game 4. That is, what do you think is the probability that your partner – Player B has chosen B1 in that game? Please note that different from Game 3 in which Player B can be anyone from either your own university or the different university, your partner in Game 4 is a student from [Hong Kong Baptist University, Hong Kong / Central University of Finance and Economics, Beijing] in particular.

Similarly, you can increase your earnings by submitting a good estimate. As before, a "Robert X" has been randomly assigned to be the potential substitute for your actual matched partner, and you have to tell us when we should use your actual partner's strategy and when we should use the substitute Robot's strategy. As a result, if and only if "Robot X" has a better chance of choosing B1, i.e. the assigned probability X% is larger than your estimate of the probability that your actual partner has chosen B1, then we will use the Robot's strategy instead of your actual partner's, and pay you an additional 100 tokens if the strategy is B1.

Therefore, you are most likely to earn the additional 100 tokens if your estimate is as accurate as possible. So, what is your estimate of the probability that your partner – Player B has chosen B1 in Game 4? [0%, 1%,. . . ,100%]

[New Screen]

Game 4: Please update your estimate after feedback 1

You have already made an initial estimation on the probability that your partner – Player B has chosen B1 in Game 4. In the following periods, we would like to you release more information about Player B's strategy to help you update your estimate.

Like in Game 3, we will use the same dice rolling mechanism to report Player B's strategy. As a result, the feedback we give to you should be correct with a probability of 2/3.

Now, the dice is rolled once and it turns out to be (B1/B2). So, what is your updated estimate of the probability that your partner – Player B has chosen B1 in Game 4? [0%, 1%,. . . ,100%]

[New Screen]

Game 4: Please update your estimate after feedback 2

In this period, the dice rolling mechanism, which correctly reports Player B1's with a probability of 2/3, works for the second time. The dice is rolled again and it turns out to be (B1/B2).

According to this new feedback as well as the previous one, we ask you to update the probability that Player B has chosen B1. As before, you are most likely to earn an additional 100 tokens if your estimate is as accurate as possible. So, what is your updated estimate of the probability that your partner – Player B has chosen B1 in Game 4? [0%, 1%,...,100%]

[New Screen]

Game 4: Please update your estimate after feedback 3

In this period, the dice rolling mechanism, which correctly reports Player B1's with a probability of 2/3, works for the third time. The dice is rolled again and it turns out to be (B1/B2).

According to this new feedback and the previous ones, we ask you to update the probability that Player B has chosen B1. As before, you are most likely to earn an additional 100 tokens if your estimate is as accurate as possible. So, what is your updated estimate of the probability that your partner – Player B has chosen B1 in Game 4? [0%, 1%,...,100%]

[New Screen]

Game 4: Please update your estimate after feedback 4

In this period, the dice rolling mechanism, which correctly reports Player B1's with a probability of 2/3, works for the last time. The dice is rolled again and it turns out to be (B1/B2).

According to this new feedback and the previous ones, we ask you to update the probability that Player B has chosen B1.

As before, you are most likely to earn an additional 100 tokens if your estimate is as accurate as possible. So, what is your updated estimate of the probability that your partner – Player B has chosen B1 in Game 4? [0%,1%,...,100%]

[New Screen]

Game 4: Please rechoose your strategy as Player A

So far, you have received four feedbacks, and already updated your estimate of the probability that your matched partner – Player B has chosen B1 in Game 4.

No matter which strategy you chose as Player A in Game 4, in this period, we give you chance of making your decision again. So, at this time, you can restate the same strategy or revise your decision in Game 4.

Since your matched partner– Player B's strategy has been fixed, any changes in outcomes and payoffs will be only affected by your new strategy as Player A. **However, at the end of the experiment, we will randomly select one strategy from your previous one and this new one, to calculate your payoff as well as your partner's payoff in Game 4.**

Decision: As Player A, which strategy will you choose? [A1 / A2]

[New Screen]

Game 4: Bid for the noiseless feedback on Player B's strategy

In this period, you have chance to purchase the noiseless feedback. Now we ask you to state your willingness to pay for the following two bundles:

(1) In order to receive a correct report about your partner Player B's strategy, how much is your bid price for this bundle? [0, 1, . . . , 100] tokens

(2) In order to receive a correct report about your partner Player B's strategy as well as the strategies of all other Player Bs who are from the same university as your partner, how much is your bid price for this bundle? [0, 1, . . . , 100] tokens

Once you complete the bids for both bundles, one bundle out of those two will be randomly selected and you will be able to purchase the corresponding bundle if and only if your bid price exceeds the randomly generated price. If so, you can get that bundle by paying for the generated price.

[New screen]

Post-experiment Survey

1. What is your age? [16,. . . , 40]
2. What is your major in the university?
3. Are you an undergraduate or graduate student?Undergraduate / Graduate
4. Which year are you in your program?

5. Where is your citizenship? [Mainland China / Hong Kong / United States / Other, please specify]

6. How many siblings (i.e. brothers or sisters) do you have? [0, 1, 2,. . . , 10]

7. In general, to what extent would you say that most people can be trusted? [Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree]

8. To what extend do you think that "every person for themselves" is a good description of how people act in the community? [Strongly Disagree / Disagree / Neutral / Agree / Strongly Agree]

9. I always care what other people think of me. How accurately does this statement describe you? [Completely Untrue / Mostly Untrue / It Depends / Mostly True / Completely True]

10. Have you ever participated in any economics or psychology experimental studies before? [Yes/ No]

11. Is it easy for you to understand the experiment instruction? [Yes / Somewhat / No]

12. Are you able to distinguish the differences among the different games clearly? [Yes / Somewhat / No]

BIBLIOGRAPHY

Adams, J. S. (1963). Towards an understanding of inequity. *The Journal of Abnormal and Social Psychology*, 67(5):422–436.

Aghion, P., Algan, Y., Cahuc, P., and Shleifer, A. (2010). Regulation and distrust. *The Quarterly Journal of Economics*, 125(3):1015–1049.

Andersen, S., Harrison, G. W., Lau, M. I., and Rutstrm, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3):583–618.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401):464–477.

Andreoni, J. and Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2):737–753.

Andreoni, J. and Vesterlund, L. (2001). Which is the fair sex? gender differences in altruism. *The Quarterly Journal of Economics*, 116(1):293–312.

Arya, A., Frimor, H., and Mittendorf, B. (2003). Pouring money down the drain? how sunk investments and signing bonuses can improve employee incentives. *The B.E. Journal of Economic Analysis & Policy*, 0(1):6.

Bandiera, O., Barankay, I., and Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics*, 120(3):917–962.

Barsky, R. B., Kimball, M. S., Juster, F. T., and Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement survey. *Quarterly Journal of Economics*, 112:537–579.

Becker, G. S. and Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy*, 84(4):S143–S162.

Benz, M. and Meier, S. (2008). Do people behave in experiments as in the field? – evidence from donations. *Experimental Economics*, 11:268281.

Bernhard, H., Fehr, E., and Fischbacher, U. (2006). Group affiliation and altruistic norm enforcement. *The American Economic Review*, 96(2):217–221.

Blake, J. (1981). The only child in america: Prejudice versus performance. *Population and Development Review*, 7(1):43–54.

Bognanno, M. L. (2001). Corporate tournaments. *Journal of Labor Economics*, 19(2):290–315.

Bohnet, I. and Huck, S. (2004). Repetition and reputation: Implications for trust and trustworthiness when institutions change. *The American Economic Review*, 94(2):362–366.

Bonatti, A. and Hrner, J. (2011). Collaborating. *American Economic Review*, 101(2):632–63.

Bonin, H., Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2006). Cross-sectional earnings risk and occupational sorting: The role of risk attitudes. IZA Discussion Papers 1930, Institute for the Study of Labor (IZA).

Booth, A. L. and Frank, J. (1999). Earnings, productivity, and performance-related pay. *Journal of Labor Economics*, 17(3):447–463.

Bull, C., Schotter, A., and Weigelt, K. (1987). Tournaments and piece rates: An experimental study. *Journal of Political Economy*, 95(1):1–33.

Butler, J., Giuliano, P., and Guiso, L. (2009). The right amount of trust. Working paper, National Bureau of Economic Research.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *The American Economic Review*, 89(1):306–318.

Cameron, L., Erkal, N., Gangadharan, L., and Meng, X. (2013). Little emperors: Behavioral impacts of china's one-child policy. *Science*, 339(6122):953–957.

Castillo, M., Petrie, R., and Torero, M. (2010). On the preferences of principles and agents. *Economic Inquiry*, 48:266–273.

Charness, G. and Kuhn, P. (2011). Chapter 3 - lab labor: What can labor economists learn from the lab? volume Volume 4, Part A, pages 229–330. Elsevier.

Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3):817–869.

Charness, G., Rigotti, L., and Rustichini, A. (2007). Individual behavior and group membership. *The American Economic Review*, 97(4):1340–1352.

Chen, Y. and Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1):431–57.

Cook, K. S. (2001). *Trust in Society*. New York: Russell Sage Foundation.

Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior*, 46(2):260–281.

Dargnies, M.-P. (2009). Does team competition eliminate the gender gap in entry in competitive environments ? Documents de travail du Centre d'Economie de la Sorbonne 09006, Universit Panthon-Sorbonne (Paris 1), Centre d'Economie de la Sorbonne.

Ding, W. and Lehrer, S. F. (2007). Do peers affect student achievement in china's secondary schools? *The Review of Economics and Statistics*, 89(2):300–312.

Dohmen, T. and Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review*, 101(2):556–90.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2010). Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. *Journal of the European Economic Association*, 3:522–550.

Eckel, C. C. and Grossman, P. J. (2005). Managing diversity by creating team identity. *Journal of Economic Behavior & Organization*, 58(3):371–392.

Ehrenberg, R. G. and Bognanno, M. L. (1990). Do tournaments have incentive effects? *Journal of Political Economy*, 98(6):1307–1324.

Engelmann, D. and Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *The American Economic Review*, 94(4):857–869.

Englmaier, F. and Wambach, A. (2010). Optimal incentive contracts under inequity aversion. *Games and Economic Behavior*, 69(2):312–328.

Eriksson, T., Poulsen, A., and Villeval, M. C. (2009). Feedback and incentives: Experimental evidence. *Labour Economics*, 16(6):679–688.

Eriksson, T. and Villeval, M. C. (2004). Other-regarding preferences and performance pay an experiment on incentives and sorting. IZA Discussion Papers 1191, Institute for the Study of Labor (IZA).

Ermisch, J., Gambetta, D., Laurie, H., Siedler, T., and Noah Uhrig, S. C. (2009). Measuring people's trust. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(4):749–769.

Ertac, S. (2011). Does self-relevance affect information processing? experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior and Organization*, 80(3):532–545.

Falk, A. and Fehr, E. (2003). Why labour market experiments? *Labour Economics*, 10(4):399–406.

Falk, A. and Ichino, A. (2006). Clean evidence on peer pressure. *Journal of Labor Economics*, 24(1):39–57.

Fehr, E. and Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, 46(4-5):687–724.

Fehr, E., Fischbacher, U., Schupp, J., von Rosenbladt, B., and Wagner, G. G. (2003). A nation-wide laboratory examining trust and trustworthiness by integrating behavioural experiments into representative surveys. CEPR Discussion Papers 3858, C.E.P.R. Discussion Papers.

Fehr, E., Klein, A., and Schmidt, K. M. (2007). Fairness and contract design. *Econometrica*, 75(1):121–154.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.

Fehr, E. and Schmidt, K. M. (2000). Fairness, incentives, and contractual choices. *European Economic Review*, 44(4):1057–1068.

Freeman, R. B. and Gelber, A. M. (2010). Prize structure and information in tournaments: Experimental evidence. *American Economic Journal: Applied Economics*, 2(1):149–64.

Gachter, S. and Fehr, E. (1999). Collective action as a social exchange. *Journal of Economic Behavior & Organization*, 39(4):341–369.

Gachterchter, S. and Fehr, E. (2002). Fairness in the labour market: A survey of experimental results. IEW - Working Papers 114, Institute for Empirical Research in Economics - University of Zurich.

Gibbons, R. (1987). Piece-rate incentive schemes. *Journal of Labor Economics*, 5(4):413–29.

Gneezy, U., Leonard, K. L., and List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, 77(5):1637–1664.

Gneezy, U. and List, J. A. (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, 74(5):1365–1384.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, 118(3):1049–1074.

Gneezy, U. and Rustichini, A. (2004). Gender and competition at a young age. *American Economic Review*, 94(2):377–381.

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. *The American Economic Review*, 96(2):212–216.

Guth, W., Ockenfels, P., and Wendel, M. (1997). Cooperation based on trust. an experimental investigation. *Journal of Economic Psychology*, 18(1):15 – 43.

Harbring, C. and Irlenbusch, B. (2003). An experimental study on tournament design. *Labour Economics*, 10(4):443–464.

Healy, A. and Pate, J. (2011). Can teams help to close the gender competition gap?*. *The Economic Journal*, 121(555):1192–1204.

Hewston, M., Rubin, M., and Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, 53:575 – 604.

Johnson, N. D. and Mislin, A. (2009). Cultures of kindness: A meta-analysis of trust game experiments. Working paper.

Joyce, B., John, D., and Kevin, M. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1):122–142.

Karlan, D., Mobius, M., Rosenblat, T., and Szeidl, A. (2009). Trust and social collateral. *The Quarterly Journal of Economics*, 124(3):1307–1361.

Karni, E. (2009). A mechanism for eliciting probabilities. *Econometrica*, 77(2):603–606.

Kluger, A. and Denisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2):–.

Krakel, M. (2008). Emotions in tournaments. *Journal of Economic Behavior & Organization*, 67(1):204–214.

Laffont, J.-J. and Tirole, J. (1988). The dynamics of incentive contracts. *Econometrica*, 56(5):1153–1175.

Lazear, E. P. (1989). Pay equality and industrial politics. *Journal of Political Economy*, 97(3):561–580.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, 90(5):1346–1361.

Leider, S., Mobius, M. M., Rosenblat, T., and Do, Q.-A. (2009). Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics*, 124(4):1815–1851.

Malhotra, D. (2004). Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties. *Organizational Behavior and Human Decision Processes*, 94(2):61–73.

Manapat, M. L., Nowak, M. N., and Rand, D. G. (2012). Information, irrationality and the evolution of trust. Working paper.

Mas, A. and Moretti, E. (2009). Peers at work. *American Economic Review*, 99(1):112–45.

Massey, C. and Wu, G. (2005). Detecting regime shifts: The causes of under- and overreaction. *Management Science*, 51(6):932–947.

Mobius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2011). Managing self-confidence: Theory and experimental evidence. Working Paper 17014, National Bureau of Economic Research.

Nagin, D. S., Rebitzer, J. B., Sanders, S., and Taylor, L. J. (2002). Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *The American Economic Review*, 92(4):850–873.

Nalbantian, H. R. and Schotter, A. (1997). Productivity under group incentives: An experimental study. *American Economic Review*, 87(3):314–41.

Nalebuff, B. J. and Stiglitz, J. E. (1983). Prizes and incentives: Towards a general theory of compensation and competition. *The Bell Journal of Economics*, 14(1):21–43.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *The Quarterly Journal of Economics*, 122(3):1067–1101.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563.

Paserman, D. (2007). Gender differences in performance in competitive environments: Evidence from professional tennis players. IZA Discussion Papers 2834, Institute for the Study of Labor (IZA).

Porta, R. L., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. W. (1997). Trust in large organizations. *The American Economic Review*, 87(2):333–338.

Rabin, M. and Schrag, J. L. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics*, 114(1):37–82.

Rand, D. G., Greene, J. D., and Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, 489:427–430.

Rockenbach, B., Sadrieh, A., and Mathauschek, B. (2007). Teams take the better risks. *Journal of Economic Behavior & Organization*, 63(3):412–422.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., and Camerer, C. (1998). Introduction to special topic forum: Not so different after all: A cross-discipline view of trust. *The Academy of Management Review*, 23(3):393–404.

Sainty, B. (1999). Achieving greater cooperation in a noisy prisoner's dilemma: An experimental investigation. *Journal of Economic Behavior & Organization*, 39(4):421–435.

Salop, J. and Salop, S. (1976). Self-selection and turnover in the labor market. *The Quarterly Journal of Economics*, 90(4):619–627.

Schotter, A. and Weigelt, K. (1992). Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results. *The Quarterly Journal of Economics*, 107(2):511–39.

Slovic, P. and Lichtenstein, S. (1971). Comparison of bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 6(6):649–744.

Soroka, S. N. (2006). Good news and bad news: Asymmetric responses to economic information. *Journal of Politics*, 68(2):372–385.

Steelman, L. C., Powell, B., Werum, R., and Carter, S. (2002). Reconsidering the effects of sibling configuration: Recent advances and challenges. *Annual Review of Sociology*, 28:243–269.

Stiglitz, J. E. (1975). The theory of "screening," education, and the distribution of income. *The American Economic Review*, 65(3):283–300.

Suen, W. (2004). The self-perpetuation of biased beliefs. *The Economic Journal*, 114(495):377–396.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47(2):143–148.

Tajfel, H. and Turner, J. (1979). *An Integrative Theory of Intergroup Conflict*. Brooks/Cole.

Tajfel, H. and Turner, J. (1986). *The Social Identity Theory of Intergroup Behavior*. Nelson-Hall.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

van Dijk, F., Sonnemans, J., and van Winden, F. (2001). Incentive systems in a real effort experiment. *European Economic Review*, 45(2):187–214.

Van Wesep, E. D. (2009). Pay before performance: The signing bonus as an incentive device. Afa 2009 san francisco meetings paper.

Wozniak, D., Harbaugh, W. T., and Mayr, U. (2010). The menstrual cycle and performance feedback alter gender differences in competitive choices. University of Oregon Economics Department Working Papers 2010-2, University of Oregon Economics Department.

Zhang, J. (2011). Do girls in china compete just as much as boys? evidence from an experiment that predicts educational choice. Technical report, University of California at Berkeley.