Essays on social preferences and strategic interactions : an experimental economics analysis

Liu, Jia

2014

Liu, J. (2014). Essays on social preferences and strategic interactions : an experimental economics analysis. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/62119

https://doi.org/10.32657/10356/62119

ESSAYS ON SOCIAL PREFERENCES AND STRATEGIC INTERACTIONS: AN EXPERIMENTAL ECONOMICS ANALYSIS LIU JIA 2014

NANYANG TECHNOLOGICAL UNIVERSITY

ESSAYS ON SOCIAL PREFERENCES AND STRATEGIC INTERACTIONS: AN EXPERIMENTAL ECONOMICS ANALYSIS

LIU JIA

SCHOOL OF HUMANITIES AND SOCIAL SCIENCES

2014

ESSAYS ON SOCIAL PREFERENCES AND STRATEGIC INTERACTIONS: AN EXPERIMENTAL ECONOMICS ANALYSIS

LIU JIA

LIU JIA

School of Humanities and Social Sciences

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2014

Acknowledgements

This thesis will be completing the journey of pursuing my Ph.D. Despite that the journey is long and challenging, it has been one of the best things ever happened in my life. I would never have been able to make it this far without the support and help of numerous nice people around me. It is my great pleasure to thank those who have contributed to the completion of this thesis.

First and foremost, I would like to express my heartfelt and deepest gratitude to my supervisor and co-author of four research papers included in this thesis, Prof. Yohanes Eko Riyanto, for his invaluable guidance and tremendous help during the last several years. I am sincerely grateful to him for agreeing to accept and fund me as his Ph.D. student and for having always been supportive, encouraging and enlightening. This thesis could not have been completed without his immense knowledge, insightful suggestions, generosity with his time in helping me and patience. I feel blessed to have him as my supervisor and one could not wish for a friendlier and better mentor.

I am also grateful to many other faculty members in the division of economics. I thank my internal thesis committee and oral examination panel members, Prof. He Tai-Sen and Prof. Hong Fuhai, Prof. Christos Sakellariou for their invaluable comments. I also thank the two external examiners for their insightful and constructive comments. My gratitude also goes to Prof. Ernie Teo for being nice in my Ph.D. admission interview, Prof. Nattavudh Powdthavee, Prof. Quang Nguyen and Prof. Zhang Jipeng for being supportive, helpful and caring during my Ph.D. study, Prof Joseph Alba, Prof Huang Weihong, Prof. Low Chan Kee and Prof. Yao Shuntian for teaching me advanced knowledge in economics. I am indebted to many visiting professors who have provided insightful comments on my research projects. I thank them for making their time available and for always being approachable.

I thank organizers of the Asia summer institute in behavioral economics at National University of Singapore for offering me a seat there. I benefited substantially from the summer institute. I also thank organizers of the joint UCSD-NHH spring school in behavioral economics for accepting me into the spring school and for the financial support during the spring school. It was a beneficial

and memorable experience. I acknowledge the financial support from division of economics for covering local and overseas conferences.

I would like to thank my fellow doctoral friends in Singapore and friends in China for support over the years. I have been lucky to have them around me in one way or another. I thank Li Jingping for friendship and for being a co-author of the third research paper. A special thanks to my peers Chen Zhenxi, Le Thaiha and Xie Taojun for all the help during my Ph.D. study. I would like to give a big thank you to my boyfriend Zhao Sixuan for companionship, encouragement and continuous support along the way.

Last but not the least, I thank my beloved parents and brother for unconditional love, understanding and support. They always give me strength to carry on during hard times. The love from my family is indispensable for the completion of my Ph.D.

Table of Contents

List of Tables

List of Figures

Executive Summary

This thesis consists of four self-contained essays on social preferences and strategic interactions.

Chapter 1 explores non-binding default options in voluntary contribution games. It is well documented that people are reluctant to switch from a default option. We experimentally test the robustness of this behavioral inertia by varying the default option type. We examine the impacts of automatic-participation and no-participation default options on subjects' participation in a public goods provision and their contributions. Our experimental results square with the evidence of behavioral inertia only when the automatic-participation default is used. This default boosts contributions in the linear public goods game but not in the threshold public goods game. The evidence of partial stickiness is robust to the variation of the game employed, but the effect on contribution is sensitive to it.

Chapter 2 studies the role of liquidation policy and disclosure of credit history in financial contracting. In the presence of contract incompleteness and asymmetric information, liquidation policy plays an important role in financial contracting. Liquidation is a double-edged sword. It deters borrowers from defaulting strategically, but it could be harsh to borrowers experiencing short-term liquidity problems. This chapter presents an experimental analysis of the impacts of (1) liquidation policy on borrowers' incentive to engage in strategic default and (2) disclosure of credit history information on lending relationships and borrowers' behaviors. We show that liquidation policy deters borrowers from defaulting strategically, and the availability of credit information softens the liquidation policy and helps reduce strategic defaults.

Chapter 3 studies altruistic punishment in the face of direct externality and selfish temptation. By giving the third party an opportunity to misappropriate the punishment-induced windfall money, we investigate to what extent a third party's willingness to punish is motivated by kind intentions. We find that a significant proportion of third parties succumb to this temptation. Interestingly, more altruistic third parties impose lesser punishments suggesting that they are aware of the temptation and want to pre-commit to reducing the misappropriation by reducing the windfall money available. We also explore the motivation behind altruistic punishment. Two broad motives are examined: the retributive motive and the distributive motive. Our results are in line with the retributive motive.

Chapter 4 experimentally investigates the role of information transparency for equilibrium selection in stag hunt coordination games. These games can be transformed from a prisoner's dilemma game by introducing a centralized reward or punishment scheme. We aim to explore the impact of the disclosure of information on how final payoffs are derived from players' incentive to coordinate on the payoff-dominant equilibrium. We find that such information disclosure significantly increases the tendency of players to play payoff-dominant action and reduces the occurrence of coordination failure. The mechanism works directly through the positive impact of disclosure on the saliency of the payoff-dominant equilibrium, and indirectly through the positive influence of disclosure on players' belief about the likelihood of payoff-dominant plays by other players.

Chapter 1 The Limit to Behavioral Inertia and the Power of Default in Voluntary Contribution Games

1.1 Introduction

There is ample empirical evidence in the literature that people are reluctant to switch from a default option. Madrian and Shea (2002), Choi et al. (2004) and Thaler and Sunstein (2003) show that the use of automatic enrollment instead of non-enrollment as the default choice in the 401(k) saving for retirement program increases the enrollment rate because people are inclined to stay with the default option provided. Johnson et al. (1993) show that people are reluctant to switch from the stated default option when choosing insurance policies. In a similar vein, Johnson and Goldstein (2003) show that the use of opt-out schemes in organ donation drives, whereby automatic enrollment is set as the default, boosts the enrollment rate. This behavioral inertia has an important implication for policy design. It suggests that when the objective is to induce people to take a particular choice, policy makers can simply set a desired choice as the default choice.

Several explanations are proposed in the literature to rationalize this behavioral inertia. People might perceive the prevailing default option as a suggestion of the course of action to follow (see Madrian and Shea, 2002; Thaler and Sunstein, 2008; Beshears et al., 2009). People might also feel that switching from the default option requires effort, while not switching is effortless. As a result, they prefer to stick with the default (see Samuelson and Zeckhauser, 1988). Finally, people might be loss averse and reluctant to switch away from the status quo because switching causes a sense of loss relative to the initial reference point (see Kahneman et al., 1991).

One potential drawback of the existing empirical studies is the presence of confounding factors that complicate statistical inference. For instance, with regards to the impacts of the automatic enrollment plan on 401(k) retirement saving behavior, other factors, such as individual characteristics and alternative saving plans, may also potentially affect saving behavior. In this respect, controlled laboratory experiments provide a viable alternative to the existing empirical studies in dealing with these confounding factors. By experimentation, we could also evaluate the robustness of the behavioral inertia found in the existing empirical studies to a variation in the format of the default option employed.

The existing explanations for people's inclination to stay with the default option mentioned above demonstrate that the format of default option employed is inconsequential. That is, the tendency for people to exhibit affinity to the status quo choice is a general phenomenon and would likely persist regardless of the format of the default option employed. For instance, in the case of organ donation, when non-enrollment is employed as the default option, the resulting enrollment rate will be low. When automatic enrollment is employed as the default option, the resulting enrollment rate will be high.

In this chapter, we delve into this issue by carefully investigating whether or not the type of default option employed is indeed irrelevant for people's decision whether or not to stay with the default option through a series of controlled laboratory experiments. To achieve this goal, we vary the type of default option given to the subjects in our experiments. We focus our analysis on public goods experiments and on the default participation option in public goods provision. Specifically, our research objectives are threefold. *Firstly*, we evaluate the robustness of the evidence showing that people tend to stick to the default option. *Secondly*, we investigate the impact of different formats of default option on people's incentive to cooperate in a collective setting. *Thirdly*, we evaluate whether the formats of the public goods game employed influence the ability of the default participation option to affect the level of cooperation.

In particular, we focus on two public goods game formats, namely the *linear public goods* game, where the unique Nash equilibrium is characterized by zero contribution, and the threshold public goods game, where a zero-contribution equilibrium and multiple interior Nash equilibria exist. The presence of a contribution threshold essentially transforms the prisoner's dilemma type situation that exists in the linear public goods game into a coordination problem.

When subjects are confident that others will contribute enough to help meet the threshold, they tend to follow suit. However, when they are not confident, they prefer to choose zero contribution. The consideration of the threshold public goods game would also allow us to further check the robustness of the results obtained under the linear public goods game. That is, under the linear public goods game, the default participation option given under the opt-in setting coincides with the unique Nash equilibrium of the game. However, under the threshold public goods game, the default participation option given under the opt-in setting is just one of the many equilibria of the game.

Some examples in real life resemble the threshold public goods game, for instance, residents are required to collect an unknown number of signatures while lobbying local government for a public project and some fund-raising websites use thresholds to determine whether a cause is funded or not. The most efficient outcome in both games from the collective standpoint requires members to make a full contribution.

A host of experimental evidence has shown that people are willing to make a positive contribution in the linear public goods game. On average, people contribute around $30\% - 40\%$ of their endowment, and the contribution decreases over several rounds of repetition to around 10% of their endowment. Several feasible mechanisms for promoting and sustaining cooperation over the long term have been proposed in the literature. They include, among others, the use of monetary or non-monetary rewards or punishments (see for instance Fehr and Gächter, 2000; Falkinger et al., 2000; Sefton et al., 2007; Nikiforakis, 2008) and the use of social comparison (Nikiforakis, 2010). The contribution rate in the threshold public goods game varies depending on the threshold level employed (see Isaac et al., 1989; Rauchdobler et al., 2010; Laury and Holt, 2008).

If indeed people tend to stick to the default option, then the use of automatic participation in public goods contribution should boost the level of cooperation. Furthermore, the automatic participation default could also potentially act as a coordination device that would subtly lead people to increase their contribution.

We conducted three between-subject experimental treatments for each variant of the public goods game, making six experimental sessions altogether.

The first treatment involves no default option, and serves as our control treatment. The second treatment is the *opt-in* treatment, in which by default people do not contribute to the public goods provision, but can decide to opt-in if they wish to do so. The third is the *opt-out* treatment, where by default people must make a non-zero contribution, but can decide to opt-out if they wish.

To the best of our knowledge, there are only a handful of recent papers to ours. Messer et al. (2007) found that the use of a status quo of giving increases contribution in public goods games initially but the effect is not longlasting. Altmann and Falk (2009) focused on the interplay between the use of the non-binding default option, cognitive skills and the incentive to cooperate in the provision of public goods. They showed that the use of the non-binding option only influenced the behavior of people with low cognitive skills. Fredrik et al., (2011) presented a field experiment of the public goods game in Vietnam. Their subjects are villagers in a rural village who have to make a collective monetary contribution to build a bridge, a vital transportation infrastructure for the villagers. They compared the use of the no default option with zero-contribution and full contribution default options. They showed that, relative to the full contribution default option, the zero-contribution default option decreases the average contribution by around 20% . Cappelletti et al. (2014) explored channels through which defaults work by varying the form of suggested contributions (i.e. either presented as defaults or advice) and the source of the default contribution level (i.e. either set by human subjects or computers).

In the papers mentioned above, with an exception of Cappelletti et al. (2014), the employed default contribution is either to contribute all endowment or not to contribute at all. Thus, in the opt-out treatment, if subjects do not choose to opt out, they are deemed to have agreed to contribute their entire endowment. In the opt-in treatment, if subjects do not decide to opt in, they are deemed to have agreed not to contribute anything. In Cappelletti et al. (2014), the default contribution amount is determined by a random outsider whose payoff is set equal to the average payoff of the group that adopts her suggested default. In this way, the default contribution amount in the opt-out treatment may potentially be less extreme than the one provided in the other papers.

It should be noted that the default option used in the opt-in treatment in those papers is similar to ours, but not the one used in the opt-out treatment. In our design, in the opt-out treatment, if subjects do not opt out of the public goods contribution, they still have to consciously choose their non-zero contribution. Thus, in our study we maintain the voluntary nature of the public goods game within our opt-out treatment by allowing people to decide the amount they wish to contribute. In other words, in our setup, the default option provided is the default *participation* in the public goods contribution and not the default *contribution amount*. In addition, each of these studies focused on one type of public goods game and they found differentiated effects of defaults, while our study focuses on both the linear and threshold public goods games, which allows us to test whether the effects are sensitive to the game format.

This paper contributes to the literature in several ways. *Firstly*, to the best of our knowledge, this is the first study showing the evidence of partial stickiness, rather than complete stickiness, of default options. By manipulating the decision environment (i.e. the absence or presence of default option) and the type of defaults (i.e. participation or no-participation default), we are able to show that under some condition subjects would not always want to stay with the default option provided. *Secondly*, this paper employs the participation and no-participation default options, rather than the contribution amount default options, in the provision of public goods. Thus, the default option provided in this paper still preserves the voluntary nature of public goods provision. That is, even when the default option employed is participation, subjects would still have to decide the (positive) amount of contribution. *Thirdly*, this paper also presents the sensitivity analysis of the impact of default participation option on the public goods contribution to the types of voluntary contribution games employed (i.e., the linear public goods game and the threshold public goods game).

We show that the nature of a default option influences subjects' decision whether or not to stay with it. We also find that the level of contribution under different default options is sensitive to the format of the public goods game employed. We show that, contrary to the existing evidence, subjects in our experiments do not always stick to the default options. They consciously make an active decision to nullify the default option under the opt-in system, but not under the opt-out system. Essentially, under the opt-in system, the default option forces subjects to free-ride on others' contributions. Subjects in our experiments tend to be averse to being labeled free-riders, and would therefore prefer to switch from the default decision. Interestingly, they would then contribute the same amount as is contributed in the baseline setting, that is, in the standard public goods game with no default option. This behavior is robust to the format of the public goods game employed.

The use of the opt-out system also results in significantly higher individual contributions and individuals' earnings beyond that under the opt-in and the baseline settings. However, this only happens in the linear public goods game but not in the threshold public goods game. A possible explanation for this result is that the threshold level itself acts as a powerful coordination device that helps subjects cooperate in meeting the threshold level. As subjects have greater incentive to contribute in the threshold public goods game, the presence of the opt-out option becomes somewhat redundant. In other words, the effectiveness of the cooperative default on the contribution level is sensitive to the structure of the voluntary contribution game.

This chapter is organized as follows. Section 1.2 presents our experimental design and procedures, followed by predictions in Section 1.3. Section 1.4 discusses our experimental results, Section 1.5 provides some explanations to our results and Section 1.6 concludes the chapter.

1.2 Experimental Design and Procedures

We consider two variants of the public goods game, namely the standard linear public goods game, where the unique inefficient Nash equilibrium exists, and the threshold public goods game, where the inefficient equilibrium and a set of efficient equilibria exist, among which the symmetric efficient equilibrium serves as the focal point. In the baseline treatment we do not impose any default option. We compare this baseline treatment with the *opt-out* treatment, where by default subjects are deemed to agree to make non-zero contribution to the public goods provision, and the *opt-in* treatment, where by default subjects do not contribute to the public goods provision. We ran these three experimental treatments in both the linear and the threshold public goods game.

Subjects formed a group of 4 and played the public goods game over 10 (ten) periods. The first 2 periods were trial periods. In every period, subjects were randomly re-matched. Thus, the group composition varied across periods. At the beginning of each round, subjects were given 100 endowment points, which they could allocate to their own private account, or to the group account, or any combination of both.

In the linear public goods game, the payoff for each subject can be expressed as

$$\pi_i = E_i - c_i + \alpha \sum_{i=1}^{4} c_i. \tag{1.1}$$

where, E_i, c_i, and α denote, respectively, player i's initial endowment, which is set at 100 points, player i's contribution to the group account, and the marginal per-capita return (MPCR), which is set at 0.5. Given that the MPCR from placing 1 point in the group account is less than the marginal return from placing it in the individual private account ($E_i - c_i$), i.e., $\partial\pi_i/\partial c_i = -1 + \alpha < 0$, the dominant strategy for a player would be to place all points in the private account and to let others contribute to the group account. Anticipating this behavior, other players would follow suit. As a result, in equilibrium, no one would make any contribution to the group account. Zero contribution is the unique and inefficient Nash equilibrium in this linear public goods game.

In the threshold public goods game, the public goods will be provided if and only if the group contribution threshold (T) is met. If the threshold is not met, the public goods will not be provided and the points contributed in the group account will be refunded subject to the refund rate of r per contribution point. The payoff for each subject (π_i) can then be expressed as

$$\pi_i = \begin{cases} E_i - c_i + \alpha \sum_{i=1}^{4} c_i & if \ \sum_{i=1}^{4} c_i \geq T \\ E_i - c_i + rc_i & if \ \sum_{i=1}^{4} c_i < T \end{cases} \quad (1.2)$$

We set the threshold T equal to 200 points. In the threshold public goods game, there will be multiple pure-strategy equilibria containing the inefficient zero-contribution equilibrium, which also exists in the linear public goods game, and a set of efficient equilibria that includes all possible combinations of contributions where the threshold is exactly satisfied. For instance, if the total contributions of other members add up to 175 points, then an individual has to decide whether or not to contribute 25 points to meet the threshold. Doing so yields 175 points, while contributing any amount less than 25 points yields a payoff that is lower than or equal to 100. Consequently, the best response for this individual is to contribute 25 points. Among these efficient equilibria, there exists a unique symmetric equilibrium where everybody contributes exactly 50 points. This symmetric equilibrium serves as a focal point that helps players coordinate their contribution decisions.

Subjects were informed of the total amount of contributions collected in the group account and their individual earnings from a particular round at the end of that round. One round out of 8 real periods was randomly selected as the binding round to determine their payment. The points earned were converted into Singapore Dollars at the rate of 1 point = 0.10 SGD.

In the opt-out treatment, by default subjects were considered to have chosen to contribute to the group account. This implied that, as long as they did not decide to opt-out from the default option, they were required to make non-zero contribution to the group account. In other words, the default option in this setup is only related to the decision whether or not to participate in the provision of public goods, but not on the amount of contribution. Further, the default option in the opt-out treatment also has a *positive* connotation from the view point of society as it encourages people to participate in public goods provision. For this reason, we label the automatic-participation default-option in the opt-out treatment as the *positive default* in this chapter. In contrast to

Altmann and Falk (2009) and Fredrik et al (2011), the default option in our *opt-out* treatment did not require subjects to contribute their entire endowment to the group account. They could, in principle, contribute an amount that was slightly above zero, which in essence is equivalent to no contribution. Our experimental design, therefore, maintains the voluntary nature of the public goods game.

In the *opt-in* treatment, by default subjects were considered to have chosen not to contribute at all to the group account, but they can opt in to the contribution plan. The zero-contribution default has a *negative* connotation from the society's perspective as it essentially use the free-riding equilibrium as the default option. We label the no-participation default-option in the opt-in treatment as the *negative default* in this chapter.

The experiment was conducted at Nanyang Technological University (NTU). The subjects were undergraduate students from various majors. The linear public goods game experiments were conducted in February, 2011, and were programmed using a web-based interactive program. The threshold public goods game experiments were conducted in October, 2012, and were programmed using Z-tree (Fischbacher, 2007).[1] In total, there were 24 subjects in each treatment, giving a total of 144 subjects. Table 1.1 summarizes the basic descriptive statistics of participants in the two game types.

Table 1.1. Descriptive statistics by game and treatment

Game type	Treatment	Subject	Observation	Mean gender (Male = 1)	Mean age (std. dev.)	Median contribution
Linear	Control	24	192	0.54	21.0 (1.43)	10
	Opt-out	24	192	0.67	20.2 (1.55)	21
	Opt-in	24	192	0.58	21.1 (1.25)	10
Threshold	Control	24	192	0.38	20.29(1.88)	60
	Opt-out	24	192	0.50	20.38(1.71)	65
	Opt-in	24	192	0.38	20.38(1.50)	60

[1]It should be noted that in this chapter we do not compare treatments across games. That is, for example, we do not compare the opt-out treatment in the linear public good game and the opt-out treatment in the threshold public good game.

The instructions were read aloud to the subjects.[2] We gave the participants two trial periods. Once the trial periods were completed, the participants played the public goods contribution game for eight periods. At the end of the experiment, the participants were asked to fill in a post-experiment questionnaire intended to capture their demographic information and the reasons behind their decisions in the experiments.

1.3 Experimental Predictions

In this section, we formulate some experimental predictions. The predictions focus on two aspects, that is, subjects stay with the default options or make active decisions, and the effects of default options on contribution levels relative to those generated in the control treatment whereby default options are absent. We formulate the rational (standard economic theory) based prediction and the behavioral prediction for each aspect.

Prediction 1*a* *Rational based prediction* Subjects faced with the positive default in the opt-out treatment will take active decisions (switch) in the linear public goods game and stay with default options in the threshold public goods game. Subjects faced with negative defaults in the opt-in treatment will always stay with default options in both games.

The unique equilibrium in the linear public goods game is the zero-contribution equilibrium. Standard economic theory suggests that subjects will always make a zero-contribution regardless of the defaults. In other words, subjects tend to go against the cooperative default in the opt-out treatment and stay with the non-cooperative default in the opt-in treatment in order to reach the equilibrium in the linear public goods game.

In the threshold public goods game there are multiple equilibria including the inefficient zero-contribution equilibrium and the efficient interior equilibria. The presence of the threshold transforms the game with a unique equilibrium to a coordination game with multiple equilibria. In such a game, the equilibria can be categorized as either risk dominant or payoff dominant (see Harsanyi and

[2] The experimental instruction that we used in the *opt-out treatment* in the linear public goods game and in the threshold public goods game can be found in the appendix. The experimental instructions for other treatments can be obtained from the authors upon request.

Selten, 1988). A large body of literature has studied the equilibrium selection problem in coordination games both theoretically and experimentally (e.g., Van Huyck et al., 1990, 1991; Kim, 1996; Rankin et al., 2000; Broseta et al., 2003; Schmidt et al., 2003). Yet, these studies cannot really give us any conclusive evidence on whether the risk dominant or the payoff dominant equilibrium is more likely to be selected. In particular, the equilibrium outcome is shown to be sensitive to the game context, the payoff structure, the group size and other factors.

In our threshold game, subjects can choose either the zero-contribution equilibrium or any of the interior efficient equilibria. We posit that people are more likely to coordinate on the equilibrium outcome, which is consistent with the suggested default, because switching entails some psychological costs. Therefore, subjects faced with the positive default tend to coordinate on the efficient equilibrium and those faced with the negative default tend to end up with the inefficient equilibrium.

Prediction $\mathbf{1b}$ *Behavioral based prediction* Subjects will stay with default options in both games, regardless of the type of defaults.

The empirical findings suggest that the default is sticky and people are likely to stay with the exerted default option. Various reasons, ranging from loss aversion, procrastination, etc, have been put forward to explain this behavioral inertia. Thus, regardless of the format of default options employed, subjects will always stay with default options regardless of their type.

Table 1.2 summarizes various prediction categories for prediction 1.

Table 1.2. Summary of predictions of staying or switching regarding default options

		Positive default (Opt-out)	Negative default (Opt-in)
Linear	Rational	Switch	Stay
	Behavioral	Stay	Stay
Threshold	Rational	Stay	Stay
	Behavioral	Stay	Stay

Prediction $\mathbf{2a}$ *Rational based prediction* Default options have no effect on contribution levels.

Since default options do not change the payoff structure of the game nor the equilibrium outcomes, individuals still choose what is the best for them.

Prediction 2*b* *Behavioral based prediction* The positive default in the opt-out treatment increases the contribution level, and the negative default in the opt-in treatment has the opposite effect.

It has been proposed in empirical studies that the default option could be perceived as a suggestion or endorsement from the policy maker. Based on the endorsement hypothesis, subjects will contribute less when faced with negative defaults in the opt-in treatment and contribute more when faced with positive defaults in the opt-out treatment. Table 1.3 summarizes various prediction categories for prediction 2.

Table 1.3. Summary of predictions of the effects of defaults on contribution levels

		Positive default (Opt-out)	Negative default (Opt-in)
Linear	Rational	No effect	No effect
	Behavioral	Positive	Negative
Threshold	Rational	No effect	No effect
	Behavioral	Positive	Negative

1.4 Experimental Results

In this section, we proceed by first giving the summary statistics of the data obtained from our experiments. We then continue by presenting the regression analysis using our experimental data. We focus on two issues in this section; firstly, the incentive of subjects to switch from- or to stay with- the default options, and secondly, the impact of default options on contribution levels. In the analysis, we will relate the results obtained to the predictions presented earlier.

Table 1.4 summarizes the proportion of active decisions and default options taken. The third column represents the percentage of subjects who made an active decision to choose either to opt-out in the opt-out treatment, or to opt-in in the opt-in treatment. The fourth column shows the percentage of subjects who stick to the default option. In the opt-out treatment, the overwhelming majority of subjects (87% and 95.8% in the linear and threshold public goods games, respectively) stick to the default option and contribute a non-zero amount to the group account. In contrast, in the opt-in treatment, the overwhelming majority of subjects (82.8% and 94.3% in the linear and threshold public goods games, respectively) do not stick to the default option and instead prefer to take an active decision to contribute.

Table 1.4. Summary of decisions and contributions across treatments

Game type	Treatment	Active decisions	Stay with default	Average contribution	Conditional average contribution	Positive contributions
Linear	Control	–	–	21.13	34.67	60.9%
	Opt-out	13.0%	87.0%	32.82	37.73	87.0%
	Opt-in	82.8%	17.2%	19.35	23.36	82.8%
Threshold	Control	-	-	58.56	61.11	95.8%
	Opt-out	4.2%	95.8%	60.99	63.65	95.8%
	Opt-in	94.3%	5.7%	63.73	67.61	94.3%

This result suggests that people did not always blindly follow the default option. When given an option to contribute nothing, which in itself is the unique equilibrium outcome in the linear public goods game and one of the equilibrium outcomes under the threshold public goods game, most subjects did not like it. They made a deliberate decision to contribute. This behavior is universal across the public goods game formats. In other words, the default options are partially sticky. All in all, our results showed that the nature of default option employed influences subjects' decision whether or not to stay with the default option.

Interestingly, in the opt-in treatment in the linear public goods game, the average contribution was 19.35 points. This was remarkably similar to the average contribution in the control treatment in the linear public goods game, which was 21.13 points. This suggests that subjects, to some extent, seemed to care genuinely about others' well-being and were therefore willing to show some level of cooperation by making an active decision to move away from the default inefficient (free-riding) equilibrium.

In the opt-out treatment in the linear public goods game, the average contribution was 32.82 points, which was around 55% more than the average contribution in the control treatment. Prediction 2a and prediction 2b find their partial support in the linear public goods game, depending on the default option employed. It appears that setting the participation default and allowing subjects to opt-out from participating if they wish to do so encouraged subjects to make larger contributions. Rather than opting out from participating in the public goods provision, they preferred to stick to the default and even contribute more.

Besides the average contribution, we also present the conditional average contribution and the percentage of non-zero contributions across treatments. The presence of the default option, regardless of its nature, boosts the participation rate in the linear public goods game relative to the control treatment. That is, in the opt-out treatment, around 87% of subjects made non-zero contribution by sticking to the default option, while in the opt-in treatment, around 82% in the opt-out treatment made non-zero contribution by switching away from the default option. In contrast, there are only around 61% of subjects made non-zero contribution. However, even though the use of the opt-in format in the linear public good game increased the participation rate, the amount of contribution per subject is the lowest among the other treatments including the control treatment. The increased in the average contribution in the opt-out treatment comes from both the higher participation rate and the higher contribution level.

A different picture emerged in the threshold public goods game. The overall average contribution in the threshold public goods game was around 61 points. The average contributions across all three treatments were even higher than the average contribution at the focal symmetric equilibrium outcome where everybody contributes 50 points to meet the threshold of 200 points. The average contributions across all three treatments were similar. Our results in the threshold game are consistent with prediction 2a but not with prediction 2b. Thus, in general, subjects showed a stronger motivation to contribute in the threshold public goods game regardless of the nature of the default option employed. This is perhaps due to the fact that in the threshold public goods game some combinations of subjects' positive contributions that would meet the threshold exactly also constitute an equilibrium. Consequently, there is always an inherent motivation for subjects to contribute to the public goods provision and to coordinate their contributions in order to meet the threshold. All in all, our results suggest that in the threshold public goods game where the threshold itself acts as a powerful coordination device, non-binding defaults have no effect in promoting cooperation.

Figure 1.1 depicts the average contribution across periods (excluding the first two trial periods). Panel A shows that the average contribution in the opt-out treatment in the linear public goods game dropped in the early periods, but subsequently remained relatively consistent across periods. In contrast, in the other two treatments, the average contribution in the first period started at a substantially lower level than that in the opt-out treatment, and then decayed across periods. The decaying trend is consistent with findings in the existing literature (e.g., Andreoni, 1995; Palfrey and Prisbrey, 1997; Houser and Kurzban, 2002). When we ran the Wilcoxon signed-rank test comparing the average contributions in the first three periods and those in the last three periods, we found that they were not statistically different (p-value $= 0.255$). In contrast, the test result showed that the difference in the average contributions between the first three periods and the last three periods was significant in the other two treatments (p-value < 0.01 in both treatments). It is apparent from the graph shown in Panel A that, in addition to promoting contributions, the opt-out treatment helps sustain contributions as well.

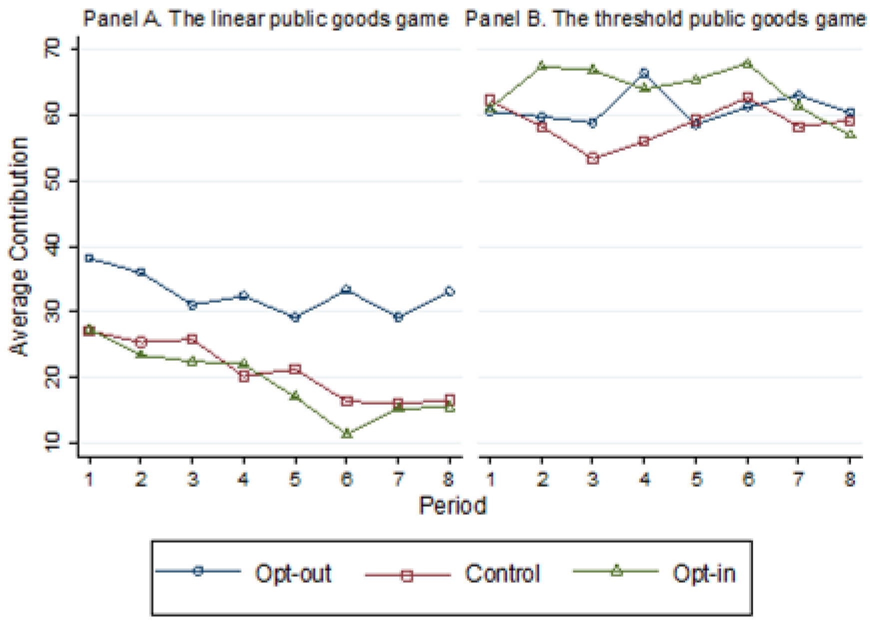


Figure 1.1. Average contribution over time

The graph shown in Panel B, in contrast, shows that the average contribution in all treatments in the threshold game was relatively stable across periods. The average contribution started at around 60 points, which was much higher than the starting average contribution in all treatments in the linear public goods game. The Wilcoxon signed-rank test, comparing the difference between average contributions in the first three periods and in the last three periods in all treatments, confirms that the difference was not statistically significant (p-value ≥ 0.55 in all three treatments).

Figure 1.2 shows the distribution of subjects' probability of switching away from the default option provided in the opt-out and opt-in treatments. Panels A and B present the linear and threshold public goods games, respectively. This probability is defined as the ratio of the number of periods wherein an active decision to nullify the default was taken by subjects to the total number of periods. It can be seen that the probability distribution of switching in the opt-out treatment was markedly different from that in the opt-in treatment, regardless of the format of the public goods game employed. This confirms our earlier result. The nature of the default option does matter for subjects' decisions whether or not participate in the public goods contribution.

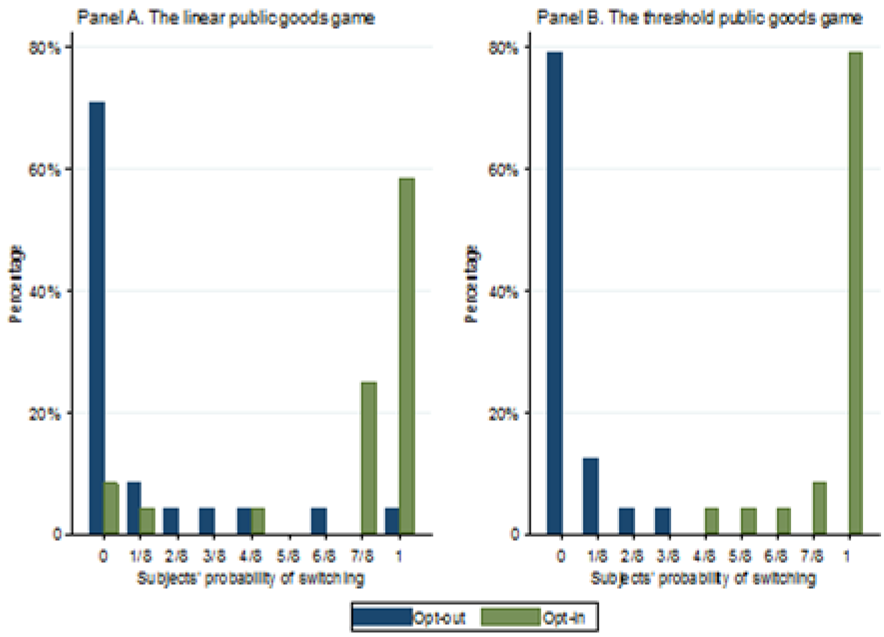


Figure 1.2. The distribution of subjects' probability of switching

About 70% of subjects in the opt-out treatment never switched from the default participation option in the linear public goods game. The proportion stands at around 80% in the threshold public goods game. In contrast, about 60% of subjects in the opt-in treatment always switched from the default participation option in the linear public goods game. The proportion was much higher (around80%) in the threshold public goods game. It is interesting to note that the distributions of the probability of switching in both treatments in the threshold public goods game did not really overlap. This further corroborates our earlier finding that people seemed to have stronger motivation to contribute to public goods in the threshold public goods game.

All in all, the evidence suggests that subjects have a natural tendency to contribute to public goods. When they were considered not to participate in the provision of public goods by default, subjects consciously decided to move away from the default. However, when the default option is consistent with their natural intention to participate in the provision of public goods, subjects stayed with the default option.

Figure 1.3 below presents the distributions of subjects' average contributions in all treatments in both games. The vertical axis represents the percentage of subjects whose average contributions fall into the range of average contribution specified in the horizontal axis. It can be seen that the distributions in the two public goods games were substantially different.

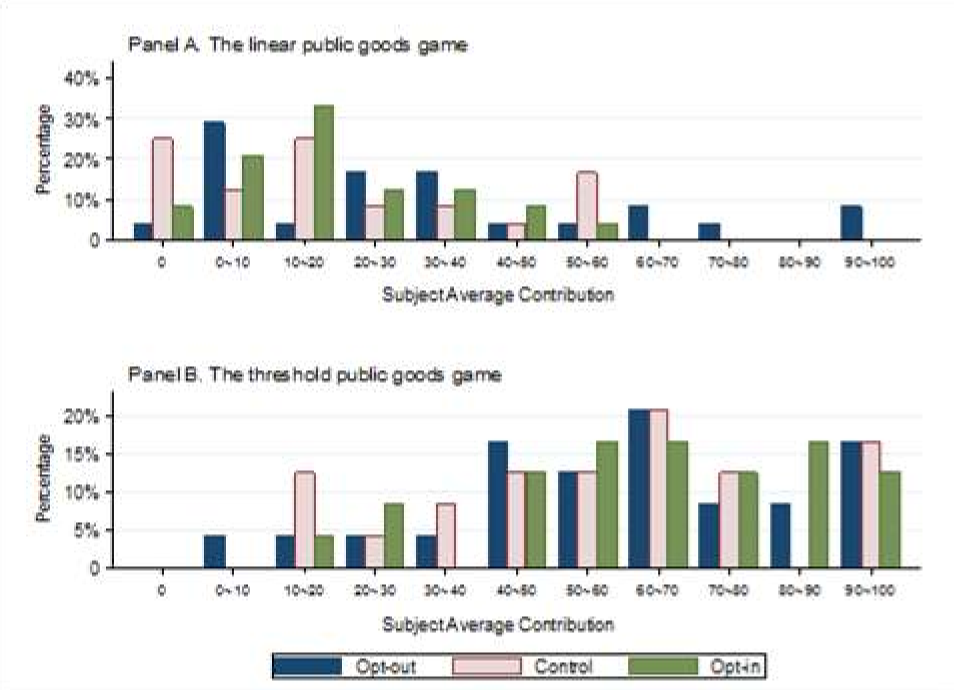


Figure 1.3. The distribution of subjects' average contributions

Note that in the linear public goods game shown in Panel A, the zero contribution bar in the opt-out treatment represents the proportion of subjects who always decided to opt-out from participating in the public goods contribution in all periods, while that in the opt-in treatment represents the proportion of subjects who always decided to stay with the default option of no participation. It is interesting to see that these percentages are relatively close to each other, and are markedly different from the percentage of subjects who never decided to contribute in the control treatment where the default option was absent.

For each subject, we count the number of times he/she participated in the provision of public goods in all eight periods. The null hypothesis, which states that the number of participations in the provision of public goods in the opt-out and control treatments are equal, is rejected (the p-value from the pair-wise two-sided Mann-Whitney test is 0.02). This is also the case for the comparison between the opt-in and control treatments (p-value $= 0.075$). However, the null hypothesis is not rejected for the comparison between the opt-in and opt-out treatments (p-value $= 0.46$). This implies that the participation rates are similar in the opt-in and opt-out treatments, despite the differing default options provided. It also suggests that people consciously evaluated the default option and decided whether or not to stay with the default option. Thus, depending on the nature of the default option provided, people may or may not exhibit behavioral inertia, contrary to the prevailing popular argument on the use of default options.

Panel A also shows that in the opt-out treatment roughly around 25% of subjects contributed more than 50 points, while in the control and opt-in treatment only around 18% and 4% of subjects did so, respectively. When we compare the opt-in and control treatments, we can see that the distribution of contributions in the opt-in treatment tended to shift to the right relative to that in the control treatment.

Panel B shows the distribution of contributions across treatments in the threshold public goods game. It is interesting to see that subjects' average contributions are clustered around a higher contribution range in all treatments. It can also be seen that no subject contributed zero, even in the control treatment. This suggests that the presence of the threshold generally decreased selfish behaviors. The use of the default option did not seem to increase contributions in the threshold public goods game. This could be because the threshold itself acted as a strong coordination device helping subjects to focus on meeting the threshold. We also found that the participation rates in all treatments were similar (the p-values from the pair-wise two-sided Mann-Whitney test are, respectively, 0.54 for the opt-out vs control treatments, 0.48 for the opt-in vs control treatments, and 0.92 for the opt-out vs opt-in treatment).

Again, if we look at the zero contribution case in Panel *B*, we observe that no subject in the opt-out treatment decided to move away from the default option in all eight periods. Likewise, no subject in the opt-in treatment decided to stay with the default option in all eight periods. This evidence once again confirms that subjects made conscious decisions whether or not to stay with the default option. When presented with a non-participation default option, subjects actively decided to nullify the default.

Next, we further examine the determinants of active decision, individual contribution, group contribution, and individuals' earnings through a series of regression analyses. Table 1.5 presents the probit estimates of the determinants of active decision, which takes a value of 1 if the subject goes against the default option to make a zero-contribution in the opt-out treatment or to make a positive contribution in the opt-in treatment, and takes a value of 0 otherwise. Panels *A* and *B* represent separate regression results for the linear public goods game and the threshold public goods game, respectively. The explanatory variables include the treatment dummy for the opt-out treatment (*Opt-out treatment*) with the opt-in treatment being the baseline treatment, contribution in the previous period (*Contribution (t-1)*), the interaction term between opt-out treatment dummy and contribution in the previous period (*Opt-out * Contribution (t-1)*), individuals' earnings in the previous period (*Earnings (t-1)*), time trend (*Period*), age (*Age*), and gender (*Gender*) taking the value of 1 for male and 0 otherwise. In addition, an indicator variable which takes a value of 1 if the threshold in the previous period was reached and 0 otherwise is included in the panel *B* regression. The lagged variables are included to capture the effect of past outcomes on the contribution decision in the current period. The table also presents marginal effects for average individuals at the sample mean of regressors and the marginal effects for the binary variables report the probability change when the binary variable changes from 0 to 1. As pointed out by Ai and Norton (2003) and Norton et al. (2004), the conventional way of calculating marginal effects and standard errors for interaction terms could be misleading in non-linear models. Therefore, we employ the alternative method proposed in their paper.

The signs of the treatment dummy and the interaction term with the treatment dummy are negative indicating that, relative to the opt-in treatment, subjects in the opt-out treatment are less likely to take an active decision in both the linear public goods game and the threshold public goods game. This confirms our earlier conclusion that the nature of the default option matters. Subjects are more likely to stay with the default option when it is consistent with their inherent motivation, instead of blindly following it. Individual contribution in the previous period has significant impacts on the probability of taking active decisions in the opt-in treatment for both games. Subjects who made a higher contribution in the previous period are more likely to take active decisions in the opt-in treatment, which is valid in both games. The individual contribution in the previous period has significantly negative effects on the probability of taking active decisions in the opt-out treatment in the linear public goods game, while it has no significant impact in the threshold public goods game.

Table 1.5. Determinants of active decision: probit estimation

	Panel A. Linear		Panel B. Threshold	
	Dependent variable: active decision			
	Probit	Marginal effects	Probit	Marginal effects
Opt-out treatment	-1.377***	-0.540***	-1.191***	-0.471***
	(0.421)	(0.168)	(0.424)	(0.169)
Contribution (t-1)	0.019**	0.007**	0.039***	0.015***
	(0.009)	(0.004)	(0.005)	(0.002)
Opt-out × Contribution (t-1)	-0.037***	-0.008	-0.047***	-0.004
	(0.013)	(0.018)	(0.008)	(0.057)
Earnings (t-1)	0.004	0.002	0.006	0.002
	(0.003)	(0.001)	(0.008)	(0.003)
Period	0.007	0.003	-0.124**	-0.049**
	(0.044)	(0.017)	(0.049)	(0.020)
Age	0.045	0.018	-0.100	-0.039
	(0.109)	(0.043)	(0.095)	(0.038)
Gender	0.073	0.028	-0.110	-0.043
	(0.296)	(0.116)	(0.324)	(0.128)
Threshold (t-1)			-0.892	-0.352
			(0.817)	(0.324)
Constant	-0.957		2.333	
	(2.254)		(2.152)	
Observations	336	336	336	336

Note: Standard errors reported in parentheses are clustered at subject level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 1.6. Determinants of contribution: pooled OLS estimation

	Panel A. Linear	Panel B. Threshold
	Dependent variable: contribution	
Opt-out treatment	17.010**	2.241
	(7.509)	(6.790)
Opt-in treatment	-2.479	5.653
	(4.242)	(6.720)
Earnings (t - 1)	-0.298***	-0.160***
	(0.072)	(0.043)
Period	-1.856***	0.047
	(0.473)	(0.477)
Age	1.316	4.134**
	(1.922)	(1.738)
Gender	-7.262	1.626
	(4.809)	(5.757)
Threshold (t-1)		24.332***
		(6.469)
Constant	42.226	-22.674
	(38.107)	(34.761)
Observations	504	504
R^2	0.168	0.126

Note: Standard errors reported in parentheses are clustered at subject level.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table 1.6 presents the pooled OLS estimates of the determinants of individual contribution.[3] Panels *A* and *B* present the linear and threshold public goods games, respectively. The regressors include: treatment dummies for the opt-out treatment (*Opt-out treatment*) and the opt-in treatment (*Opt-in treatment*), individuals' earnings in the previous period (*Earnings (t-1)*), time

[3] The analysis of contribution here is not conditional on the participation decision. Unless otherwise stated, the contribution is treated in this way throughout this chapter. We also tried the estimation of the participation decision (Probit) and the contribution conditional on participation in both games. In the linear game, the presence of both default types increases the propensity of contribution while only the default in the opt-out treatment has positive effects on the contribution level conditional on participation. In the threshold game, the presence of defaults mostly does not have impact except for the puzzling positive effect of the default option in the opt-in treatment.

trend (*Period*), age (*Age*), gender (*Gender*) which takes the value of 1 for male and 0 otherwise, and a dummy variable indicating whether the threshold was met in the previous period (*Threshold (t-1)*). The control treatment serves as our baseline. We also ran Tobit regressions with the same control variables, and they produced qualitatively the same results to those obtained from the OLS regressions.[4]

It can be seen that being in the opt-out treatment led to an increase of around 17 points in contribution relative to being in the control treatment. This impact is statistically significant at the 5% level. In contrast, being in the opt-in treatment did not have any significant effect on contribution. These regression results are consistent with the previous findings presented in the earlier parts of this section. It is important to note that if subjects preferred to stay with the default option in our opt-out treatment, they would need to make a non-zero contribution to the public goods provision. The amount of contribution that subjects had to make was not specified so as to maintain the voluntary nature of contributions made by subjects. This is an important aspect that distinguishes our setup from those in the studies conducted by Altmann and Falk (2009) and Fredrik et al. (2011). In other words, our default option is on participation in the provision of public goods, rather than on the amount of contribution.

The regression result presented in Panel *B* also confirms our previous results. In the threshold public goods game, the use of default participation options did not have any impact on subjects' contributions. As mentioned earlier, this could be due to the fact that the threshold itself serves as a powerful coordination device for subjects to make a collective contribution that would meet the threshold. This effect rendered the use of a non-binding default option to boost contributions redundant.

Table 1.7 shows the estimates of the pooled OLS regressions on individual earnings. Panels *A* and *B* present the linear and threshold public goods game, respectively.[5] It can be seen that being in the opt-out treatment, where participation in the public goods provision was given as the default option, led

[4]The Tobit regression results are available upon request.

[5]We also ran random effects panel data regressions and we obtained qualitatively similar results. These are available upon request.

to an increase of around 5% in individual earnings. This suggests that the use of the opt-out treatment with its cooperative default option in the linear public goods game helped to some extent to soften the social dilemma problem.

Interestingly, individual earnings were affected by group composition. When a subject was grouped together with generous group members, the individual earnings were higher. A 1 point increase in the contribution of other group members in the previous period would, on average, increase individual earnings by around 1.1 and 1.6 points in the linear and threshold public goods game, respectively.

Table 1.7. Determinants of earnings: pooled OLS estimation

	Panel A. Linear	Panel B. Threshold
	Dependent variable: earning	
Opt-out treatment	5.430***	2.346
	(1.519)	(3.794)
Opt-in treatment	-1.120	0.860
	(1.357)	(3.278)
Contribution	-0.529***	-0.099*
	(0.023)	(0.056)
Period	0.078	-0.239
	(0.298)	(0.789)
Gender	2.046*	0.293
	(1.110)	(3.107)
Others' average contribution (t-1)	1.121***	1.646***
	(0.050)	(0.111)
Constant	105.307***	53.654***
	(2.694)	(8.699)
Observations	504	504
R^2	0.722	0.382

Note: Standard errors reported in parentheses are clustered at subject level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

1.5 Discussions

The evidence presented in the preceding sections indicates that people do respond differently according to the type of default in both games instead of blindly following it. However, the effects of non-binding defaults on contribution in the voluntary contribution game depend to a large degree on context. It is noteworthy that we only specify the nature of defaults, that is, the socially desirable and the less desirable default. Subjects still have to decide how much to contribute if they stick with the cooperative default. In other words, the default exerted in our experiment is a hybrid system of active decisions and regular defaults which require no additional decision if those default options are taken.

The experimental results indicate that people do not blindly follow non-binding defaults and that they consciously evaluate the defaults. In the voluntary contribution game, whether they take active decisions depends on their motivation to contribute. The probability of staying with the cooperative default and the probability of taking active decisions against the uncooperative default are both higher when subjects have stronger motivation to contribute. On one hand, non-binding defaults could be influential in affecting the contribution amount; on the other, there are certain limitations. The nature of default options matters in terms of affecting contributions in the linear public goods game, even when the games played are essentially the same. Our experimental results show that the cooperative default substantially increases contributions while the uncooperative default genuinely has no impact on contribution in the linear game. In addition, both types of default options increase the participation rate in contribution and the participation rates under the two default schemes are comparable in the linear game. However, we find no effect of defaults on individual contribution in the threshold public goods game, which sheds light on the limitation of defaults. The threshold itself acts as a powerful coordination device and the effect dominates the possible effects of non-binding defaults.

Madrian and Shea (2002) propose several possibilities in explaining the impacts of automatic enrollment on 401(k) saving behavior. These include status quo bias because of procrastination, endowment effect, anchoring, framing of decision, and perception of the automatic enrollment as a suggestion from the setter. Beshears et al. (2009) also discussed two types of procrastination and the default option as an endorsement. We now discuss status quo bias and the default acting as suggestions, which are more plausible in our experimental context. If status quo bias is the dominant effect, subjects in the opt-out and the opt-in treatment tend to stick with the default options and thus the number of default options taken in these two treatments are expected to be similar. However, the proportion of the default option taken across periods differs between the opt-out treatment and the opt-in treatment in both games, which contradicts the prediction resulting from status quo bias. If the default options act as implicit suggestions, the predicted results are similar to those resulting from status quo bias. The experimental results also reject the suggestion hypothesis.

Experimental study makes it possible to examine the effects of various default options in a controlled setting, which is a substantial advantage of our study over existing empirical studies. Our results show that people do evaluate the default options consciously and respond differently according to their nature. The intention to contribute is not affected by the type of default options implemented, which explains why the participation rates in the opt-out and opt-in treatments are similar and why they are higher than the participation rate in the control treatment in the linear public goods game.

However, the amounts contributed are indeed affected by the nature of the default options and the effects are asymmetric in the linear game. If people start with the cooperative default, the motivation to contribute is strengthened and they tend to contribute more. Conversely, if people start with the uncooperative default, they take active decisions to go against the default scheme and make positive contributions. Moreover, the amounts contributed are just as high as those in the control treatment. Consider the control treatment as the reference, the cooperative default and the uncooperative default can be classified as positive and negative defaults, respectively. The evidence indicates that if people are put in the positive default, they tend to stay with the positive default and further strengthen their pro-social behavior. However, if people are negatively positioned initially, instead of strengthening the negative trend, they switch from the behavior indicated by the negative default to do what they would in normal circumstances. The above mechanism applies to the linear public goods game where there is no anchoring point when subjects are deciding how much to contribute. In the threshold public goods game, the existence of the focal point (i.e. equal contribution among group members when the threshold is met exactly) dominates the possible influence of non-binding defaults. As a result, the cooperative default does not have the same positive effect on individual contribution in the threshold game as that in the linear game.

The differentiated effects of default options might depend on whether people have well-defined preferences. Default options could be powerful in the absence of well-formed preferences. It is cost-minimizing to stay with the default if people don't have a clear preference for particular outcomes. Otherwise, default options may only have limited influence given that the switching cost is not too large. This is supported by the fact that people take active decisions when faced with the zero contribution default as making contribution is preferred to free-riding. A possible direction for future research is to elicit preferences and explore the effect of defaults in response to heterogenous preferences. Whether this could be generalized in other domains opens for future research.

1.6 Concluding Remarks

In this chapter, we tested the robustness of the evidence showing that people are reluctant to switch from a given non-binding default option and evaluated the impact of non-binding default participation options on the incentive of people to cooperate in the provision of public goods.

Moreover, we examined the extent to which these defaults affect cooperative behavior in different contexts. In particular, we focused on the voluntary contribution game in which people collectively provide public goods for the benefit of everybody in the group. In such a setting, one person's decision has a bearing on other people's well-being. We implemented a linear public goods game where a unique inefficient equilibrium existed and a threshold public goods game in which there were an inefficient zero-contribution equilibrium and a set of efficient equilibria. The most efficient outcome for the group as a whole involves all group members making a full contribution; unfortunately, the socially optimal outcome is hard to achieve.

We conducted three experimental treatments for each game. The first was the control treatment where no defaults are exerted. The second was the opt-in treatment in which subjects are considered to have preferred not to contribute by default. If they wished to contribute, they could decide to opt-in. The third was the opt-out treatment in which subjects are considered to have agreed to make a non-zero contribution by default and had to decide how much to contribute if they stuck with the cooperative default, but could decide to opt-out if they wished to do so.

We show that the nature of default options matters in both games. Contrary to the existing status quo bias and loss aversion rationales for the use of default options, subjects in our experiments do not always stick with the default options. They make active decisions to nullify the default option under the opt-in system, but not under the opt-out system. Essentially, under the opt-in system, the default option forces subjects to free-ride on others' contributions. Subjects in our experiments tended to be averse to being labeled free-riders, and preferred to switch from the default decision. Active decisions against the zero-contribution default indicated that subjects made contributions because of intrinsic motivation rather than confusion.

Interestingly, subjects then contributed the same amount in the opt-in treatment as they would under the baseline setting, that is, in the standard public goods game without any default option. This behavior is robust to the format of the public goods game used. We also found that the use of the opt-out system results in significantly higher individual contributions beyond those under the opt-in system and the baseline. However, this only happens in the linear public goods game, not in the threshold public goods game. A possible explanation for this result is that the threshold level itself acts as a powerful coordination device helping subjects cooperate in meeting the threshold level. As subjects have a greater incentive to contribute in the threshold public goods game, the presence of the opt-out option becomes somewhat redundant. Our study reveals the power as well as the limitations of the cooperative default in promoting cooperation. As defaults are partial sticky and can be influential in promoting cooperation, one possible direction for future research is to study a default system specifying different concrete contribution levels and explore the optimal default contribution level in voluntary contribution games.

1.7 Appendix: Experimental Instructions

1. Linear Public Goods Game

INSTRUCTIONS

General Information

You are now taking part in an interactive study on decision making. **Please pay attention to the information provided here and make your decisions carefully. If at any time you have questions to ask, please raise your hand and we will attend to you in private.**

Please note that unauthorized communication is prohibited. Failure to adhere to this rule would force us to stop the simulation and you may be held liable for the cost incurred. You have the right to withdraw from the simulation at any point, if you decide to withdraw, payments earned will be forfeited.

By participating in this study, you can potentially earn a considerable amount of money. The amount depends on the decisions you and others make.

At the end of this session, this money will be paid to you privately and in cash. It will be sealed in an envelope (marked with your unique user ID) together with a receipt form for you to acknowledge that you have been given the correct payment amount. After signing, please put the receipt form back into the envelope.

General Instructions

Each of you will be given a unique user ID required for logging on to the computer terminal. Your **anonymity will be preserved** for the study. You will **never be aware of** the personal identities of other participants **during or after** the study. Similarly, other participants will also **never be aware** of your personal identity **during or after** the study. You will be identified solely by your user ID in our data collection. All information collected will **strictly be kept confidential** and preserved for the sole purpose of this study.

We have randomly assigned you to a group. In each group, there are four participants (including you). The composition of your group will **differ from round to round**. Participants in each group will interact with other group members for several rounds of decision making. Your earnings will depend on the decisions you make and the decisions made by the other members of your group.

Specific Instructions

At the beginning of each round, you and the other members of your group will be given **100 endowment points**. There are two types of account to which you can allocate your endowment points. The first is a PRIVATE ACCOUNT and the second is a GROUP ACCOUNT (COMMON POOL).

For every **1 point** you place in your PRIVATE ACCOUNT you will earn **1 point**. Thus, if you choose to place ALL your endowment points in your PRIVATE ACCOUNT you will earn 100 points in that round.

For every **1 point** placed in the GROUP ACCOUNT you will earn **0.50 points.** **All members** of the group can place points in the GROUP ACCOUNT.

The amount of money you earn in each round depends on how many points you place in your private account, how many points you place in the group account and the total points placed in the group account.

Your earnings will be calculated in points. They will be converted into Singapore Dollars at the rate of

1 point = 0.10 SGD (10 cents).

Your earnings (and each group member's earnings) from each round are calculated as follows:

- Your **earnings from the PRIVATE ACCOUNT** = (100 – Your contribution to the GROUP ACCOUNT)
- Your **earnings from the GROUP ACCOUNT** = 0.50*(total contributions by your group members in the GROUP ACCOUNT)

Your TOTAL earnings for each round will be calculated as:

0.10*[(100 – Your contribution to the GROUP ACCOUNT) + 0.50*(total contributions by all four group members in the GROUP ACCOUNT)]

This amount will be rounded up or down to the nearest 50 cents.

There will be **two practice rounds** before you begin the actual study to familiarize you the whole procedures and the earnings calculation.

You will have to make the following steps in your decisions.

Step 1

By default, you are considered to have chosen to contribute to the GROUP ACCOUNT. This implies that you **MUST** make a **NON-ZERO** contribution to the GROUP ACCOUNT (COMMON POOL).

If you <u>DO NOT WANT to contribute</u> to the GROUP ACCOUNT, you will have to indicate this explicitly by ticking the appropriate box shown on your screen.

<u>Step 2</u>

Read carefully.

- If you HAVE indicated in step 1 that you do not want to contribute to the GROUP ACCOUNT, this is the end of the round for you. You will receive as your earnings from this round all the points you have placed in the PRIVATE ACCOUNT.

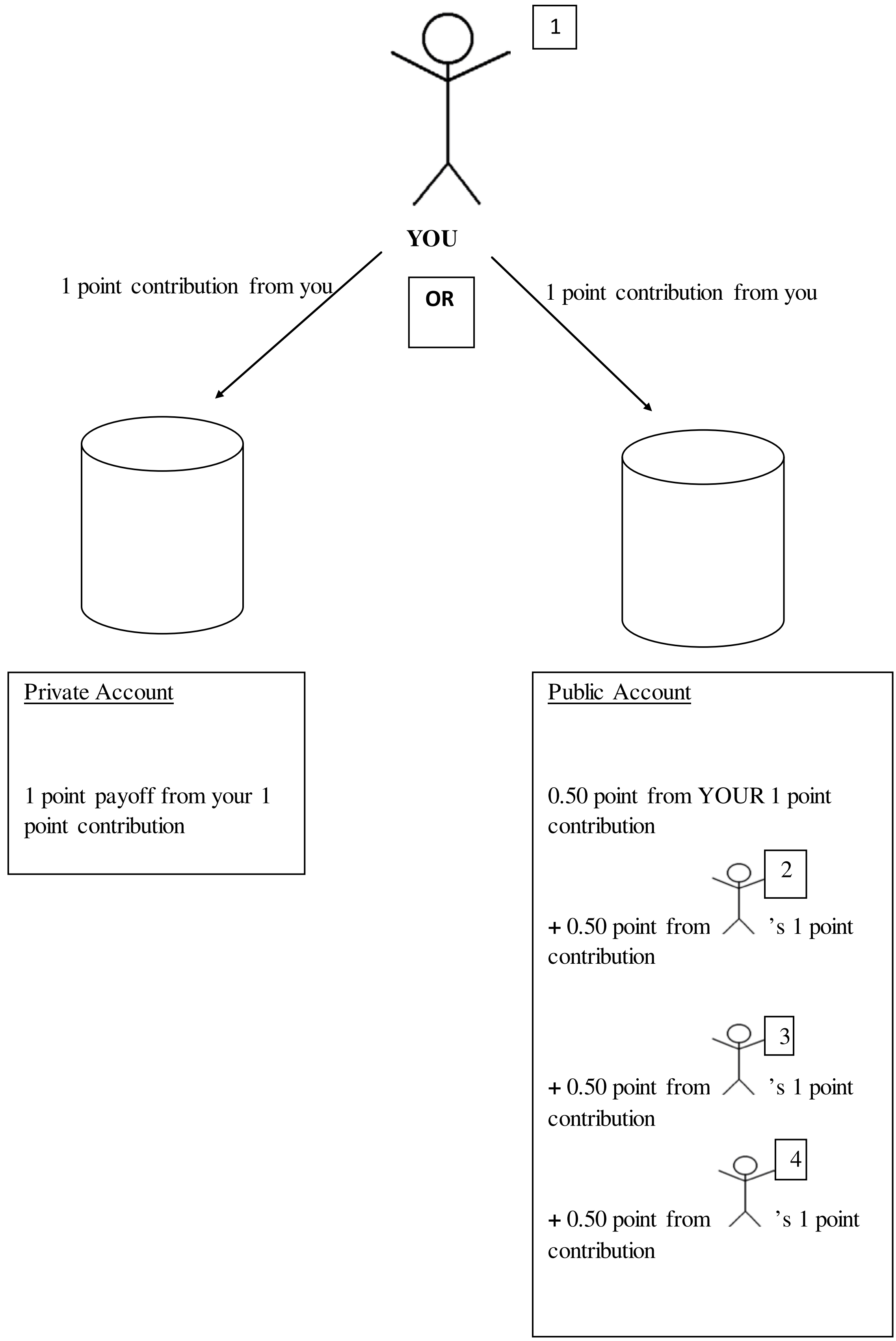
- If you HAVE NOT ticked the box indicating that you do not want to contribute to the GROUP ACCOUNT, you must decide how many points you would like to place in the GROUP ACCOUNT. **Note that you will not be allowed to place 0 points in the GROUP ACCOUNT**

After all participants have made their decisions for the round, the computer will tabulate the results. You will be informed of the total contribution to the group account and your total earnings for the round. This is reflected on your computer screen.

This is the end of the decision making process for one round. After this, a new round begins.

Do refer to the above steps as a guide. You will have to do the same exercise for several rounds. **At the end of the experiment, one of the rounds will be randomly selected as a binding round. Your actual payment will be based on the decision you made in this binding round.**

Illustration



2. Threshold Public Goods Game

General Information

You are now taking part in an interactive study on decision making. **Please pay attention to the information provided here and make your decisions carefully. If at any time you have questions to ask, please raise your hand and we will attend to you in private.**

Please note that unauthorized communication is prohibited. Failure to adhere to this rule would force us to stop the simulation and you may be held liable for the cost incurred. You have the right to withdraw from the simulation at any point, if you decide to withdraw, payments earned will be forfeited.

By participating in this study, you can potentially earn a considerable amount of money. The amount depends on the decisions you and others make.

At the end of this session, this money will be paid to you privately and in cash. It will be sealed in an envelope (marked with your unique user ID) together with a receipt form for you to acknowledge that you have been given the correct payment amount. After signing, please put the receipt form back into the envelope.

General Instructions

Each of you will be given a unique user ID shown on your computer terminal. Your **anonymity will be preserved** for the study. You will **never be aware of** the personal identities of other participants **during or after** the study. Similarly, other participants will also **never be aware** of your personal identity **during or after** the study. You will be identified solely by your user ID in our data collection. All information collected will **strictly be kept confidential** and preserved for the sole purpose of this study.

We have randomly assigned you to a group. In each group, there are four participants (including you). The composition of your group will **differ from round to round**. Participants in each group will interact with other group members for several rounds of decision making. Your earnings will depend on the decisions you make and the decisions made by the other members of your group.

Specific Instructions

At the beginning of each round, you and the other members of your group will be given **100 endowment points**. There are two types of account to which you can allocate your endowment points. The first is a PRIVATE ACCOUNT and the second is a GROUP ACCOUNT (COMMON POOL).

For every **1 point** you place in your PRIVATE ACCOUNT you will earn **1 point**. Thus, if you choose to place ALL your endowment points in your PRIVATE ACCOUNT you will earn 100 points in that round.

All members of the group can place points in the GROUP ACCOUNT. For every **1 point** placed in the GROUP ACCOUNT regardless of who placed it you will earn **0.50 points IF the points in the group account that you and other members contributed equal or exceed 200 points.**

However, **IF the points in the group account that you and other members contributed are below 200 points, no one would earn any point from the group account,** and **instead** the points raised in the group account would be **refunded** back to you at the rate of **0.50 point for every 1 point** you raised in the group account.

The amount of money you earn in each round depends on: 1) the total points placed in the group account by you and the other members and 2) the remaining points in your private account after you contributed into the group account.

Your earnings will eventually be converted into Singapore Dollars at the rate of

1 point = 0.10 SGD (10 cents).

Your earnings (and each group member's earnings) from each round are calculated as follows:

<u>If the points in your group account equal or exceed 200 points</u>

- Your **earnings from the PRIVATE ACCOUNT** = 100 – Your contribution to the GROUP ACCOUNT

- Your **earnings from the GROUP ACCOUNT** = 0.50*(total contributions by your group members in the GROUP ACCOUNT)

<u>Your TOTAL earnings for each round will be calculated as:</u>

S$0.10*[(100 – Your contribution to the GROUP ACCOUNT) + 0.50*(total contributions by all four group members in the GROUP ACCOUNT)]

This amount will be rounded up or down to the nearest 50 cents.

<u>If the points in your group account lie below 200 points</u>

- Your **earnings from the PRIVATE ACCOUNT** = 100 – Your contribution to the GROUP ACCOUNT

- The points you placed in **the GROUP ACCOUNT** will be refunded. The amount of refund would be 0.50* Your contribution to the GROUP ACCOUNT

<u>Your TOTAL earnings for each round will be calculated as:</u>

S$ 0.10*[(100 – Your contribution to the GROUP ACCOUNT) + 0.50 * Your contribution to the GROUP ACCOUNT]

This amount will be rounded up or down to the nearest 50 cents.

There will be **two practice rounds** before you begin the actual study to familiarize you with the whole procedures and the earnings calculation.

You will have to make the following steps in your decisions.

Step 1

By default, you are considered to have chosen to contribute to the GROUP ACCOUNT. This implies that you **MUST** make a **NON-ZERO** contribution to the GROUP ACCOUNT (COMMON POOL).

If you DO NOT WANT to contribute to the GROUP ACCOUNT, you will have to indicate this explicitly by ticking the appropriate box shown on your screen.

Step 2

Read carefully.

- If you HAVE indicated in step 1 that you do not want to contribute to the GROUP ACCOUNT, this is the end of the round for you. You will receive as your earnings from this round all the points you have placed in your PRIVATE ACCOUNT.

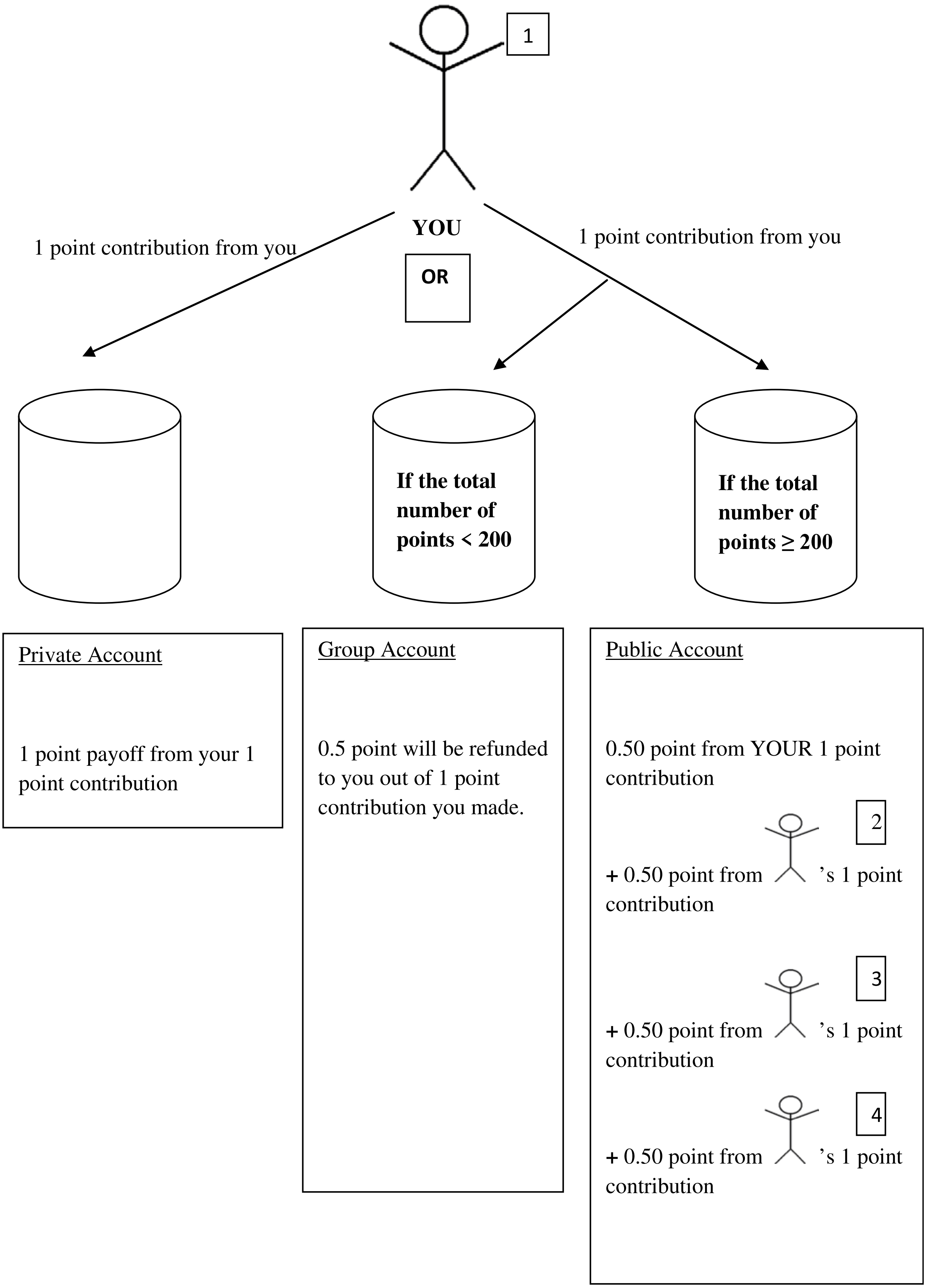
- If you HAVE NOT ticked the box indicating that you do not want to contribute to the GROUP ACCOUNT, you must decide how many points you would like to place in the GROUP ACCOUNT. **Note that you will not be allowed to place 0 points in the GROUP ACCOUNT.**

After all participants have made their decisions for a round, the computer will tabulate the results. You will be informed of the total contribution to the group account and your total earnings for the round. This will be shown on your computer screen.

This is the end of the decision making process for the round. After this, a new round begins.

Do refer to the above steps as a guide. You will have to do the same exercise for several rounds. **At the end of the experiment, one of the rounds will be randomly selected as a binding round. Your actual payment will be based on the decision you made in this binding round.** In principle, every round has the chance to be picked as the binding round, so you should treat every round seriously as if it is going to be picked as the binding round.

Illustration



Chapter 2 Liquidation Policy and Disclosure of Credit History in Financial Contracting: An Experiment

2.1 Introduction

Financial contracts govern the relationship between lenders and borrowers. Lenders provide loans to borrowers, and in return borrowers must repay the loans with interest payments. These financial contracts are accompanied by loan covenants stipulating the rights and obligations of both parties and actions to take under various contingencies designed to provide sufficient incentive for borrowers to repay the loans and to provide protection for financiers in the event of default. As complete as these contracts can be, there will always be unforeseen contingencies that are not covered by the contracts.

When these contingencies arise, both parties could potentially end up in a protracted dispute requiring a third party to adjudicate. For example, a borrower may opportunistically claim that she is facing financial distress and unable to repay the loan, then ask for financial leniency from the lender despite being actually financially sound. The incentive to commit to such a strategic act is exacerbated in the presence of asymmetric information between the borrower and the lender. The borrower usually knows whether he is facing financial distress, but the lender does not. Even when the lender is able to partially verify the borrower's information at some costs, there is no assurance that the lender can force the borrower to repay.

This reason demonstrates the importance of the issue of how to design optimal financial contracts. It has also been the subject of much attention in the economics and finance literature (see Hart, 2001 for an excellent review). The literature generally considers two types of asymmetric information problems that lenders face. They are (1) the adverse selection problem, which is the problem of not knowing borrowers' credit history prior to signing a loan agreement and (2) the moral hazard problem, which arises due to the inability of lenders to observe borrowers' actions once the loans have been disbursed. An optimal financial contract should therefore be designed to optimally screen borrowers when the adverse selection problem is present and provide enough incentive for borrowers to repay the loan when the moral hazard problem is present.

In terms of the modeling strategy, the literature adopts two main approaches. The first is the complete contract approach, wherein it is assumed that a financial contract can be based on all possible states of eventualities. It is comprehensive and leaves no room for re-negotiation. The second is the incomplete contract approach, which argues that most contracts in reality are 'incomplete' because it is impossible or even if it is possible, it would be prohibitively costly to write a complete contract (see, for example, Townsend, 1979; Spier, 1992).

Under the incomplete contract approach, the role of control rights over borrowers' assets when default occurs becomes particularly important. These rights would give power to the holder to decide the course of action to take when default occurs and thus would provide enough incentive for the holder to lend money. In the absence of the assignment of control rights, the asymmetric information problem and contract incompleteness (which prevents lenders from basing loan contracts on cash-flow realization) would place lenders in a vulnerable position. Borrowers can strategically divert cash flow to themselves instead of repaying lenders. In anticipation of this possibility, lenders might be unwilling to fund projects even if these projects are profitable. This insight was first developed by Grossman and Hart (1986) and subsequently by Hart and Moore (1990 and 1998) and Aghion and Bolton (1992).

In particular, Aghion and Bolton (1992) applied the incomplete contract approach to analyze the optimal loan contract between a lender and a borrower. In their framework, the optimal loan contract calls for a state-contingent loan contract that allows lenders to terminate funding and take control over borrowers should borrowers fail to repay (see also Dewatripont and Tirole, 1994; Gromb, 1994).

Bolton and Scharfstein (1996) extended the Aghion and Bolton state-contingent loan contract to examine the role of probabilistic liquidation policy upon transfer of control rights from borrowers to lenders when non-repayment occurs. Liquidation policy can be designed optimally to exert disciplinary muscle over errant borrowers with high propensity to commit strategic default. The termination of funding and repossession of assets that might prevail in a strategic default would keep borrowers from enjoying the continuation payoffs in the event that their companies are liquidated. This may reduce borrowers' incentive to misbehave. Unfortunately, this deterrence benefit also comes with costs in the form of inefficiency arising from a loss of value because assets are often sold for less than their actual value when they are liquidated.[6] Bolton and Scharfstein (1996) hence argue that an optimal liquidation policy must be designed to balance the trade-off between the deterrence of strategic default and the loss in economic value arising from liquidation. The optimal liquidation policy calls for probabilistic liquidation; borrowers will be liquidated with positive probability if they fail to repay the loan regardless of the underlying reason for default. This probabilistic liquidation policy could be too harsh on borrowers when failure to repay occurs because of a genuine (temporary) liquidity shock. In such a situation, lenders also have to forego the opportunity to obtain future gains from continuing the lending relationship. The key assumption here is that the borrowers can recover from the temporary liquidity shock.

This chapter delves into the role of liquidation policy in debt contracting. In particular, it adopts Bolton and Scharfstein's framework of probabilistic liquidation policy as the centerpiece and presents an experimental analysis of the impact of liquidation policy on borrowers' incentive to engage in strategic default. One potential drawback of the existing empirical studies is confounding factors that complicate statistical inference. An experimental economics methodology can deal with these confounding factors, presenting a viable

[6] See also Hotchkiss et al. (2007) for a thorough survey of literature on corporate bankruptcy and liquidation.

alternative methodology to the conventional empirical methodology. Specifically, in this study, experimental methodology allows us to clearly distinguish between strategic default and liquidity default, something that would be difficult if not impossible to do using the existing empirical data. This chapter also serves to test Bolton and Scharfstein's (1996) model.

In addition, this chapter evaluates the role of credit history in lending relationships. Specifically, we compare the incidence of strategic default and lenders' choice of liquidation policy when information on borrowers' credit history is provided to- and shared by all lenders. Similar to liquidation policy, information on credit history could play an important disciplinary role in deterring borrowers from defaulting strategically. Information on credit history gives lenders better knowledge about borrowers' credit history, thereby softening the adverse selection problem. Lenders can also completely shun bad borrowers from access to credit. This also gives lenders stronger bargaining power when negotiating the lending terms with borrowers.[7]

Pagano and Jappelli (1993) developed a theoretical model of credit information sharing among lenders in the presence of adverse selection. Their model showed that information sharing among lenders attenuates adverse selection and reduces the strategic default rate. In related papers, Padilla and Pagano (1997 and 2000) also showed that, in the presence of moral hazard, credit information sharing disciplines borrowers and reduces the strategic default rate. The sharing of credit information among lenders would open up an opportunity for borrowers to develop a good reputation and increase their desire to maintain a good reputation would lessen the extent of conflict of interest between lenders and borrowers (Diamond, 1989). In the context of the model of multiple-bank lending, Bennardo et al. (2009) pointed out that credit information sharing mitigates the over-borrowing problem, and this has the benefit of reducing the rate of defaults. In sum, theoretical models on credit

[7] See Jappelli and Pagano (2006) for an excellent survey on the role and effects of sharing credit information among lenders on lending relationship.

information sharing, regardless of their modeling framework, point to the benefit of sharing credit information in reducing the strategic default rate.

On the empirical front, Jappelli and Pagano (2002) found evidence that credit information sharing is associated with lending expansion and reduced credit risk. Djankov et al. (2007) used aggregate cross-country data and showed that credit information sharing improves borrowers' access to credit. A host of other papers have shown that credit information sharing improves credit market performance by fostering lending and reducing default rates (e.g., Brown et al., 2009; De Janvry et al., 2010; Cheng and Degryse, 2010; Degryse et al., 2012; Doblas-Madrid and Minetti, 2013.

Our study contributes to the literature on financial contracting and credit information sharing in two ways. First, to the best of our knowledge, this is the first experimental study focusing on the interplay between the liquidation policy and borrowers' strategic behavior. In particular, through a series of laboratory experiments, we evaluate how lenders set their liquidation policy, what factors influence the harshness of the liquidation policy, how this liquidation policy affects borrowers' incentive to engage in opportunistic behavior, and how the liquidation policy influences credit market performance.

Second, to the best of our knowledge, our study is one of only two experimental studies exploring the effects of the disclosure of credit information on lending relationships and credit market performance. Brown and Zehnder (2007) analyzed the effects of credit reporting on borrowers' repayment decisions and credit market performance in a lending game framework similar to an experimental trust game. However, they only focused on the impact of the induced treatment effects arising from an environment with credit information sharing relative to a baseline environment without it. In their experimental framework, lenders did not impose any liquidation penalties on defaulters, and the lending relationship was exogenously determined. That is, participants who were assigned the role of lender were randomly matched with participants who were assigned the role of borrower. In contrast, our experimental framework has an endogenous matching process between lenders and borrowers. Borrowers offer a financial contract to lenders stipulating how frequently the borrowers will be liquidated if they fail to repay the loans.

By having both liquidation policy and credit information sharing within the same experimental framework, we are able to evaluate whether or not credit information sharing is a substitute for the liquidation policy. That is, credit information sharing significantly reduces the problem of information asymmetry between lenders and borrowers. Lenders are better informed about the type of borrowers they are facing. As such, there is no need for lenders to safeguard their interests by imposing a harsh liquidation policy if default occurs. Thus, credit information sharing and liquidation policy could be substitutes.

The rest of the chapter is organized as follows. Section 2.2 introduces Bolton and Scharfstein's model, and Section 2.3 presents our experimental design and procedure. Section 2.4 presents our experimental predictions, followed by the results in Section 2.5. Section 2.6 concludes the chapter.

2.2 The Bolton and Scharfstein Model

Bolton and Scharfstein (1996) developed a two-period model of a lending relationship between a bank and a firm that has no initial wealth. At time $t = 0$, the firm needs an initial investment K to implement a project with uncertain payoff (e.g., cash flow). The project is either successful with probability θ or failed with probability $(1 - \theta)$. At time $t = 1$, if the project is successful, the cash flow is equal to x. But if it fails, the cash flow will be equal to 0. The firm borrows K from the bank. Both the bank and the firm are assumed to be risk neutral.

Similar to other incomplete contract models (for example, Grossman and Hart, 1986), cash flow is assumed to be *observable* by both parties, but *not verifiable* by a third party (e.g., a court). Consequently, the loan contract cannot be made contingent on the realization of cash flow, and instead it should specify the allocation of control rights over assets in the presence of default. More specifically, the loan contract specifies that if the firm repays an amount of R_x (i.e., the repayment when the cash flow is x), the bank has the right to liquidate the firm's assets with probability β_x. If the firm repays an amount of R_0 (i.e., the repayment when the cash flow is 0), the bank has the right to liquidate the firm's assets with probability β_0. The repayment and the liquidation decisions take place at the end of period $t = 1$. Essentially, if the liquidation takes place, the control rights over the firm's assets are transferred from the firm to the bank. If the firm survives liquidation at period $t = 1$, the firm proceeds to period $t = 2$ and receives the continuation cash flow of y with certainty. The firm's expected payoffs (given the firm is truth telling) can then be expressed as,

$$\theta[x - R_x + (1 - \beta_x)y] + (1 - \theta)[0 - R_0 + (1 - \beta_0)y]. \qquad (2.1)$$

The first and the second part represent the accrued payoffs when the project is successful and fails, respectively.

The bank's expected payoffs can be expressed as

$$\theta(R_x + \beta_x L_x) + (1 - \theta)(R_0 + \beta_0 L_0) - K, \qquad (2.2)$$

where L_x and L_0 represent the liquidation value of assets when cash flow is x and 0, respectively.

Given that the firm has no initial wealth, the repayment at period $t = 1$ cannot exceed the amount of funds available, which implies that $R_0 \leq 0$ and $R_x \leq x$. Under the risk neutrality assumption, the loan contract is designed to be incentive compatible to ensure that the manager has an incentive to repay R_x rather than R_0 when cash flow is x. In other words, the firm has an incentive to repay instead of defaulting on the loan when the project is successful. We define the default decision when the project is successful as *strategic default*. The incentive constraint can then be expressed as,

$$x - R_x + (1 - \beta_x)y \geq x - R_0 + \beta_0 S + (1 - \beta_0)y, \qquad (2.3)$$

where S refers to the utility that the firm's manager receives by paying R_0 when the actual cash flow is x and the assets are subject to liquidation. The left side of the inequality represents the manager's payoff when he behaves, and the right side represents the payoffs when he misbehaves.

In addition to satisfying the above incentive constraint, the optimal contract must also give an incentive for the firm to repay R_0 rather than R_x when cash flow is 0. However, it is easy to see that this constraint is not binding because the firm cannot repay a positive amount of R_x when cash flow is 0. Finally, at the optimum, the bank's payoffs from lending must be non-negative.

$$\theta(R_x + \beta_x L_x) + (1-\theta)(R_0 + \beta_0 L_0) - K \geq 0. \tag{2.4}$$

It is optimal to set $R_0 = 0$ and $\beta_x = 0$. That is, when cash flow is 0, the repayment amount must also be 0. When the firm repays R_x given that cash flow is x, the firm should not be liquidated. It is straightforward to establish that the incentive constraint and the nonnegative profit constraint must be binding. Substituting $R_0 = 0$ and $\beta_x = 0$ into (2.3) yields,

$$R_x = \beta_0(y - S). \tag{2.5}$$

Substituting (2.5) to (2.4) and assuming that (2.4) is binding yields,

$$\beta_0[\theta(y-S) + (1-\theta)L_0] - K = 0. \tag{2.6}$$

The optimal β_0 can then be derived as;

$$\beta_0 = \frac{K}{\theta(y-S) + (1-\theta)L_0}. \tag{2.7}$$

β_0 is increasing in the amount of investment outlay K and decreasing in the continuation cash flow y at period $t = 2$, and the liquidation value of assets L_0 when cash flow is 0. Since probability has to be within $[0,1]$, we require that $K \leq \theta(y-S) + (1-\theta)L_0$ and $\theta(y-S) + (1-\theta)L_0 > 0$. Under some parameter values, there will be a strictly positive probability of liquidation when the repayment is R_0. This implies that regardless of the reason for the lack of repayment, the bank may liquidate the firm. This may be too harsh when the lack of repayment is caused by a liquidity default rather than by a strategic default, but it also has the benefit of deterring the firm from making no repayment when cash flow is x.

2.3 Experimental Design and Procedures

Our experiment has three aims. First, we investigate factors influencing the determination of the optimal liquidation policy. Second, we evaluate the role of liquidation policy in deterring the firm from engaging in strategic default. Third, we experimentally analyze the effects of the disclosure of borrowers' credit history information on liquidation policy and borrowers' and lenders' behaviors.

The Benchmark Optimal Liquidation Probability

For the purposes of our experiment, we assigned the following parameter values to the above model: $K = 14$, $x = 30$, $(y - S) = 36$, $R_x = 18$, $L_0 = 12$ and $\theta = 2/3$. It is straightforward to verify that these values satisfy (2.5) and (2.6), and thus there exists a feasible solution for β_0. Substituting these parameter values to (2.7) yields

$$\beta_0 = 0.5. \tag{2.8}$$

Thus, whenever firms do not make any of their repayments, regardless of whether this is due to a liquidity problem or the moral hazard problem, banks would liquidate them half of the time.

The Experimental Treatments

We conducted three experimental treatments. The first treatment is the *baseline* treatment, whereby the liquidation policy is absent. In this treatment, banks and firms are matched randomly. Banks extend loan financing to firms and specify the loan repayment amount. Firms invest using the funds received from the bank, and the investment may or may not be successful. When it is successful, the cash flow generated is 30; otherwise the cash flow generated is 0. Firms must then make a binary decision of whether or not to repay the loan. Since there is no punishment for not repaying, the setting is identical to that of the standard binary trust game.

In the second and third treatments, there is a liquidation policy. Banks and firms have to go through a matching process. In particular, firms make an offer of liquidation probability to banks, that is, how severe they would want banks to be when they implement the liquidation policy. We introduce competition to attract borrowers among banks by having more banks than firms. This competition would drive the non-negative constraint closer to 0 and reduce the bargaining power of banks.[8] In contrast to the second treatment, in the third treatment we provide banks with past information on firms' credit history. That is, we give them information on how often firms have defaulted in previous

[8]In order to provide enough incentive to participants playing the role of the bank in our experiment, we need to ensure that they would on average obtain positive earnings from the experiment. We do this using a combination of the show-up fee and the earnings from the experiment.

periods. We coin this treatment as the *with credit history* (*CH*) treatment and the second treatment as the *without credit history (NCH)* treatment.

The Matching Process

The matching process between borrowers and banks in the CH and NCH treatments, which takes place at the beginning of each period, can be described as follows. Firms require funds K to implement a project and approach banks for a loan. Firms post their desired liquidation probability β_i for all banks to observe. Banks then decide which firm to select.[9] If a firm is selected by only one bank, the pair is matched successfully and the agreed β_i will be incorporated in the loan contract. If a firm is selected by multiple banks, a random draw decides which bank is going to be matched with the firm. This completes one matching cycle. Unmatched firms and banks after the first matching cycle enter the next cycle and repeat the matching process all over again. We allow for up to 5 matching cycles in the matching process. If they remain unmatched after 5 matching cycles, they will have to be inactive until a new round begins. An unmatched bank would obtain 6 points and an unmatched firm would obtain 0 points.[10]

The Game Structure and the Type of Defaults

At the beginning of each round, each bank is endowed with 14 points to lend to a firm. The firm receiving the loan will use the funds to cover its capital outlay K. The project may succeed with probability $2/3$ yielding 30 points, or fail with probability $1/3$ yielding 0 points. The probability of project success is common knowledge. If the project succeeds, the firm can either repay the bank with 18 points covering the amount loaned to the firm and the interest payment

---

[9]We opt for a setup whereby firms make an offer of β_i to banks to be in line with the assumption that firms have stronger bargaining power than banks. There are more banks than firms, and we do not allow multiple banking relationships, so there will be some banks that will not get matched and miss out on the opportunity to earn higher payoffs. Banks compete to attract firms and since firms know this, they would in theory offer β_i, which will just make the non-negative profit constraint of banks binding. Letting banks make an offer of β_i to firms should in theory not alter the results.

[10]Alternatively, we could choose different payoff magnitude for the unmatched banks. The choice of 6 points is to provide banks with a bit of bargaining power albeit much lower than that of firms, and it allows them to have on average some reasonable earnings from the experiment. We believe that changing this payoff magnitude would not qualitatively alter the results of our experiment.

or default on the loan. When the firm defaults even though the project is successful, the firm is said to be engaging in *strategic default*. When the project fails, the firm has no choice but to default on the loan. We define this as *liquidity default*.

If default occurs regardless of the reason, the bank is entitled to liquidate the firm and sell its assets and may do so with some probability agreed upon earlier between both parties and stated in the loan contract. If the firm repays the loan, the bank cannot liquidate the firm. The bank obtains 12 points if the matched firm is liquidated for defaulting. The game ends when liquidation occurs. Otherwise, the game proceeds to $t = 2$. If the firm survives liquidation and proceeds to $t = 2$, it obtains a sure second period payoff of 36 points. Both banks and firms have complete information on the project status in period 1 (i.e., whether or not the project is successful). Figure 2.1 illustrates the game tree.

Banks must use the available 14 points to loan to firms in each round. If they decide not to get involved in a lending relationship, they will have to sit out of that particular round and will only receive 6 points.[11] When they are successful in lending 14 points to a borrowing firm and the firm repays the loan, they receive 18 points.[12]

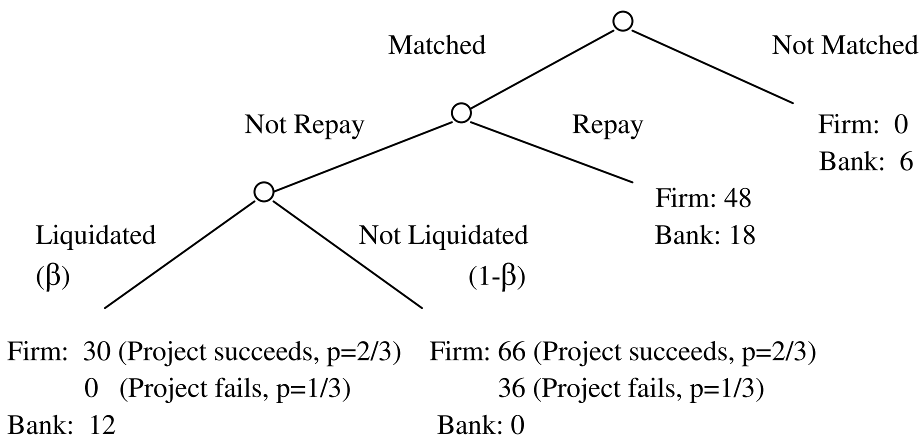


Figure 2.1. The game tree

[11]We can consider these 6 points as the amount of money available to banks after covering their cost of funds (e.g., the principal and the interest payments to depositors).

[12]These 18 points are the amount of money available to banks from their interest earned minus their cost of funds (e.g. the principal and the interest payments to depositors).

This experimental design is implemented in both the *CH* and *NCH* treatments. The only difference between the two treatments is the presence of credit history in the *CH* treatment. Banks in the *CH* treatment are given historical information at the beginning of each round containing the number of successful projects out of the total number of projects implemented up to the proceeding round by all firms, the percentage of defaults committed by all firms, and the frequency of strategic default committed by all firms. The latter is an indicator of the tendency for firms to engage in moral hazard.

The baseline treatment, wherein there is no liquidation, is implemented to capture firms' *intrinsic* trustworthiness. Since there is never any liquidation in period 1, and there is always a sure payoff in period 2 that is independent of firms' behavior in period 1, it suffices to focus on firms' behavior in period 1. For this reason, we omitted period 2 and only implemented a one-period investment game in this baseline treatment. However, the initial endowment given to banks, the probability of success of the project, the corresponding expected payoff from the project, and the choice of whether to default on the loan or to repay the loan remains the same as specified in the other two treatments. It should be noted that banks, who were randomly paired with firms rather than being asked to go into the matching process, have no decisions to make in the baseline treatment. The bank side is totally passive and has no interaction with firms in the baseline treatment. The bank side in the baseline treatment was played by experimenters. Participants were not told about this to ensure that their behaviors were not affected.[13] The amount of 14 points was automatically disbursed to firms in every round.

[13] We would like to take a detour here to clarify that we did not use deception in this experiment, though deception was used in some economics experiments (e.g., Weimann, 1994; Blount, 1995; Scharlemann et al., 2001; Sanfey et al., 2003). Deception is defined as explicit and intentional misstatement of information (see Ortmann and Hertwig, 2002, for a detailed review). In this treatment, we did not give any information on bank players' identity and no single subject ever asked any question regarding banks' identity. As pointed out by Hey (1998), "there is a world of difference between not telling subjects things and telling them the wrong things. *The latter is deception, the former is not."* (see the similar definition of deception in Baumrind, 1985; Adair et al., 1985; Nicks et al., 1997). In addition, banks in this treatment are completely passive. They have no interaction with firms and have no impact on firms' behavior. Thus, we argue our setting has mininal effects on firms' behavior. There are also experiments using confederates in the literature (see, for instance, Smith, 1991; Hoppe and Sadrieh, 2007).

The experiment was conducted via the *www* interface at the computer lab at Nanyang Technological University (NTU). Participants were NTU undergraduate students from various academic backgrounds, such as business, science, engineering, arts, and social sciences. There were 23 firms and 25 banks in the *NCH* treatment, 22 firms and 25 banks in the *CH* treatment and 24 firms in the baseline treatment.[14] We provided paper instructions,[15] which were read aloud at the beginning of each session. Participants played the game for 10 rounds. The first two rounds were practice rounds. Participants were not informed of the total number of rounds.[16] Participants were randomly assigned to firms or banks, and they were only identified by numeric IDs. Out of all real rounds, one was randomly selected to determine participants' payment in addition to a 5-dollar show-up fee. All points earned during the experiment were converted to Singapore dollars at the exchange rate 5 points = 1 SGD. Participants were paid in cash privately at the end of the experiment. The average duration of one session was 2 hours.

2.4 Experimental Predictions

In this section, we formulate our hypotheses.

Prediction 1. *Borrowers in the baseline treatment where the liquidation threat is absent are more likely to strategically default on a loan than borrowers in the NCH treatment. Nevertheless, even in the absence of liquidation threat, there should still be a significant proportion of borrowers who choose not to default on a loan when the project is successful.*

A strategic default occurs when a project is successful, but the borrower fails to repay the loan. In such a situation, liquidation serves as a penalty on borrowers who commit such an act. The presence of liquidation should therefore increase borrowers' incentive to repay the loan when the project is

---

[14]The difference in the number of firms is due to recruitment conditions. We believe this small difference does not affect the experiment process as well as the experimental results.

[15]A sample instruction for banks and firms in the CH treatment is provided in appendix B. Other instructions are available upon request.

[16]In the baseline treatment, participants actually played 15 rounds in total. However, the trustworthiness revealed in terms of the strategic default rate remains almost the same when either 15 rounds or the first 10 rounds were used. For the sake of comparison, we report the first 10 rounds. The complete data is available upon request.

successful. When the liquidation threat is completely absent, it is straightforward to see that the incentive compatibility constraint (2.3) would be violated. This implies that, in theory, the borrowers would never repay the loan when the project is successful. Notice that, in the absence of liquidation, borrowers would essentially face the standard trust game of Berg et al. (1995). Evidence from trust game experiments consistently shows that trustees are still willing to reciprocate with the trustors by sending back some positive amount to trustors. We should therefore expect that under our baseline treatment whereby the liquidation policy is absent, a significant proportion of borrowers will still repay the loan.

The next prediction concerns the optimal choice of the liquidation probability in both the CH and NCH treatments.

Prediction 2. *In equilibrium, banks would liquidate borrowers that are unable to repay the loan 50% of the time in both the CH and NCH treatments.*

The optimal liquidation probability is depicted by expression (2.7) shown earlier. Substituting the relevant parameter values used in our experiment into (2.7) yields $\beta = 0.5$ as shown in (2.8). This value maximizes the borrower's expected payoff and satisfies the borrower's incentive constraint and the bank's nonnegative profit constraint regardless of whether or not the credit history is disclosed.

The next prediction touches on the role of the borrowers' credit history information on the borrowers' incentive to default on the loan strategically when their project is successful.

Prediction 3. *In the NCH treatment, wherein the information on the borrowers' credit history is not available to banks, borrowers will repay the loan when the project is successful if and only if their chance of being liquidated (β) after default is at least 50%. However, when the borrowers' credit history information is available to banks, keeping other things constant, borrowers would be less likely to default on the loan in order to enhance their reputation and give them more bargaining power when they negotiate the liquidation policy with the banks.*

Borrowers are indistinguishable when credit history information is not available. Their incentive constraint is satisfied if $\beta \geq 0.5$, and the borrowers' best strategy would be to always repay the loan when the project is successful. However, when the credit history information is revealed to creditors, the borrowers' current behavior may have some bearing on the future lending relationship. Consequently, borrowers are less likely to default when the project is successful.

The next prediction is related to the bargaining process between banks and borrowers over the severity (probability) of the liquidation policy (β) to be implemented when default occurs. Recall that a bank and borrower pair matches and a lending relationship is formed when the bank accepts the firm's offer of β.

Prediction 4. *Banks will only accept an offer of* β *from a firm provided that* $\beta \geq 0.5$.

The above prediction is derived from the fact that the bank's expected profits (2.2) are non-negative when $\beta \geq 0.5$.

The last prediction is related to the use of the offer of β as a signaling device for a good-quality borrower to credibly signal her credit history. Banks are not informed about the borrowers' credit history. Without the availability of the credit history information, there is no way for banks to ascertain the type of a borrower seeking a loan. Knowing this, good-quality borrowers would use the offer of β as a strategic tool to separate themselves from bad-quality borrowers. In particular, good-quality borrowers can offer a sufficiently high β, which bad-quality borrowers are unwilling to offer. By examining the repayment history, we can infer borrowers' type from their propensity to repay the loan when the project is successful.

Prediction 5. *Under the NCH treatment, the good-quality borrowers will tend to offer a significantly higher* β *than bad-quality borrowers in an attempt to signal their type. Under the CH treatment, the availability of credit history information mitigates the adverse selection problem faced by banks. Consequently, the offer of* β *in this treatment tends to be smaller than that in the NCH treatment.*

2.5 Experimental Results

In this section, we present our experimental results. First, we go through the summary statistics. Then, we present an analysis of the choice of β in both the CH and NCH treatments and the effects of the credit history information on the choice of β. Subsequently, we investigate borrowers' incentive to engage in strategic default in both treatments. We then evaluate the matching process between banks and borrowers focusing on the impact of the credit history information on the efficiency of the matching process.

2.5.1. The Summary Statistics

The frequency with which the borrowers defaulted (γ) when their project is successful gives us a measure of the borrowers' propensity to engage in strategic default. This can be expressed as

$$\gamma = \frac{\sum(Default|Success)}{\sum Success}$$

Table 2.1 presents the summary statistics of all treatments. We exclude the first two trial rounds and focus only on the eight main rounds. In both the *CH* and *NCH* treatments, there is an endogenous matching process between borrowers and banks. To be consistent with the theoretical framework of Bolton and Scharfstein (1996), which assumes that the banking sector is perfectly competitive, we have more banks than the potential borrowers. Consequently, in every round, some of these banks would fail to match with a borrower and therefore have to be inactive in that particular round.

The average mutually agreed liquidation probability (β) in both the CH and NCH treatments fluctuates around 0.5, which is the optimal β as predicted by the theory. The average β is lower than 0.5 in the *CH* treatment and it is higher than 0.5 in the *NCH* treatment.

Table 2.1. The Summary Statistics by Treatment

Treatment	Number of borrowers	Number of Banks	Number of matching pairs	Average β	Number of successful project	Number of defaults	Number of strategic defaults	Propensity to default strategically (γ)
CH	22	25	170[a]	0.468	119	91	40	33.6%
NCH	23	25	177[a]	0.529	122	103	48	39.3%
Baseline	24	24[b]	192	NA	122	160	90	73.8%

Note: a. One pair with a valid agreed β is considered as 1 observation. Matching failure is excluded.

b. The bank side is played by experimenters in this treatment.

The number of matching pairs shown in column 4 of Table 2.1 also gives us the number of projects financed by banks. Recall that in our experimental design we set p = 2/3, implying that 2/3 of the projects would be successful. In the CH treatment, 119 out of 170 projects (70%) were successful and generated positive cash flow. In the NCH treatment, 122 out of 177 projects (69%) were successful. In the baseline treatment, 122 out of 192 projects (64%) were successful. Thus, indeed roughly about 2/3 of the funded projects were successful.

Column 7 gives the total number of defaults, which includes both the liquidity and strategic defaults. In the CH treatment, around 44.0% of the defaults (40 out of 91) are strategic defaults. In the NCH treatment, 46.6% (48 out of 103) of defaults are strategic compared to 56.3% (90 out of 160) in the baseline treatment. The higher proportion of strategic defaults in the baseline treatment is expected because of the absence of liquidation.

Column 9 presents the propensity to engage in strategic default. It is calculated by dividing Column 8 by Column 6. The propensity to engage in strategic default in the baseline treatment where the liquidation threat is absent (73.8%) is higher than that in both the CH (39.3%) and NCH treatments (33.6%) where the liquidation threat is present. It shows that the liquidation threat substantially deters borrowers from engaging in strategic default. Interestingly, the propensity to default strategically in the baseline treatment is substantially lower than 100%, indicating that even in the absence of liquidation threat, some borrowers still show some degree of trustworthiness. The presence of the credit history information in the CH treatment helps reduce the incidence of strategic default relative to the NCH treatment (33.6% vs. 39.3%). These results are consistent with Prediction 1.

2.5.2. The Optimal Choice of Liquidation Policy

In this section, we evaluate the mutually agreed liquidation policy β and then compare it with the theoretical optimal liquidation policy derived from the Bolton and Scharfstein model given the parameter values chosen in our experiments. We will also examine the impact of credit history information disclosure on the choice of β.

Figure 2.2 shows the evolution of the mutually agreed β over time. The average mutually agreed βs across rounds in both treatments fluctuate around the theoretical prediction of β, which is 0.5. The average of mutually agreed β in the *CH* treatment is, however, significantly lower than 0.5 ($p = 0.0107$ using a one-sample t-test). In the *NCH* treatment, it is significantly higher than 0.5 ($p = 0.0351$, one-sample t-test). Overall, the average mutually agreed β in the *CH* treatment is significantly lower than in the *NCH* treatment ($p = 0.0251$, Wilcoxon signed-rank test), illustrating that the presence of credit history information softens the liquidation policy. Thus, we find the mutually agreed β deviates slightly from the value predicted by the theory, contrary to the β in Prediction 2.

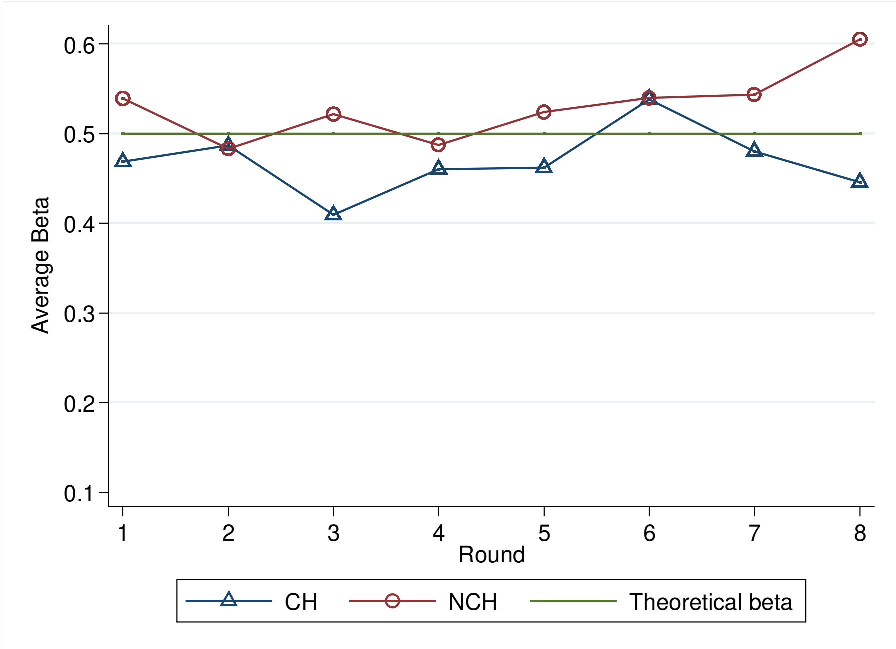


Figure 2.2. The average mutually agreed β across rounds

Figure 2.3 shows the distribution of the mutually agreed β in the CH and NCH treatments. Around 40% of the mutually agreed βs are lower than 0.5. Some of them are even less than 0.1, indicating that banks do accepts βs that violate the banks' non-negative profit constraint. This is not in line with the theoretical prediction 4 presented earlier. The mode of β in both treatments, however, is 0.5, which accounts for, respectively, 31% of the lending relationships in the *CH* treatment and 27% in the *NCH* treatment. The distribution of β is also highly skewed to the right in the *NCH* treatment, suggesting that there is a relatively large proportion of high β values. However, the distribution is more balanced in the *CH* treatment, concentrating on the range of 0.4 to 0.6. The Kolmogorov-Smirnov test of the equality of the distributions also shows that the two distributions are significantly different from each other ($p = 0.011$). In appendix A, we also present the mutually agreed β over all rounds for each individual borrower in both the *CH* and *NCH* treatments. It can be clearly seen that there are more borrowers whose mutually agreed β is above 0.5 in the *NCH* treatment than in the *CH* treatment.

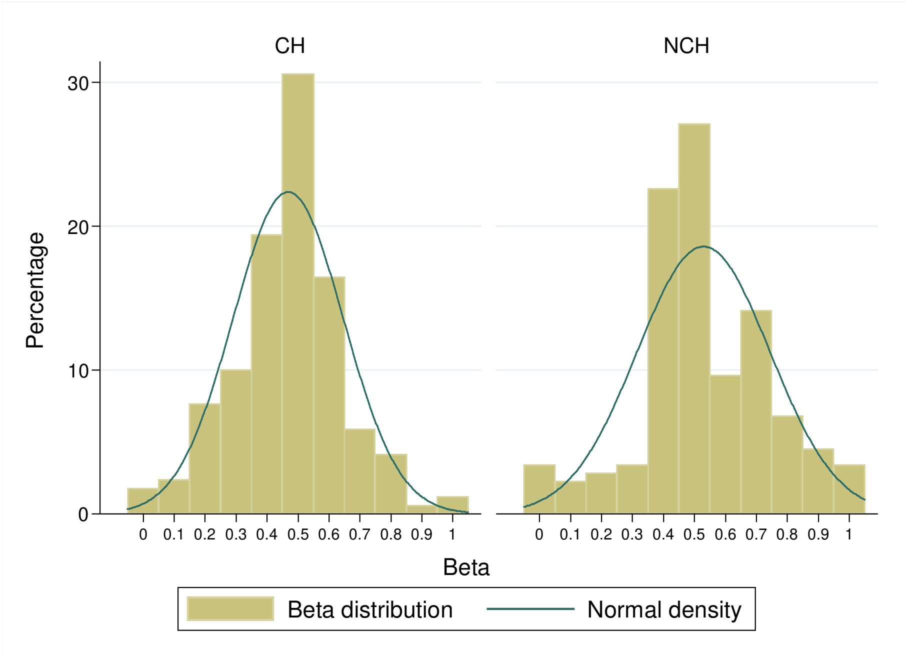


Figure 2.3. The distribution of beta and normal density

Table 2.2 reports the results of OLS regressions. The dependent variable is the mutually agreed liquidation probability (β), which also captures the severity of the liquidation policy that both parties agreed on. The independent variables include, among others, a treatment dummy (*CH*) that takes the value of 1 when credit history information is available and 0 otherwise; a variable indicating the round (*Round*); the lagged propensity to engage in strategic default expressed in percentage form (*Firm PreSdefault Percent*) and its interaction with the treatment dummy (*CH*×*Firm PreSdefault Percent*); the number of matching cycles needed by the respective borrower to match with a lender in the previous round (*Firm Cycles (t-1)*) and its interaction with the treatment dummy (*CH*×*Firm Cycles (t-1)*); the number of cycles needed by the respective bank to match with a borrower in the previous round (*Bank Cycles (t-1)*) and its interaction with the treatment dummy (C*H*×*Bank Cycles (t-1)*); the average strategic default the respective banks experience over all of the previous rounds (*Bank PreSdefault Percent)* and its interaction with the treatment dummy (C*H*×*Bank PreSdefault Percent*); a set of indicator variables representing the gender pairing composition of the subjects playing the role of bank and borrower, with mixed gender pairings (*Gender Pair (D)*) and two-female pairings (*Gender Pair (F)*); an indicator variable (*Course Pair*) taking the value of 1 if the pair is from the same course major and 0 otherwise; and an indicator variable taking the value of 1 if the pair has the same nationality (*Nationality Pair*). Note that, *Firm Cycles (t-1)* and *Bank Cycles (t-1)* can be interpreted as a measure of toughness of either the firm or the bank in the negotiation process. Longer *Firm Cycles (t-1)* imply that the firm initially offers a very low β, and no bank is willing to accept the offer, and the firm is reluctant to revise its offer upward in the early matching cycles. Likewise, longer *Bank Cycles (t-1)* imply that the bank initially is only willing to accept a sufficiently high offer of β and is reluctant to give in in the early matching cycles.

In all variations of the regression models presented in Table 2.2, the treatment dummy *CH* is significant, as are its interactions with the lagged propensity to engage in strategic default (*CH×Firm PreSdefault Percent*) and with the number of cycles needed by the borrower to match with a bank in the previous round (*CH×Firm Cycles (t-1)*). The treatment dummy also has a negative sign suggesting that the presence of the credit history information reduces the severity of the liquidation policy (i.e., the level of the mutually agreed β) given low values of variables involved in the interaction terms (*Firm PreSdefault Percent and Firm Cycles (t-1)*). This is intuitive because the presence of credit history would allow the bank to infer the borrower's type. A borrower with a good reputation would have strong bargaining power vis-à-vis the bank, and would be likely to end up with a less severe liquidation policy.

When we evaluate the impact of the borrower's type on the resulting liquidation policy (see the independent variable *Firm PreSdefault Percent* in Table 2.2) in the absence of credit history, we can infer that the better the borrower's type is, the more severe the liquidation policy will be. One can interpret this as evidence of a signaling attempt by a good borrower to the bank that she is willing to be punished harshly if she does not repay the loan. Bad borrowers would not be able to afford such a severe liquidation policy, and thus by

Table 2.2. OLS Regressions of the Determinants of Equilibrium β

	Dependent variable: Equilibrium β		
	(1)	(2)	(3)
CH	-0.286***	-0.321***	-0.318***
	(0.055)	(0.061)	(0.070)
Round	0.012**	0.013**	0.013*
	(0.006)	(0.007)	(0.007)
Firm PreSdefault Percent	-0.169***	-0.173***	-0.169***
	(0.040)	(0.042)	(0.046)
CH × Firm PreSdefault Percent	0.297***	0.322***	0.310***
	(0.067)	(0.071)	(0.079)
Firm Cycles (t-1)	-0.065***	-0.059***	-0.067***
	(0.015)	(0.015)	(0.017)
CH × Firm Cycles (t-1)	0.064***	0.054**	0.066***
	(0.021)	(0.021)	(0.024)
Bank Cycles (t-1)	0.027**	0.024*	0.026*
	(0.012)	(0.013)	(0.013)
CH × Bank Cycles (t-1)	-0.023	-0.018	-0.014
	(0.018)	(0.019)	(0.021)
Bank PreSdefault Percent		-0.044	-0.057
		(0.049)	(0.051)
CH × Bank PreSdefault Percent		0.094	0.072
		(0.071)	(0.080)
Gender Pair (D)			0.002
			(0.030)
Gender Pair (F)			-0.044
			(0.037)
Course Pair			0.022
			(0.028)
Nationality Pair			-0.034
			(0.027)
Constant	0.624***	0.633***	0.652***
	(0.049)	(0.057)	(0.066)
Observations	238	222	194
R^2	0.295	0.305	0.331

Note: Standard errors in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

undertaking this costly action the good borrower can separate herself from a bad borrower. With the availability of the credit history (see the independent variable $CH \times$ *Firm PreSdefault Percent* in Table 2.2), the asymmetric information problem faced by the bank disappears, and hence the opposite is true. That is, a good borrower would be able to get a milder liquidation policy than a bad borrower.

In the absence of credit history information, a tougher firm in the bargaining process (i.e., a firm with a longer matching cycle *Firm Cycles (t-1)*) would tend to end up with a lower β, and a tougher bank in the negotiation process (i.e., a bank with a longer matching cycle *Bank Cycles (t-1)*) would tend to end up with a higher β. When the credit history information becomes available in the CH treatment, the borrower's and the bank's number of matching cycles in the previous round do not have any impact on the equilibrium β. The possible reason could be that credit history serves as the main determinant of the equilibrium β and renders other potential explanatory variables irrelevant.

In model 2 (see the third column of Table 2.2), the percentage of strategic default that a bank has experienced previously and its interaction with the treatment dummy are added into the regressions. The qualitative aspects of the results remain unchanged. In model 3, we include more demographic variables, and again the results are qualitatively unchanged.

2.5.3. The Incidence of Strategic Defaults

Figure 2.4 presents the strategic default rate conditional on the value of the mutually agreed β in all treatments. Note that, in the baseline treatment the liquidation policy is absent and so we can focus our attention on the corresponding strategic default rate conditional on $\beta = 0$. The baseline treatment is almost identical to the standard trust game; the only difference is that we frame the problem as a formal lending relationship. The solid dot shown in Figure 2.4 indicates that the strategic default rate in the baseline treatment is 73.8%. In other words, in the remaining 26.2% of cases where the funded projects were successful, borrowers repaid back the loan even in the absence of liquidation threat. This rate is comparable to the one found in the standard trust games, which is around 37% (Johnson and Mislin, 2011). All in all, this shows that borrowers were willing to show some degree of 'intrinsic' trustworthiness even when the liquidation threat was absent.

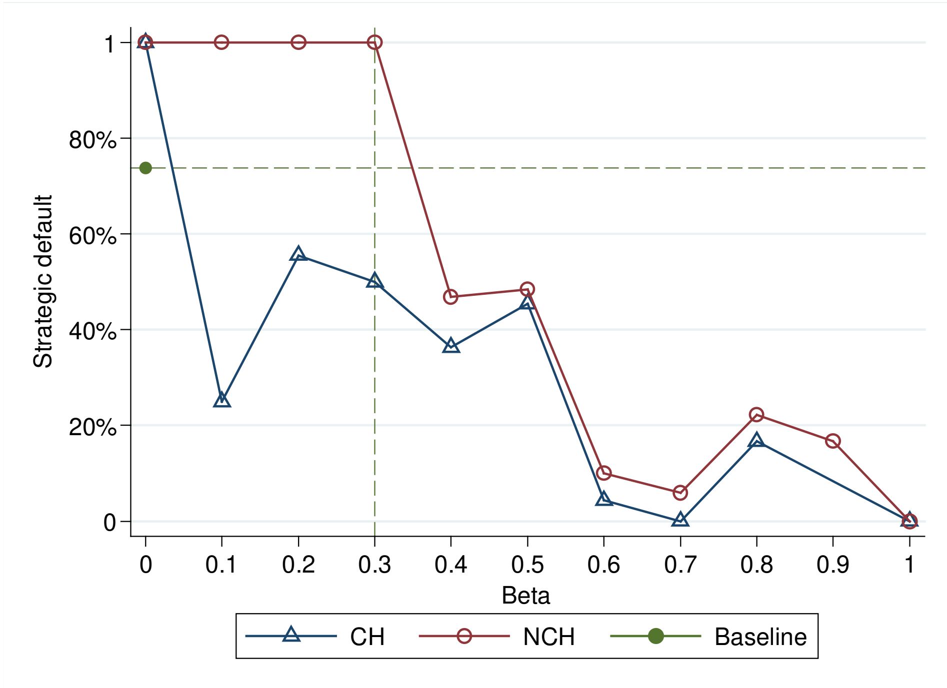


Figure 2.4. The strategic default rate conditional on the mutually agreed liquidation policy (β)

Interestingly, the intrinsic trustworthiness vanishes completely in the *NCH* treatment where the liquidation threat is present, but the credit history information is not available. This can be seen from Figure 2.4 above. When the prevailing mutually agreed β is sufficiently low such that $\beta \leq 0.3$, the strategic default rate in the *NCH* treatment goes up to 100%. This suggests that when the liquidation policy is formally in force, but is set too lenient, intrinsic trustworthiness, which exists in the absence of any punishment (in the baseline treatment), is completely crowded out. This result is consistent with the results of Gneezy and Rustichini (2000) and Fehr and Rockenbach (2003), who found that punishment could induce a detrimental effect on behavior. The reason might be that the presence of sanction threats leads to the perception of business settings rather than social contexts governed by social norms. As a result, people may perceive the sanction as the price of the business transaction (Houser et al., 2008; Li et al., 2009). Rather than deterring people from taking an undesirable action, it instead encourages them to take the undesirable action. As financial decisions are naturally in business settings, the detrimental effect is expected to be more pronounced. It is only when the punishment for defaulting strategically is severe enough ($\beta \geq 0.4$) that the strategic default rate in the *NCH* treatment becomes lower than that in the baseline treatment.

Relative to the baseline case, the rate of strategic default falls when the credit history information becomes available, and the fall is prominent for $0 \leq \beta \leq 0.5$. Recall that the borrower's incentive compatibility constraint is satisfied only when $0.5 \leq \beta \leq 1$. Consequently, we should expect that the rate of strategic default in the two treatments would be more or less identical at this range. Figure 2.4 shows that it is indeed the case. If instead $0 \leq \beta \leq 0.5$, and thus the incentive compatibility constraint is not satisfied, it is then optimal for borrowers to default strategically. The only thing that would dampen borrowers' incentive to default strategically is the availability of credit history information as shown in Figure 2.4. When the credit history information is available, borrowers have less incentive to default strategically even when the incentive compatibility constraint is violated. Prediction 3 is partially supported by our results.

Table 2.3 reports the results of probit regressions of the likelihood of the borrower to default strategically in both the *CH* and *NCH* treatments. The regressors include, among others, the mutually agreed β (*Beta*); a treatment dummy variable (*CH*) that is equal to 1 if the credit history information is present and 0 otherwise; the number of cycles needed for the pair of bank and borrower to get matched (*Cycles*); an interactive dummy variable between the *CH* and *Cycles* (C*H*×*Cycles*); a variable for the round (*Round*); an indicator variable (*Liquidation experience*) that is equal to 1 if the borrower has ever been liquidated in the previous rounds and 0 otherwise; an indicator variable (*Gender*) that takes value of 1 for female and 0 otherwise, and a set of dummy variables capturing the nationality of the participants with the baseline nationality being Singaporean. The table also provides the marginal effects of the regressors. As argued by Ai and Norton (2003) and Norton et al. (2004), the conventional way of calculating the marginal effects of the interaction terms might be invalid in non-linear models. Therefore, we use an alternative way of calculating the marginal effects proposed in their paper.

Note that β has a significant negative effect on the likelihood of strategic default. That is to say that when the liquidation policy becomes more severe, the likelihood of the borrower committing strategic default decreases. Further, the availability of credit history information would lower the likelihood of strategic default by around 43%.

Table 2.3. Probit regressions of the likelihood of strategic default

	Dependent variable: Strategic default	
	Probit	Marginal effects
Beta	-4.136***	-1.489***
	(0.632)	(0.226)
CH	-1.192***	-0.429***
	(0.403)	(0.141)
CH × Cycles	0.359***	0.092
	(0.135)	(0.089)
Cycles	-0.054	-0.019
	(0.108)	(0.039)
Round	-0.059	-0.021
	(0.053)	(0.019)
Liquidation experience	0.274	0.099
	(0.263)	(0.095)
Gender	-0.416	-0.150
	(0.275)	(0.098)
Chinese	0.728**	0.262**
	(0.345)	(0.124)
Malaysian	1.224***	0.441***
	(0.250)	(0.076)
Others	-0.355	-0.128
	(0.277)	(0.103)
Constant	1.928***	
	(0.455)	
Observations	241	241

Note: Standard errors reported in parentheses are clustered at subject level.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

2.5.4. The Matching Process

In what follows, we delve further into the matching process between banks and borrowers. If an offer of β made by a borrower is accepted by a bank, a lending relationship is initiated. Table 2.4 reports the random-effects probit regression estimates of the likelihood that a bank accepts an offer of β made by a borrower and their marginal effects. The independent variables include, among others, the magnitude of β offered by a borrower (*Beta offer*); the number of borrowers offering the same value of β (*No. of Beta*); the number of previous successful projects the borrower has had previously (*Previous Success No*) and its interaction with the treatment dummy (*CH× Previous Success No*); the number of strategic defaults the borrower has committed previously (*PreSdefault No*) and its interaction with the treatment dummy (*CH×PreSdefault No*); a variable indicating the round (*Round*); and a variable indicating how many cycles have passed when the offer of β is made (*Match Round*).

Table 2.4. Random-effect probit regressions of the likelihood of an offer of β being selected

	Dependent variable: Being selected	
	RE probit	Marginal effects
CH	-0.090	-0.179
	(0.236)	(0.175)
Beta offer	5.588***	5.588***
	(0.491)	(0.491)
No. of Beta	-0.220***	-0.220***
	(0.038)	(0.038)
Previous Success No.	-0.105	0.147*
	(0.100)	(0.088)
CH × Previous Success No.	0.509***	
	(0.100)	
PreSdefault No.	0.120	-0.447***
	(0.094)	(0.077)
CH × PreSdefault No.	-1.144***	
	(0.165)	
Round	-0.083	-0.083
	(0.058)	(0.058)
Match Round	0.091	0.091
	(0.061)	(0.061)
Constant	-1.151***	
	(0.297)	
Observations	772	772

Note: Standard errors in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

From a bank's point of view, the most important factor influencing the acceptance decision is the magnitude of β offered by the borrower. A higher offer of β would result in a higher likelihood of the borrower being selected simply because a higher β would give the bank greater control over the borrower in the event of a default. Table 2.4 also shows that stiffer competition from other borrowers, who offer the same value of β, would result in a lower likelihood of the borrower being selected.

Holding all else constant, being in the treatment wherein the credit history information is available does not affect the propensity of the borrower being selected by a bank. It is not the availability of credit history information per se that matters for the selection decision of a bank, but rather the borrower reputation, which is only identifiable if the credit history information is provided. This is illustrated by the statistical significance of the variable $CH \times$ *PreSdefault No*. Conditional on the availability of the credit history information, the worse the borrower's reputation is (i.e., the higher the number of strategic defaults committed previously) the lower the likelihood that the offer of β from the borrower is going to be selected by the bank. Note that, the marginal effects of the interaction term $CH \times$ *Previous Success No* and $CH \times$ *PreSdefault No* are not reported because the conventional way of computing the marginal effects for these interaction variables would yield imprecise estimates.[17]

2.5.5. Factors Influencing the Borrower's β Offer

In what follows, we investigate the determinants of the borrower's offer of β. The random-effects regression estimates are reported in Table 2.5. All variations of the regression models give qualitatively similar results. A borrower in the CH treatment, on average, tends to offer a lower β. The offer of β also tends to be stable over time as indicated by the non-significant effect of *Round*. *Match Round* is statistically significant and has a positive sign suggesting that the borrower tends to offer a higher β in the later matching cycles for fear that she would not be able to strike a lending relationship with a bank.

[17]The estimated marginal effects of *PreSdefault No*. at the sample mean of regressors are -1.023 ($p < 0.001$) in the CH treatment and 0.120 ($p = 0.198$) in the NCH treatment. The estimated marginal effects of *Previous Success No*. at the sample mean of regressors is 0.404 ($p < 0.001$) in the CH treatment and -0.105 ($p = 0.291$) in the NCH treatment.

Table 2.5. Random effects regressions of the determinants of β offers

	Dependent variable: β Offer		
	(1)	(2)	(3)
CH	-0.119**	-0.134**	-0.112**
	(0.056)	(0.059)	(0.056)
Round	0.005	0.004	0.002
	(0.004)	(0.004)	(0.004)
Match Round	0.034***	0.035***	0.035***
	(0.007)	(0.007)	(0.007)
PreSdefault Percent	-0.181***	-0.178***	-0.159***
	(0.059)	(0.059)	(0.046)
CH × PreSdefault Percent	0.152**	0.126	0.072
	(0.071)	(0.078)	(0.074)
Cycles (t-1)		-0.011*	-0.008
		(0.006)	(0.006)
CH × Cycles (t-1)		0.009	0.011
		(0.010)	(0.012)
Gender			0.001
			(0.056)
Engineering			0.126*
			(0.068)
Science			0.106*
			(0.064)
Chinese			0.034
			(0.059)
Malaysian			0.091
			(0.085)
Others			-0.036
			(0.052)
Constant	0.440***	0.468***	0.366***
	(0.051)	(0.054)	(0.071)
Observations	637	603	569

Note: Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

The percentage of strategic default committed by the borrower previously has a significant and negative impact on the offer of β in the *NCH* treatment, but it becomes non-significant in the *CH* treatment, where credit history information is provided. This evidence is consistent with the signaling hypothesis. Prediction 5 is supported. In the absence of credit history information, banks face asymmetric information about the borrower's creditworthiness. In the presence of asymmetric information, a good borrower would have an incentive to signal her trustworthiness by offering a sufficiently high β, such that a bad borrower would not be able to mimic the good borrower. Obviously, when credit history information is provided, the asymmetric information problem disappears. Consequently, there is no need for the borrower to offer such a high β. This is why β tends to be higher in the *NCH* treatment than in the *CH* treatment (see also Figure 2.2).

2.6 Concluding Remarks

In this chapter, we present an experimental analysis of the impacts of bank liquidation policy and credit history information on borrowers' incentive to commit to strategic default. Our experimental design is motivated by the Bolton and Scharfstein model of optimal liquidation policy in an incomplete contracting framework (Bolton and Scharfstein, 1996). We distinguish two types of financial default: liquidity default and strategic default. Liquidity default is when a borrower fails to repay the loan she receives from a bank because of a genuine liquidity problem, for instance when the borrower's project funded by the bank fails to generate positive cash flow. Strategic default is when the borrower fails to repay the loan when the cash flow generated from the project is positive. Bolton and Scharfstein (1996) showed that the optimal financing contract is characterized by a probabilistic liquidation policy. Whenever the borrower fails to repay the loan—regardless of the reason—the borrower would be liquidated with some positive probability.

We have designed an experimental treatment that incorporates the important elements of Bolton and Scharfstein (1996). We coin this treatment as the no-credit history (*NCH*) treatment. Following the spirit of the model, in this treatment, borrowers make an offer of liquidation probability (β) to banks.

Banks would decide whether or not to accept any particular offer from a borrower. Borrowers and banks who fail to match will have to go through the next matching cycle consisting of all unmatched banks and borrowers. We allow for up to five matching cycles, and if they still fail to match, they have to sit out that particular round.

We have also run two other treatments: in one treatment, the liquidation policy is absent. In that case, when a borrower fails to repay, she does not face liquidation. This treatment resembles a standard trust game. The other treatment, which we call the credit history (*CH*) treatment, has a similar setup as the *NCH* treatment, except that we provide the credit history information to banks. Essentially, under the *NCH* treatment, banks face an asymmetric information problem. They are not able to ascertain borrowers' creditworthiness. By comparing the *NCH* treatment and the *CH* treatment, we are able to evaluate the impact of the disclosure of credit history information on the lending relationship and borrowers' behaviors. The comparison between the *CH* treatment and the *NCH* treatment would also allow us to test the signaling hypothesis. The signaling device here would be the offer of β. Hence, good borrowers would signal their integrity by choosing a sufficiently high β, which is difficult for bad borrowers to mimic.

The results show that, in the absence of liquidation threat, borrowers do repay their debt sometimes, showing some degree of intrinsic trustworthiness consistent with the existing experimental evidence from the standard trust games. The presence of the liquidation threat discourages borrowers from defaulting strategically. Next, the average chosen liquidation probability in the absence of credit history is slightly higher than the probability predicted by the theory. In the NCH treatment, a 'good' borrower (defined as a borrower who has a lower propensity to default strategically) would tend to offer a higher β than the one offered by a 'bad' borrower. This result is consistent with the signaling hypothesis.

Interestingly, when the actual chosen liquidation probability is very low, such that its deterrence effect is very weak, the strategic default rate rises to 100%. In other words, the presence of a weak threat of being punished for defaulting crowds out the intrinsic trustworthiness relative to the case where the liquidation threat is non-existent. Finally, the presence of credit information significantly drives down the liquidation probability and the incidence of strategic default.

2.7 Appendix A

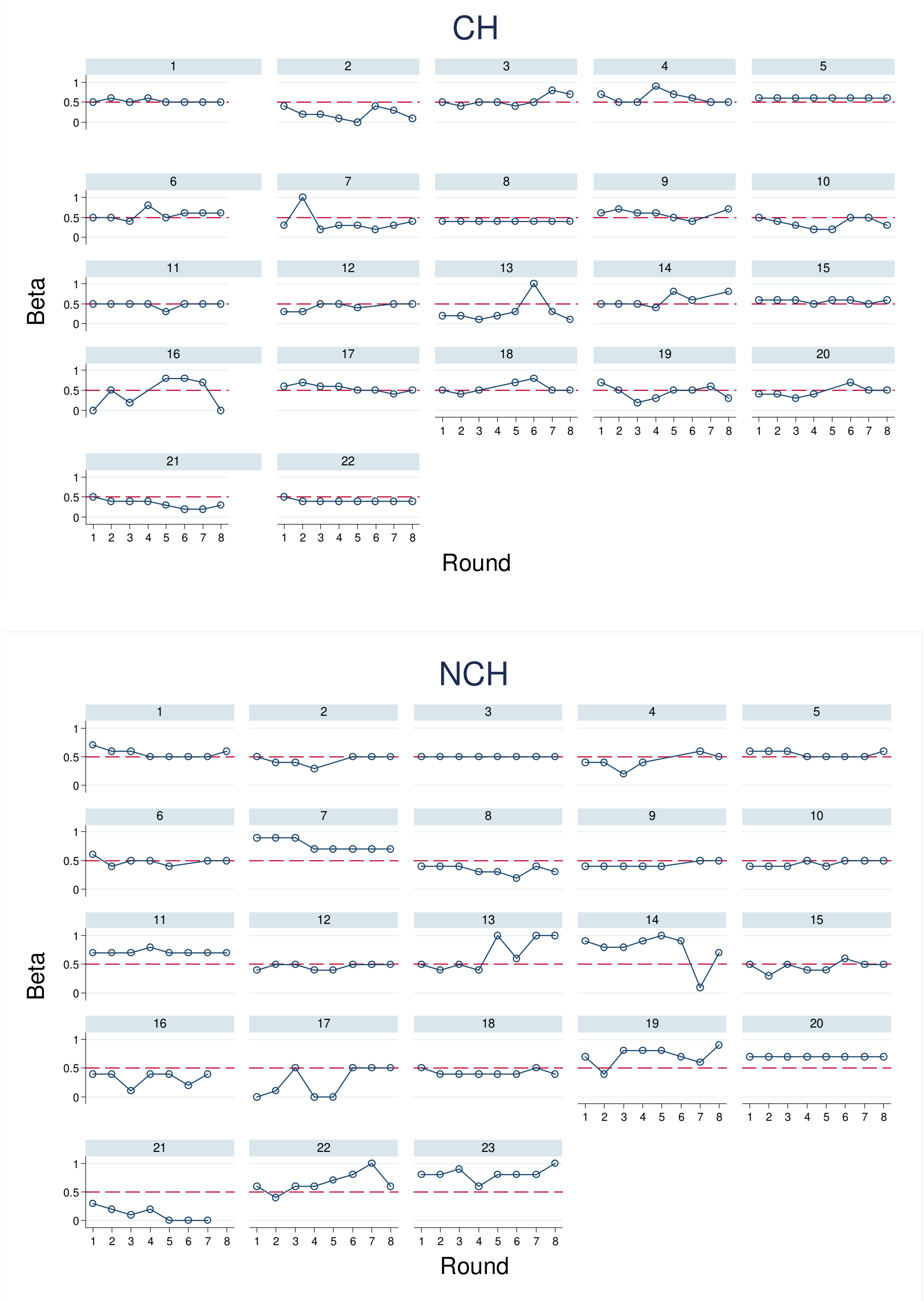


Figure 2.A1. Equilibrium beta over time by firms

2.8 Appendix B: Experimental Instructions

B.1 Firm in CH

General Information

Thank you for participating in this interactive study on decision making. **Please pay attention to the information provided here and make your decisions carefully**. If you have any questions during the session, please raise your hand and we will attend to you in private. Please **DO NOT** communicate with any other participant during the experiment. Failure to adhere to this rule would force us to stop the session and you will be asked to leave the study.

The purpose of this study is to help us better understand decision-making by observing your decisions. Information that could identify you will remain confidential. If you choose to participate, you will be asked to make a number of decisions. If you choose to withdraw after listening to the instructions, you are under no further obligation to us.

If you choose to stay, you are entitled to some amount of money during the course of the experiment. You will receive **5 SGD** as a show-up fee. In addition, you can also earn substantially more, and the amount depends on the decision you make in the experiment.

At the end of the study, your total points earned will be converted into Singapore dollar equivalent at the conversion rate of:

5 points= 1 SGD (rounded up/down to the nearest 50 cents)

Your total payment will thus be equal to your dollar equivalent of your earned points plus your 5 SGD show-up fee. Payment is made in private and you will be asked to sign a payment receipt. The receipt is for accounting purposes only and will not be linked to your responses.

To preserve anonymity, each of you will be given a unique user ID to log on to the computer terminal. Neither we nor other participants would be aware of your personal identity. We will also treat the data obtained from this study with upmost confidentiality.

Participants in this study will be assigned a role as "**Participant A**" (the firm/ entrepreneur) or "**Participant B**" (the bank creditor). Once the role has been assigned, it will be **fixed** throughout the study.

There are **23 firms** and **25 banks**. One bank will be matched with one firm through a matching process. The study will last for several rounds. In each round, a new matching process starts. At the end of the study, we will randomly draw a binding round and your payment will be based on the earnings you obtain in this binding round.

Specific Instructions for Participant A

This is a <u>two-period</u> investment game.

Your role is a **firm manager**.

You have to make an investment decision for your company. The investment will cost you a certain amount of money; however, you have no money to begin with. You can get this amount of money from B, who is a **creditor** (**bank**). If a bank has agreed to extend you a loan and you fail to repay the loan, the bank **<u>may</u>** liquidate your firm. The probability of being liquidated is denoted by β.

Example:

If you repay nothing:	
β=0	Your asset will **not be liquated for sure.**
β=0.5	Your asset will **be liquidated half of the times.**
β=1	Your asset will **be liquated for sure.**

There are **23 firms** and **25 banks**. One firm will be matched with one bank through the following process: at the beginning of each round, you **post the liquidation probability β you desire from the bank**. **Once all firms post their offers of β, all banks will be informed about all firms' offers of β. Banks will also be informed about your repayment history (i.e. how many times you defaulted when project was successful).**

Upon observing these offers, each bank can either select one and only one offer or reject all offers. **If a bank accepts** your offer, you will be matched with this bank. If there is more than one bank accepting your offer, one bank will be selected randomly as your bank partner. Thus, banks have equal chance of being selected when they choose the same firm.

If no bank accepts your offer, you need to revise your proposed β value and **enter the next cycle** of matching process. If your offer is still not accepted, the matching cycle will repeat until all firms are matched. We allow for **5 matching cycles**. If you are still not matched after 5 cycles, you will have to sit out for that particular round and your earnings from this round are **0 points**. Once the round is completed, you can again participate in the next round.

Each firm will not be informed of other firms' choices of β and their matching outcome. Thus, you will only know your own matching outcome.

With the loan of 14 points from your matched bank creditor, you will then proceed to invest in **project X** which costs exactly 14 points. This project X lasts for **2 periods**. At the end of each period you will receive a payoff for that period.

At the end of period 1, Project X **may or may not be successful**. The chance of success is **2/3** and the project yields **30 points**. The chance of failure is **1/3** and the project yields **0 points**. The following table summarizes the payoffs:

	Success	Failure
Probability	2/3	1/3
Firm Payoff	30	0

If your **project fails**, you have **no choice** but to **default on the loan**.

If your **project is successful**, you need to make your **repayment** decision. There are 2 possible repayment choices:

1. **Repaying** the bank with **18 points**;
2. Not repaying the bank (**0 points**), i.e. to **default on the loan**

If you **repay** the loan, your matched bank **cannot liquidate** you and you will **proceed to project period 2, and in period 2 the project yields a sure payoff of 36 points regardless of whether the project was successful or failed in period 1**.

If you decide **not to repay** the loan, you will **be liquidated according to probability β you and your partner selected earlier**. The computer will decide whether there is liquidation or not randomly according to this β value.

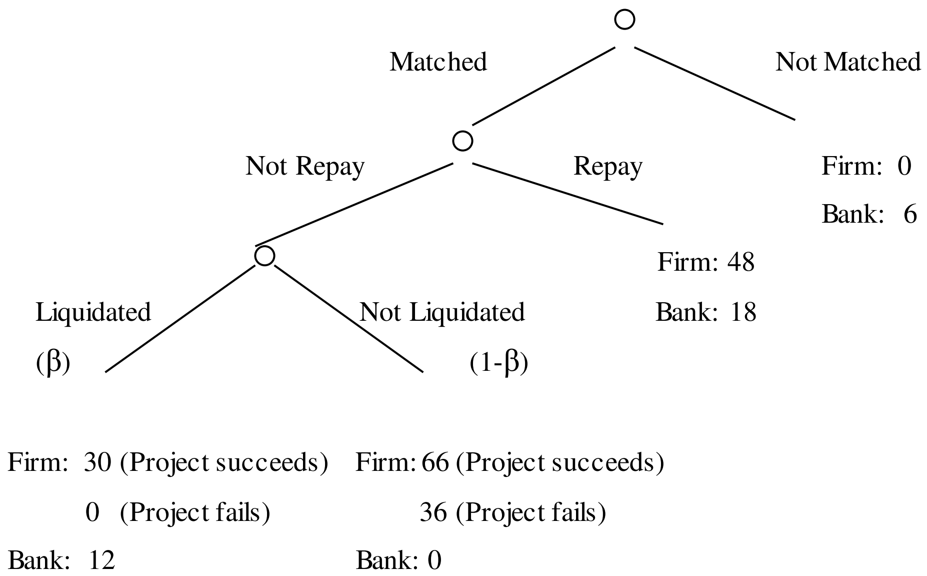
- If you are **liquidated**, the **game ends** and your total payoff is just your first period payoff.

- If you **avoid liquidation**, you will be able to continue with the project to the second period and in period 2 **the project yields a sure payoff of 36 points regardless of whether the project was successful or failed in period 1**.

At the end of the first period, your matched bank will be informed about the outcome of your project, i.e. whether it is successful or failed and the

respective payoff. Note also that banks will know your repayment history prior to engaging in the matching process.

<u>You may refer to the following Game Tree for the calculation of your earnings:</u>



<u>Note</u>: The first two rounds are trial rounds. Your results from these two rounds are not going to be taken into account in the drawing of your binding payment.

===***===

B.2 Bank in CH

General Information

Thank you for participating in this interactive study on decision making. **Please pay attention to the information provided here and make your decisions carefully**. If you have any questions during the session, please raise your hand and we will attend to you in private. Please **DO NOT** communicate with any other participant during the experiment. Failure to adhere to this rule would force us to stop the session and you will be asked to leave the study.

The purpose of this study is to help us better understand decision-making by observing your decisions. Information that could identify you will remain confidential. If you choose to participate, you will be asked to make a number of decisions. If you choose to withdraw after listening to the instructions, you are under no further obligation to us.

If you choose to stay, you are entitled to some amount of money during the course of the experiment. You will receive **5 SGD** as a show-up fee. In addition, you can also earn substantially more, and the amount depends on the decision you make in the experiment.

At the end of the study, your total points earned will be converted into Singapore dollar equivalent at the conversion rate of:

2 points= 1 SGD (rounded up/down to the nearest 50 cents)

Your total payment will thus be equal to your dollar equivalent of your earned points plus your 5 SGD show-up fee. Payment is made in private and you will be asked to sign a payment receipt. The receipt is for accounting purposes only and will not be linked to your responses.

To preserve anonymity, each of you will be given a unique user ID to log on to the computer terminal. Neither we nor other participants would be aware of your personal identity. We will also treat the data obtained from this study with upmost confidentiality.

Participants in this study will be assigned a role as "**Participant A**" (the firm/ entrepreneur) or "**Participant B**" (the bank creditor). Once the role has been assigned, it will be **fixed** throughout the study.

There are **23 firms** and **25 banks**. During the session, one bank will be matched with one firm through a matching process. In each round, a new matching process starts. At the end of the study, we will randomly draw a binding round and your payment will be based on the earnings you obtain in this binding round.

Specific Instructions for Participant B

This is a <u>two-period</u> investment game.

Your role is a **bank creditor**.

You have an endowment of **14 points**. The firm manager wants to borrow this amount for an investment project, which will be called **project X**. If you have agreed to extend participant A (the firm manager) a loan, **you have the right to liquidate the firm's investment asset with a predetermined liquidation probability β (0≤β≤1) if the firm fails to repay you the money you loaned (default on the loan).**

Example:

If the firm repays nothing:	
β=0	You will **not liquidate the firm's asset for sure.**
β=0.5	You will **liquidate the firm half of the times.**
β=1	You will **liquidate the firm's asset for sure.**

There are **23 firms** and **25 banks**. One firm will be matched with one bank through the following process: at the beginning of each round, the firms will post the liquidation probability β they desire from the bank. Once all firms post their choices of β**, all of these firms' choices of β will be made known to all banks, as well as all firms' repayment history (i.e. how many times they defaulted when project was successful).**

Upon observing these offers, as a bank you can either **select one and only one offer** or **reject all offers**. **If you accept an offer, you will be matched with the firm making that offer**. If besides you, there are other banks accepting that firm's offer, a bank will be selected randomly as the matched partner of the firm. Thus, you stand equal chance of being selected as the matched partner of the firm.

If you reject all offers, or if you fail to match with the firm you choose because there is more than one bank accepting the firm's offer, you will **enter the next cycle** of matching process involving all banks and firms that are still unmatched. In this next cycle, firms will have to come up with a new offer of β. If you are still unmatched in this new cycle, a new matching cycle will repeat again until all firms are matched. We allow for **5 matching cycles**. **If you are still not matched after 5 cycles, you will have to sit out for that particular round and your earnings from this round are 6 points**. Once the round is completed, you can again participate in the next round.

Once matched, you, as a bank, will extend the loan of 14 points to your matched firm, which will then invest in project X. Project X lasts for 2 periods. The following is the information on the payoff obtained at the end of each period.

At the end of period 1, Project X **may or may not be successful**. The chance of success is **2/3** and the project yields **30 points**. The chance of failure is **1/3** and the project yields **0 points**.

	Success	Failure
Probability	2/3	1/3
Firm Payoff	30	0

If **project X fails**, the firm will have **no choice** other than to **default on the loan received**. If **project X is successful**, the firm will make a **repayment** decision, with the following choices:

1. **Repaying** you with **18 points**
2. Not repaying you (**0 points**), i.e. to **default on the loan**

If your matched firm **repays** the loan, you **cannot liquidate** the firm and the firm will proceed to project period 2.

If your matched firm decides **not to repay you regardless of whether project X is successful or not**, you **will liquidate** your matched firm **according to the probability β** agreed earlier. The computer will decide whether there is liquidation or not randomly according to the β value.

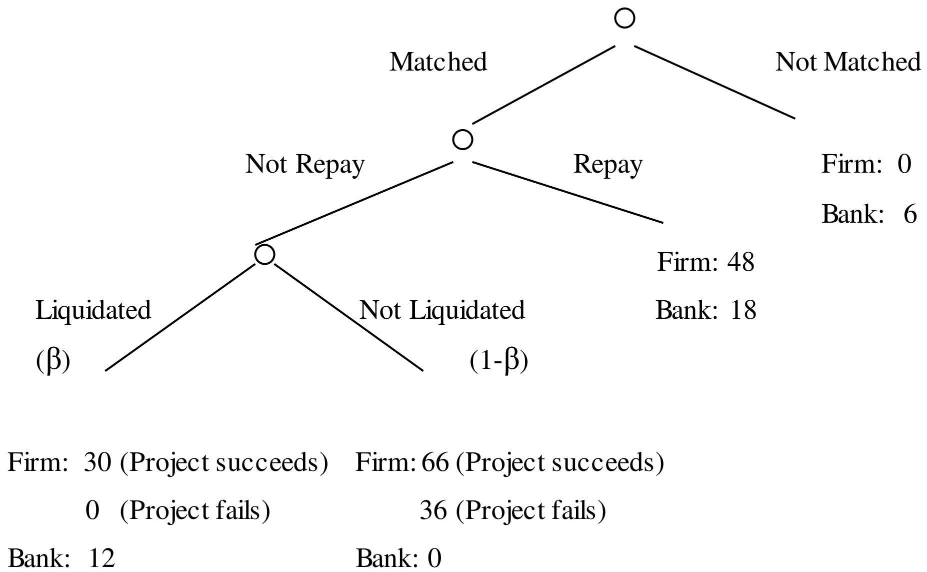
- If the firm **is liquidated**, you will get **12 points** from selling the firm's asset and the game ends.

	Liquidated	Not Liquidated
Probability	β	1-β
Bank Payoff	12	0

- If your matched firm **avoids liquidation**, the firm manager will be able to continue with the project to the second period and **the project yields a sure payoff of 36 points regardless of whether the project was successful or failed in period 1**.

At the end of the first period you will be informed about the outcome of the firm's project, i.e. whether it is successful or failed and the respective payoff.

<u>You may refer to the following Game Tree for the calculation of your earnings:</u>



<u>Note</u>: The first two rounds are trial rounds. Your results from these two rounds are not going to be taken into account in the drawing of your binding repayment.

===***===

Chapter 3 Altruistic Punishment in the Face of Direct Externality and Selfish Temptation

3.1 Introduction

> *"Everyone is a moon, and has a dark side which he never shows to anybody." — (Mark Twain, 1897, Chapter LXVI)*

> *"People couldn't become truly holy, he said, unless they also had the opportunity to be definitively wicked." — (Pratchett and Gaiman, 2006, p.39)*

Ample experimental evidence has shown that we care not only about our own welfare but also societal welfare (see Cooper and Kagel, 2009 for a review). We are willing to take actions that directly benefit others at our own expense and to impose costly sanctions on people deemed to have violated social norms even when the violation does not affect us personally (Fehr and Fischbacher, 2004). Anecdotal evidence on this type of pro-social behavior abounds. For example, there is the story of two passers-by in Stockton, California who chased down a suspected hit-and-run driver and managed to subdue him while they waited for police to arrive (Fitzgerald, 2014). These two "citizen heroes" did not back off even though the suspect threatened to hurt them. They performed this good deed at some personal cost without reaping direct (material) benefit from it.

In other circumstances, however, we are capable of doing the opposite: behaving selfishly or taking actions that inflict loss on others. For example, there is the story of a migrant worker window cleaner in Fuzhou, a city in Fujian Province, China (January 8, 2014). One day while she was cleaning the eighth floor window of a high rise building, she accidentally dropped her fanny-pack carrying 6000 RMB worth of notes, which is a huge sum for a person with an income as low as hers. Her money scattered all over the busy street below and passing pedestrians quickly grabbed the bills. She desperately attempted to collect her money, but only managed to collect 300 RMB (Zhou, 2014). Under normal circumstances, these individuals who took her money could very well be kind-hearted individuals who are willing to take actions benefiting others at their own expense just like those passersby in Stockton in our earlier example.

These two examples presented above illustrate two contrasting sides of human beings: the good and the ugly. In many situations, however, it is the good side rather than the ugly side that is more apparent, either because we are often compelled to never show our ugly side or simply because our environment provides little opportunity for us to reveal our wicked side. In the migrant worker example, money falling from the sky is a rare occurrence. Consequently, the opportunistic behavior of the pedestrians is rarely seen under normal circumstances. We therefore cannot know whether people would be truly kind and remain righteous in the face of a wicked opportunity to benefit at the expense of another. The true measure of people's righteousness is their ability to remain righteous in the face of a wicked temptation.

To a certain extent, the environment depicted in many experimental studies showing that people are kind-hearted, inherently care about upholding social (fairness) norms and are willing to impose costly punishment on people who have treated someone else unfairly is also devoid of any opportunity to behave wickedly. Many of these studies typically use the third-party punishment game (TPP) developed by Fehr and Fischbacher (2004) and its variants. In its simplest setup, the TPP game involves three roles: the dictator, the recipient, and the third-party observer. In this game, the dictator must decide how to divide the money between themselves and the recipient who is not able to reject the chosen division. The third-party observer, upon observing the amount the dictator chooses to give to the recipient, can decide whether or not to impose costly monetary punishment on the dictator. Every dollar spent on punishment by the third party reduces the payoff of the dictator by a certain amount. This type of punishment, which does not bring any direct benefit to the person who imposes it, is often referred to as altruistic punishment (Fehr and Fischbacher, 2004; Marlowe et al., 2008). It has been suggested that TPP is a unique trait developed in human society (see for example Riedl et al., 2012). Fehr and

Fischbacher (2004) further found that the severity of the punishment imposed by the third party is positively related to the dictator's degree of selfishness.[18]

However, in those studies, the third party's alternative option to imposing costly punishment on the unfair dictator is to play ignorant. Although playing ignorant is not a loyal act, it also cannot be definitively categorized as a devious act. For some people, being ignorant in this circumstance may be acceptable. They might think that since they were not the person who harmed the unfortunate victim, they cannot be held accountable. At the very least, it is not the same as stealing someone else's money. In nearly all of the TPP game experiments, the third party is never presented with an option that allows them to reveal their ugly side. Consequently, it may be misleading to interpret these TPP experiments as indicating people's inherent pro-sociality and intrinsic desire to uphold social norms.

This chapter has two objectives. *First*, we aim to test this line of inquiry using a controlled laboratory experiment. We hope to shed light on the interplay between the good and the ugly sides of human beings. Specifically, we want to explore the following questions: How robust is the evidence that people care about the enforcement of social norms in the presence of an 'evil' temptation? We often come across similar questions like this albeit in different contexts. For instance, when a government official runs a clean office, people might ask whether he is genuinely a man of integrity or he is not corrupt because his working environment gives little room for him to be corrupt. Further, we are also interested in examining the heterogeneity of people's pro-social motives. How large is the proportion of people who are willing to impose altruistic punishment in the absence of evil temptation relative to those whose behavior is akin to that of the rational economic man? Are there differences in the patterns of altruistic punishment, for example the punishment schedule belonging to those who are willing to impose punishment on the dictator? Would there be people who impose anti-social punishment on the dictator even if the dictator

[18] The incentive to be righteous by punishing the unfair dictator has been shown to be robust to the variations of game settings. For instance, it exists in the dictator game (Fehr and Fischbacher, 2004 ; Nikiforakis and Mitchell, 2013) the trust game (Charness et al., 2008) and the ultimatum game (Chavez and Bicchieri, 2013). It has also been shown that the incentive to punish is higher in larger societies where there is less chance for people to interact in the future (Marlowe et al., 2007 Heinrich et al., 2010).

has already treated the recipient fairly? How significant is the proportion of people who are tempted to switch from the ‘good’ side to the ‘ugly’ side?

Second, we explore the motivation behind altruistic punishment in the absence of an ‘evil’ temptation. There are various motives behind altruistic punishment; however, they can be generally grouped into two broad motives. The first one is the *retributive justice* motive. The main focus here is on *the dictator’s wrongdoing* that warrants punishment. The punishment assigns moral blame to the wrongdoer and it should “fit the crime.” The second is the *distributive justice* motive. The main focus here is on *the resulting unequal payoff distribution* due to the dictator’s action. The punishment given to the dictator is aimed at reducing the extent of the payoff inequality. This second motive is in line with the inequality-aversion based social-preferences model of Levine (1998), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Charness and Rabin (2002).

We use the following experimental treatments in our analysis. Our control treatment (which we term “the TPTT treatment”) is a modified TPP game. In contrast to the standard TPP game, the dictator’s payoff reduction from the punishment imposed by the third party is automatically transferred to compensate the recipient. Thus, the difference between this treatment and the standard third party punishment game is the presence of direct positive externality of punishment on the recipient. This externality alters the payoff distribution of the three players.

The second treatment, the TPTR treatment, is different from the TPTT in one crucial aspect. That is, in the TPTR treatment, we allow the third party to expropriate the dictator’s payoff reduction. Specifically, the third party must decide how much of the payoff reduction will be transferred to the recipient and how much of it is to be expropriated. Essentially, by allowing the third party to enrich herself and comparing these results with those obtained under the TPTT treatment, we are able to examine the stability of the third party’s incentive to uphold social norms in the face of a wicked temptation. If the third party truly has righteous intentions, she would not be tempted to enrich herself using the pretext of punishing the unfair dictator. Also, we would be able to evaluate whether or not people who behave as a rational economic man, for example, and impose zero punishment on the dictator in the TPTT treatment would be tempted to punish a kind-hearted dictator who shares her endowment generously with her recipient in the TPTR treatment. If the third party succumbs to the temptation, then the motive behind punishment may not be truly noble.[19]

Our third treatment is the standard TPP game of Fehr and Fischbacher (2004). By comparing the results obtained from the TPP treatment with those obtained from the TPTT treatment, we are able to analyze the underlying motives behind the third party's altruistic punishment. Some explanations have been put forward to explain the apparent pro-social and norm-upholding behavior of the third party. Fehr and Fischbacher (2004) argue that the violation of social norms by the dictator may trigger an emotional response from the third party that will induce the third party to impose punishment on the dictator. This emotional response can take the form of anger or guilt (Nelissen and Zeelenberg, 2009). Carpenter et al. (2004) argue that social reciprocity motive could also explain why the third party is willing to engage in altruistic punishment. That is, people use costly punishment to express their social disapproval towards the norm-violating action. In another paper, Carpenter and Matthews (2012) argue that altruistic punishment is an expression of the third party's indignation towards the dictator who has violated social norms. The aforementioned explanations center on the 'wrongdoing' of the dictator, which may exert emotions, such as anger and indignation, feelings of solidarity towards the victim, and a strong feeling that social justice needs to occur. However, the third party's punishment can also center on the consequences the dictator's action has on income distribution rather than the action itself. The argument rests on a premise that people are averse to payoff inequality between themselves and others and care about fairness. Consequently, people are willing

[19] In most studies using the TPP game, it is common to associate the punishment meted by the third party on the dictator with glorifying labels such as "altruistic" (Fehr and Fischbacher, 2004; Nelissen and Zeelenberg, 2009) or "moralistic" (Kurzban, 2007). There are almost no studies that explore the less bright side of the TPP, except Leibbrandt and López-Pérez (2011). They show that the TPP has some undesirable features. The third party often punishes the dictator if the later becomes the "richest" party, even though perhaps the dictator has made a socially efficient or egalitarian allocation. This evidence points to a situation where envy could potentially induce the third party to impose anti-social punishment.

to spend resources to mitigate any advantageous or disadvantageous payoff inequality. Some of the notable papers along this line of reasoning are by Levine (1998), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Charness and Rabin (2002).

Our TPP and TPTT treatments can be compared to see whether the punishment behavior of the third parties is consistent with the retributive motive or the distributive motive. Under the TPTT treatment, there is a direct positive externality of the altruistic punishment in the form of an automatic transfer to the recipient. The presence of the transfer in the TPTT treatment alters the relative payoff distribution after punishment compared to that of the TPP treatment in which such a transfer is absent. If what matters is the dictator's wrongdoing rather than the resulting payoff distribution after the dictator's allocation has been made, then the amount of punishment imposed by the third party in the two treatments should not be statistically significantly different. Otherwise, if what matters is the payoff distribution rather than the dictator's wrongdoing, we should expect that the amount of punishment imposed in both treatments would be statistically significantly different.[20]

Our results can be summarized as follows. From the comparison between the TPTT treatment and the TPTR treatment, we find that the behavior of the third party changes drastically when they are presented with an opportunity to misappropriate the reduction in the dictator's payoff. The third party diverts a large sum of money and only transfers a small sum to the recipient. Also surprisingly, the third party tends to impose a large punishment on the dictator in the face of evil temptation to a level that approaches the maximum allowable punishment amount. This severe punishment is also imposed on dictators who have treated the recipient fairly and generously. In the TPTT treatment where it is not possible for the third party to misappropriate the money and the money is

[20] Note that it is possible that the two motives are present simultaneously under both the TPP and the TPTT treatments. But notice that the single factor that distinguishes the two treatments is the presence of transfer to the recipient in the TPTT treatment, and this transfer would make it easier for the third party to decrease the payoff inequality between the dictator and the recipient, and as such the third party need not punish as much. So even if both motives are present simultaneously, if the distributive justice motive is more important than the retributive justice motive, we should still expect the punishment amount under these two treatments to be different.

completely transferred to the recipient, the third party shows great care and altruism towards the recipient. However, the third party ignores the recipient's disadvantageous position and even deliberately inflicts losses on the dictator in order to enrich herself in the presence of selfish temptation. This behavioral inconsistency suggests that one needs to be cautious in drawing a sweeping conclusion that people are pro-social, care about norm adherence, and are willing to impose costly intervention to help uphold social norms.

Next, from the comparison between TPP treatment and TPTT treatment, we find that the presence of a direct externality has no significant effect on the punishment spending of the third party. It suggests that retributive justice seems to motivate altruistic punishment. The result squares with findings in public goods experiments, which suggest it is the norm violation itself that triggers punishment rather than payoff differences (Price et al., 2002; Carpenter et al., 2004; Carpenter and Matthews, 2012. Interestingly, although the punishment spending is not affected, the propensity of the third party to punish the dictator becomes significantly higher in the presence of a direct externality in the TPTT treatment.

This chapter contributes to the literature in several ways. First, to the best of our knowledge, this is the first study to test the robustness of the third party's behavior to uphold social norms in the face of selfish temptation. The results from our study show that individuals are capable of doing good and bad things. When presented with the temptation to misbehave at the expense of adhering to social norms, people often fall into the temptation. Their behavior changes drastically. People behave pro-socially in the absence of temptation, but become very selfish and are often willing to misuse their ability to punish and extract the dictator's payoffs for their own benefits. This drastic change in behavior, to some degree, is in line with the finding of List's (2007) study on dictators' giving behavior. He shows that fewer dictators give positive amounts when the dictators' choice set is extended to include taking away the recipient's monetary entitlement.

However, there is a notable difference between this study and List's (2007). In List's experimental design the agent (the dictator) is required to choose either to give (to be generous) or to take (to be mean). However in our design, we allow the third party to appear to care about the enforcement of social norms by punishing the (unfair) dictator and transferring the reduction in the dictator's payoff to the recipient, while at the same time also enriching herself by allocating some portion of the reduction in the dictator's payoff to herself. Thus, the third party does not have to completely give up their money-earning opportunity to show their good side. In contrast, List (2007) found that the dictator has to choose an allocation amount to give to the recipient from a range of possible allocation amounts in which the lower limit is extended to include negative allocation amounts (i.e., taking away instead of giving away money to the recipient). In such as case, the dictator must choose to either be good or be bad. Our design thus allows us to explore to what extent the presence of selfish temptation changes the third party's pro-social behavior and the intricate moral battle between the desire to be good and the temptation to be bad. We are also able to evaluate whether the more altruistic third party would behave differently from the less altruistic third party in the face of temptation.

This chapter is also related to the strand of literature on corruption. It is closely related to Xiao (2013), which studies profit-seeking punishment using sender-receiver games. Our findings on the third party's punishment behavior is in line with findings in Xiao (2013), which suggests that the enforcer may punish for no reasonable clause and tends to abuse the authority if such punishments bring material benefits to the enforcer. However, the study in this chapter is different from Xiao (2013). Our design gives the third party a wider span of choices. Not only does the third party decide whether to punish but also the level of punishment. It minimizes the distortion on behavior due to limited choices. In addition, all players' payoff functions are common knowledge and the third party is able to adjust the other two players' payoffs directly, which may motivate the punishment behavior differently.

Second, this chapter presents an important attempt to unveil the motivation behind TPP. Most papers on TPP focus on the punishment behavior of the third party and its robustness under various alternative settings, rather than on the underlying motive behind the third party punishment.

This chapter is organized as follows: Section 3.2 describes our experimental design and procedures, Section 3.3 presents the results, and Section 3.4 concludes the chapter.

3.2 Experimental Design

The basic game structure used in this chapter is Fehr and Fischbacher's (2004) TPP game. The game is played by three players: the dictator (player A), the recipient (player B), and the third party (player C). First, the dictator must decide on the division of money between him or herself and the recipient. The recipient cannot reject the division of money made by the dictator. The third party is able to observe the dictator's decision. This third party can then decide whether or not to punish the dictator. Punishment is costly for the third party, and if imposed, the dictator's payoff will be reduced. In this study, every point of punishment reduces the dictator's payoff by 2 points. In contrast to the standard TPP game of Fehr and Fischbacher (2004), where the punishment solely decreases the dictator's payoff, we vary the consequences of the punishment across treatments.

We have three experimental treatments. In the next sub-section, we present the details of the three experimental treatments. Because our TPTT treatment and TPTR treatment are built upon the standard TPP game, we begin with the TPP treatment.

3.2.1. The Experimental Treatments

The TPP treatment

Player A (the dictator) is given an endowment of S$20 (roughly equivalent to 16USD) and has to decide how much of it to transfer to Player B (the recipient). The amount of transfer (t) has to be within the range of $0 \leq t \leq 20$, in increments of S$1. Player B has no initial endowment and only receives the amount transferred by Player A. Player C, who is an observer, is endowed with S$15 and can use her endowment to punish Player A. The punishment amount (p) imposed by Player C has to be within the range of $0 \leq p \leq 15$, in increments of S$1. Player C can spend any amount of her endowment to punish Player A, but the punishment must not reduce Player A's payoff to negative.

Every S$1 spent on punishment by Player C will reduce Player A's payoff by S$2. Player A's earnings can be expressed as:

$$\pi_A = 20 - t - 2p \tag{3.1}$$

Player B's payoff can be expressed as:

$$\pi_B = t \tag{3.2}$$

Player C's payoff is:

$$\pi_C = 15 - p \tag{3.3}$$

We adopt the strategy method to elicit Player C's punishment spending conditional on Player A's allocation. Thus, Player C must decide how much to spend on punishment for all possible transfer amounts from Player A. There are 21 possible transfer amounts and each Player C will have to make 21 decisions. This strategy method is widely used in previous studies (see, for instance, Fehr and Fischbacher, 2004; Almenberg et al., 2010; Fischbacher and Gächter, 2010; Nikiforakis and Mitchell, 2013).[21] As an alternative to using the strategy method, one can also use the direct response method, in which Player C is directly asked how much punishment she wants to impose on Player A upon observing the actual amount of allocation made by Player A to Player B. The strategy method provides us a more complete picture of the third party's punishment preferences and allows us to do direct comparison of punishment behavior across treatments without relying on the actual transfer amount made by Player A.

The TPTT Treatment

In this treatment, the reduction in Player A's payoff due to the punishment imposed by the third party would be automatically transferred to Player B. The punishment, therefore, creates a direct positive externality on Player B. Similar to the previous treatment, S$1 of punishment would reduce Player A's payoff by $2. If Player C spends $0 \le p \le 15$ on punishment, each player's earnings can be expressed as:

[21] Brandts and Charness (2011) show that both methods give the same qualitative results.

$$\pi_A = 20 - t - 2p \tag{3.4}$$

$$\pi_B = t + 2p \tag{3.5}$$

$$\pi_C = 15 - p \tag{3.6}$$

The TPTR Treatment

In this treatment, Player C has the option of sharing the reduction in Player A's payoff with Player B. Note that in both the TPTT and TPTR treatments, the implication of punishment on social efficiency is the same. That is, every dollar spent by Player C to punish Player A will increase the total welfare by 1 dollar. The resulting reduction in player A's payoff due to punishment is denoted as $2p$. A part of $2p$, $0 \leq r \leq 2p$, is redistributed to player B and the remaining $(2p - r)$ is appropriated by player C. Each player's earnings can then be expressed as:

$$\pi_A = 20 - t - 2p \tag{3.7}$$

$$\pi_B = t + r \tag{3.8}$$

$$\pi_C = 15 + p - r \tag{3.9}$$

In the TPTT and TPTR treatments, the punishment inflicts direct positive externality to Player B. If the third party is genuinely motivated by good intentions for example, the desire to uphold social norms and punish the dictator who has treated the recipient unfairly the third party should not be tempted to divert the reduction in the dictator's payoff to her own benefit. In addition, if the third party never punishes the fair-minded dictator in the TPTT treatment, she should also behave the same and not be tempted to punish the fair-minded dictator in order to increase the potential payoff diverted to herself.

By comparing the TPP treatment and the control treatment (e.g. the TPTT treatment), we are able to investigate how the presence of direct externalities from punishment alter the third party's incentive to punish the dictator. That is, under the TPTT treatment, knowing that the punishment imposed on the dictator directly benefits the recipient and hence improves the payoff distribution between the dictator and the recipient, would the amount of punishment imposed by the third party be affected? If yes, in which direction would the effect be?

3.2.2. The Experimental Procedures

The experiment was programmed using z-tree (Fischbacher, 2007) and conducted at Nanyang Technological University. There were in total 69 student participants from various majors ranging from science, engineering, business, economics, and the social sciences. The duration of the experiment was approximately one hour. The average earnings, inclusive of a S$3 show-up fee, were S$12. Table 3.1 presents the summary of statistics.

Table 3.1. Descriptive statistics

	The change of payoff from $1 on punishment			No. of subjects	Mean age	Male
	Dictator	Recipient	Third party			
TPP	-2	0	-1	21	21.67	62%
TPTT	-2	+2	-1	24	22.21	54%
TPTR	-2	R[a]	1-R[a]	24	22.38	54%

[a] The third party decides how to share the 2 dollars between the recipient and herself. R represents the amount transferred to the recipient.

We had a between-subject experimental design and ran a separate session for each treatment. We provided hard copy written instructions, read the instructions aloud, and presented the instructions on each subjects' computer screen.[22] Each participant went through three rounds in total and played a different role in each round. This role reversal method has been widely used in the literature and it seems have no impact on subjects' behavior (Charness and Rabin, 2005; Coffman, 2011). We elicited player C's punishment decision using the strategy method. In total, player C had to make 21 punishment decisions conditional on all possible transfers that could be made by player A. The group composition was reshuffled from round to round. Participants were only identified by their subject ID, and all decisions were done anonymously. To avoid any order effect, we randomly assigned the sequence of roles to

[22] The experimental instructions can be found in the appendix. The wording ("punishment" used in instructions) and the illustration chart followed coffman (2011).

subjects across the three rounds. No feedback was given to players until all rounds were completed.

The endowment amounts allocated to all subjects and the experimental procedure were made common knowledge to the participants. Before the experiment started, but after the experimental instructions were read aloud, we asked participants to answer control questions to test their understanding of the experiment they were about to go through. It was only when all control questions had been answered correctly that they were allowed to proceed to the main experiment. After the experiment was completed, we asked participants to complete a post-experiment questionnaire with questions on their socio-demographic characteristics and their decisions during the experiment.

3.2.3. The Experimental Predictions

In the TPP and TPTT treatments, if player C is a rational payoff maximizer, she would never punish player A because punishment is costly and generates no future benefit. When a rational payoff-maximizer Player C is presented with an opportunity to enrich herself, she would punish player A to the maximum allowable punishment and divert the whole reduction in player A's payoff to herself.[23]

If behavior remains stable, we should expect that the proportion of the rational payoff-maximizer participants in the TPTR treatment would be similar to that in the TPTT treatment. What differs is only the manner in which the payoff is maximized. In the TPTT treatment it is achieved by not punishing player A, while in the TPTR treatment it is achieved by punishing player A to the maximum allowable punishment amount. Also, if participants care about upholding social norms and are genuinely altruistic, adding an opportunity to enrich oneself should not diminish their incentive to help the recipient by transferring the money resulting from player A's payoff reduction to player B.

[23]Since the third party is not allowed to deduct the dictator's payoff to a negative value, there will be a maximum amount the third party can spend on punishments for each transfer level. Note that all inputs are restricted to whole numbers, the dictator thus has either 0 or 1 dollar left in the maximum punishment case. For instance, if the dictator transfers 10 dollars, she would have 0 left after the maximum punishment spending (5 dollars) is imposed on her; if the dictator transfers 9, she would have 1 dollar left after the maximum punishment spending (5 dollars) is imposed on her.

In other words, the proportions of altruistic punishers in the TPTT and TPTR treatments should not be different. All in all, relative to the TPTT treatment, we have the following prediction in the TPTR treatment.

Prediction 1: *If the third party is inherently altruistic and cares about the manner with which player A allocates money to player B, the presence of selfish temptation in the TPTR treatment should not alter her altruistic punishment behavior in the TPTR treatment relative to that of the TPTT treatment.*

Next, by comparing the TPP and TPTT treatments, we can evaluate the motive behind altruistic punishments imposed by Player C on Player A. If what matters for player C is the 'wrongdoing' of Player A (i.e., Player A's allocative behavior towards Player B), then conditional on Player A's allocation towards Player B we should expect no statistical difference between Player C's punishment amounts in the TPP and TPTT treatments. The presence of the direct externality of punishment in the TPTT treatment should not draw any wedge in the punishment behavior in both treatments. However, if the underlying motive is the distributive motive, then we should expect the third party's punishment behavior in both treatments conditional on Player A's allocation towards Player B to be statistically different.

This is because the compensation to the recipient resulting from the punishment imposed by Player C on Player A does affect the payoff distribution between all parties and the overall social efficiency. However, it should be noted that the direction of influence can go either way depending on whether or not the third party's goal is to wipe out the payoff difference between the dictator and the recipient or whether the third party is willing to expand her budget to punish knowing that punishment becomes more effective in enforcing the norm. All in all, we have the following prediction.

Prediction 2: *If the underlying explanation is retributive motive instead of distributive motive, or any other motives that are different from the retributive motive, then the punishment spending in the TPTT treatment, conditional on the amount transferred by player A, should not be different from that in the TPP treatment.*

3.3 The Experimental Results

This section is organized as follows. First, we discuss the patterns of punishment spending imposed by the third party, conditional on the dictator's transfer amount to the recipient, across treatments. Second, we evaluate the third party's decision of whether to punish the dictator across treatments. Third, we discuss the third party's redistribution behavior in the TPTR treatment. Fourth, we analyze the individual heterogeneity of third parties in terms of their social behaviors. We classify them into different types based on their altruistic punishment and redistribution patterns. Finally, we briefly discuss the impact of the anticipation of the third party punishment on the dictator's giving behavior across treatments.

3.3.1. Punishment Spending

It should be noted that the punishment in the TPP and TPTT treatments does not generate any direct personal benefit to the third party, and it is also costly for the third party. In the TPTT treatment, the beneficiary of the dictator's payoff reduction is the recipient, while in the TPP treatment no one receives any direct benefit from punishment. In the TPTR treatment, however, the third party can potentially benefit from the punishment. Consequently, the third party's motivation for punishing the dictator in the TPTR treatment is rather unclear. Specifically, the punishment can be driven by altruism or selfishness, or even a combination of both. To infer this punishment motive we have to evaluate the way the third party redistributes the punishment-induced windfall money. Figure 3.1 presents the third party's average punishment spending conditional on the dictator's transfer amount for all treatments (Panel A) and the average compensation the recipient receives in TPTT and TPTR (Panel B).[24] Recall that there is an upper limit for the punishment amount that the third party can impose on the dictator. That is, the third party cannot punish the dictator to the extent that the dictator's net payoff after transfer and punishment becomes negative. The maximum spending amount on punishment, conditional

[24] Recall that, in the TPTT treatment, the compensation is an automatic by-product of punishment, while in the TPTR treatment, it is an outcome of the redistribution decision made by the third party.

on the dictator's transfer amount, is captured by the graph connecting the crossed points shown in panel A.

It can be seen that the average punishment spending in all treatments decreases as the dictator's transfer amount increases. The same result was also found by Fehr and Fischbacher (2004) and Nikiforakis and Mitchell (2013). The punishment spending patterns under the TPP and TPTT treatments are remarkably similar: They decrease with the dictator's transfer amount to the recipient. Specifically, when the dictator's transfer amount is larger than the egalitarian transfer amount of S$10, the punishment amount is virtually zero. In contrast, the average punishment spending in the TPTR treatment exhibits a markedly different pattern from that in the TPP and TPTT treatments. In particular, the third party punishes the dictator much more severely. The punishment spending is very close to the maximum allowable punishment spending. Even when the dictator has already given more than S$10 to the recipient, the dictator is still being punished relatively severely by the third party.

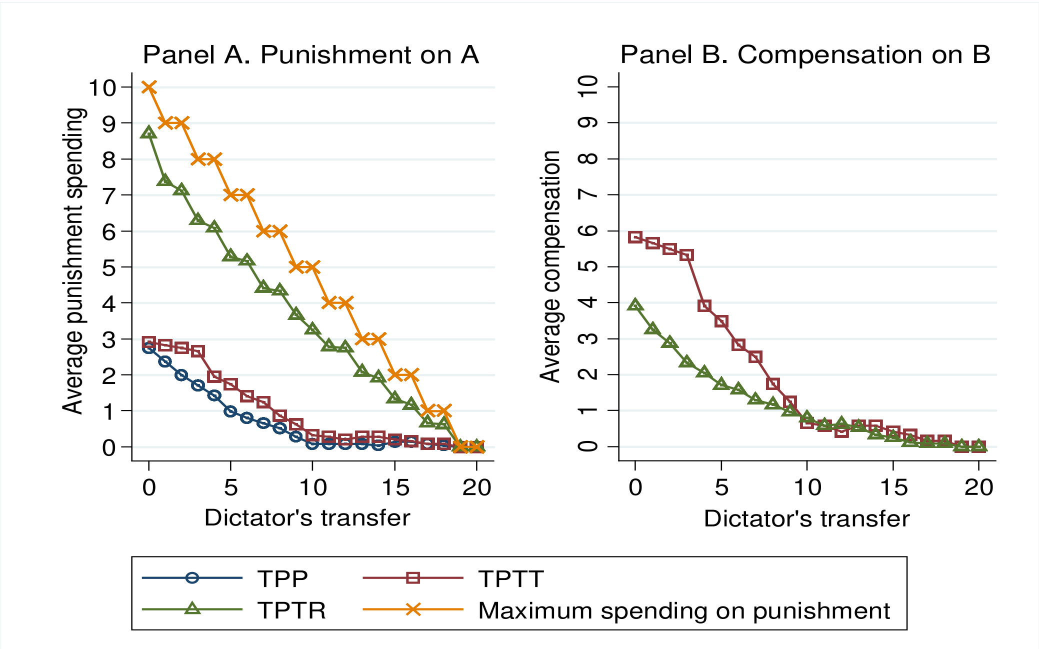


Figure 3.1. The average punishment spending and compensation by treatment

When we compare the punishment spending pattern, conditional on the dictator's transfer amount, in the TPTR treatment to that in the TPTT treatment, we find that the presence of selfish temptation in the TPTR treatment induces the third party to punish the dictator more severely (see Panel A of Figure 3.1). Next, we evaluate how the third party allocates the windfall money generated from the punishment in the TPTR treatment. How much of it is given as compensation to the recipient? Is the allocation to the recipient in the TPTR treatment, which is solely at the discretion of the third party, larger than that in the TPTT treatment, which is not at the discretion of the third party? Panel B of Figure 3.1 shows that the compensation amount to the recipient, conditional on the dictator's transfer amount, is much lower in the TPTR treatment than that in the TPTT treatment. This suggests that the main motivation underlying the punishment in the TPTR treatment is the third party's desire to redirect the windfall money to her rather than the desire to help the recipient improve her relative income vis-à-vis the dictator. The evidence does not seem to support prediction 1.

Next, the difference between the average punishment spending in the TPP treatment and the TPTT treatment is not statistically significant (the p-value from the Wilcoxon-Mann-Whitney test is $p = 0.298$).[25] The average punishment spending in the TPTT treatment, where the punishment would automatically bring direct benefit to the recipient and alter the relative income distribution between the dictator and the recipient by changing both agents' payoff, is remarkably similar to that in the TPP treatment in which the punishment alters the relative income distribution between the dictator and the recipient by changing the dictator's payoff only. This result is consistent with Prediction 2. It is noteworthy that efficiency improves in addition to the fact that the reduction from punishment goes to the recipient. As a result, the third party might be willing to punish the dictator for retributive justice and efficiency considerations. Intriguingly, these considerations do not lead to more punishment.

Table 3.2 presents the OLS regressions of the punishment spending in all treatments. The dependent variable is the third party's punishment spending

[25] Using each subject's punishment spending as an independent observation, we also conducted non-parametric tests conditioned on each transfer level. It shows the difference is not significant for most of transfer levels (Wilcoxon-Mann-Whitney, two-sided, for TPP vs. TPTT, only marginally significant when the transfer is 6 with $p = 0.095$).

conditional on each amount of the dictator's transfer to the recipient. The right hand side regressors include: D_{neg}, which is the negative deviation of the dictator's allocation from the egalitarian allocation (S$10), D_{pos}, which is the positive deviation of the dictator's allocation from the egalitarian allocation, *Transfer as Dictator*, which is the allocation amount to the recipient when the third party plays as the dictator, and *Round 2* and *Round 3*, which are dummy variables indicating in which round the subject plays as the third party. We also include socio-demographic control variables: age, gender, nationality, and religion.

The variable *Transfer as Dictator* serves as our proxy for the altruistic tendency of the third party. That is, the amount that the third party allocates to the recipient when she plays as the dictator in one of the three rounds.

Table 3.2. The third party's punishment spending in individual treatments

Dependent variable: punishment spending						
	TPP		TPTT		TPTR	
D_{neg}	0.251***	0.251***	0.280***	0.283***	0.497***	0.490***
	(0.058)	(0.058)	(0.055)	(0.058)	(0.033)	(0.034)
D_{pos}	0.007	0.007	-0.036*	-0.040*	-0.330***	-0.331***
	(0.015)	(0.015)	(0.021)	(0.023)	(0.045)	(0.047)
Transfer as dictator	0.129***	0.102***	0.127***	0.113***	-0.182***	-0.234***
	(0.046)	(0.029)	(0.028)	(0.033)	(0.069)	(0.066)
Round 2		0.293		0.582		-0.018
		(0.231)		(0.391)		(0.617)
Round 3		0.092		-0.330		0.609
		(0.224)		(0.338)		(0.533)
Socio-demographic variables	No	Yes	No	Yes	No	Yes
Constant	-0.669**	-5.121***	-0.201	3.318	4.038***	4.194**
	(0.260)	(1.500)	(0.214)	(3.136)	(0.386)	(2.007)
Observations	441	441	504	462	504	483
R^2	0.405	0.475	0.421	0.480	0.751	0.785

Note: OLS regressions with clustering on individuals. Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

It can be observed that the coefficient of D_{neg} is always positive and highly statistically significant at the 1% significance level in all treatments. For every one dollar increase in the negative deviation (the extent to which the dictator's transfer falls below the egalitarian amount), the punishment amount imposed by the third party increases by around 25 cents in the TPP treatment, 28 cents in the TPTT treatment, and 49 cents in the TPTR treatment. The effect of the negative deviation on punishment spending in the TPTR treatment is much stronger than that in the other two treatments.

The positive deviation does not have any significant effect on punishment spending in the TPP treatment, but it has a small but significant negative effect in the TPTT treatment in which the punishment would automatically lead to an increase in the recipient's payoff. In the TPTR treatment, the coefficient of this independent variable has a negative sign, is statistically significant, and is of a substantial magnitude.

As mentioned earlier, *Transfer as dictator* measures the degree of altruism of the third party. It has a positive effect on punishment spending in the TPP and TPTT treatments suggesting that the more altruistic third party in these two treatments imposes larger punishment spending, which is fairly intuitive. However, in the TPTR treatment where the third party faces temptation, the more altruistic the third party is the lower the spending on punishment becomes. This suggests an interesting behavior. By lowering punishment spending, the third party lowers the reduction in the dictator's payoff due to the punishment, thereby reducing the money available. By doing so, the third party is able to credibly commit to reducing the misappropriation. All in all, our results show that the third party is aware of the misappropriation temptation and would want to pre-commit to reducing the misappropriation by restricting the amount of money available for misappropriation. The more altruistic the third party is, the stronger the incentive to pre-commit would be.

Table 3.3 presents the results of the OLS regressions for the pairwise treatment comparisons of the punishment spending. The first comparison is between the TPTT and TPTR treatments. Recall that both treatments are identical except that in the TPTR treatment the reduction in the dictator's payoff due to the punishment is given to the third party to decide on how it should be redistributed. In the TPTT treatment, it is automatically given in full to the recipient. This comparison allows us to evaluate prediction 1.

The treatment dummy *TPTR* is positive and statistically significant indicating that the presence of selfish temptation in the TPTR treatment increases the punishment spending relative to that in the baseline TPTT treatment. Relative to the baseline TPTT treatment, the more altruistic the third party is (when *Transfer as dictator* increases) the smaller the punishment amount will be in the TPTR treatment. This evidence shows that the more altruistic third party commits to lessening the misappropriation by reducing punishment spending.

Table 3.3. The third party's punishment spending for comparisons between treatments

Dependent variable: punishment spending				
	TPTT vs TPTR		TPP vs TPTT	
D_{neg}	0.280***	0.283***	0.251***	0.251***
	(0.054)	(0.057)	(0.057)	(0.057)
D_{pos}	-0.036*	-0.040*	0.007	0.007
	(0.021)	(0.022)	(0.015)	(0.015)
Transfer as dictator	0.127***	0.127***	0.129***	0.131***
	(0.028)	(0.035)	(0.045)	(0.048)
Round 2		0.131		0.286
		(0.390)		(0.250)
Round 3		0.089		-0.247
		(0.337)		(0.190)
TPTR	4.239***	4.342***		
	(0.436)	(0.423)		
D_{neg} × TPTR	0.217***	0.207***		
	(0.063)	(0.066)		
D_{pos} × TPTR	-0.294***	-0.291***		
	(0.049)	(0.052)		
Transfer as dictator × TPTR	-0.309***	-0.340***		
	(0.074)	(0.079)		
TPTT			0.468	0.566
			(0.333)	(0.401)
D_{neg} × TPTT			0.029	0.033
			(0.078)	(0.081)
D_{pos} × TPTT			-0.043*	-0.047*
			(0.026)	(0.027)
Transfer as dictator × TPTT			-0.002	-0.021
			(0.053)	(0.057)
Socio-demographic variables	No	Yes	No	Yes
Constant	-0.201	0.773	-0.669***	-0.942
	(0.211)	(2.199)	(0.257)	(2.015)
Observations	1,008	945	945	903
R^2	0.738	0.742	0.420	0.438

Note: OLS regressions with clustering on individuals. Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

The second comparison is between the TPP and TPTT treatments. The treatment dummy $TPTT$ is not statistically significant. Controlling for other variables and conditional on the dictator sharing less than or equal to half of her endowment (S$10), the presence of compensation as a by-product of punishment does not affect the third party's punishment spending in the TPTT treatment relative to that in the TPP treatment. This can be seen from the coefficient of $D_{neg} \times TPTT$, which is not statistically significant. However, when the dictator gives more than half of her endowment, being in the TPTT treatment has a negative effect on punishment spending. This can be seen from the coefficient of $D_{pos} \times TPTT$, which is negative and marginally statistically significant.

3.3.2. The Third Party's Propensity to Punish

Figure 3.2 presents the third party's propensity to punish in all treatments, which is measured by the proportion of subjects imposing non-zero punishment on the dictator, conditional on the dictator's transfer amount. One common feature among the three treatments is the presence of a kink in the distribution of the propensity to punish, which is situated at the dictator's egalitarian transfer amount of S$10. In all treatments, the propensity to punish is smaller when the transfer amount is above S$10 than when it is below S$10. However, it can be seen that the kink is more pronounced in the absence of other punishment motives than the altruistic punishment motive. All in all, the egalitarian transfer norm does seem to play an important role in driving the altruistic punishment decision in all treatments, although it is weaker in the TPTR treatment because of the presence of the misappropriation temptation.[26]

[26] Using the same set of control variables as in the punishment spending, we ran Probit regressions of the punishment decision for individual treatments as well as pairwise comparisons between treatments. Most of the conclusions in the punishment spending still stand for the punishment decision with the following exceptions. The positive deviation always has significant negative effects on the propensity to punish in all treatments. The magnitude of the coefficient of positive deviation is comparable to the coefficient of negative deviation in the TPP and TPTT treatments and the magnitude of the difference is larger in the TPTR treatment, which is consistent with observations shown in Figure 3.2. As for pairwise comparisons, in contrast with the non-significant treatment dummy of the TPTT treatment for the punishment spending, it is positive and significant for the punishment decision suggesting that the presence of direct externality increases the third party's propensity to punish.

There seems to be a larger proportion of third parties who punish even if the dictator's transfer exceeds 10 in the TPTT treatment compared to the TPP treatment, which might be explained by efficiency considerations.

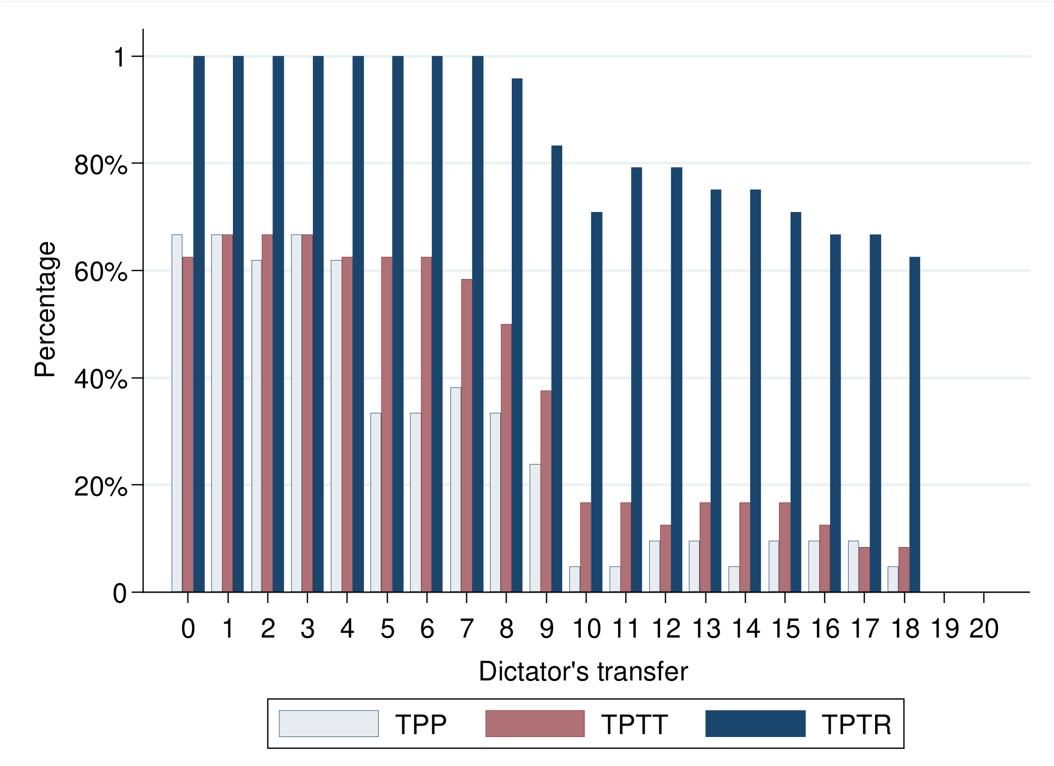


Figure 3.2. The percentage of the third parties imposing non-zero punishment by treatment

3.3.3. The Third Party's Redistribution Decision in the TPTR Treatment

Figure 3.3 depicts the average percentage of money distributed to the recipient and to the third party herself at each transfer amount. It suggests that the selfish motive dominates the altruistic motive. Overall, the third party on average redistributes more than 70% of the money available to herself and leaves only a small proportion to the recipient. It is fair to say that punishment spending in the TPTR treatment is mostly driven by selfishness. Nevertheless, the positive redistribution amount to the recipient does seem to suggest that altruistic considerations still play a modestly important role. The percentage of money redistributed to the recipient begins to decrease when the dictator's transfer amount is closer to S$10. Beyond S$10, the amount redistributed to the recipient becomes much lower, presumably because the third party thinks that the recipient has already received a sizable amount of money from the dictator and hence needs no further redistribution.

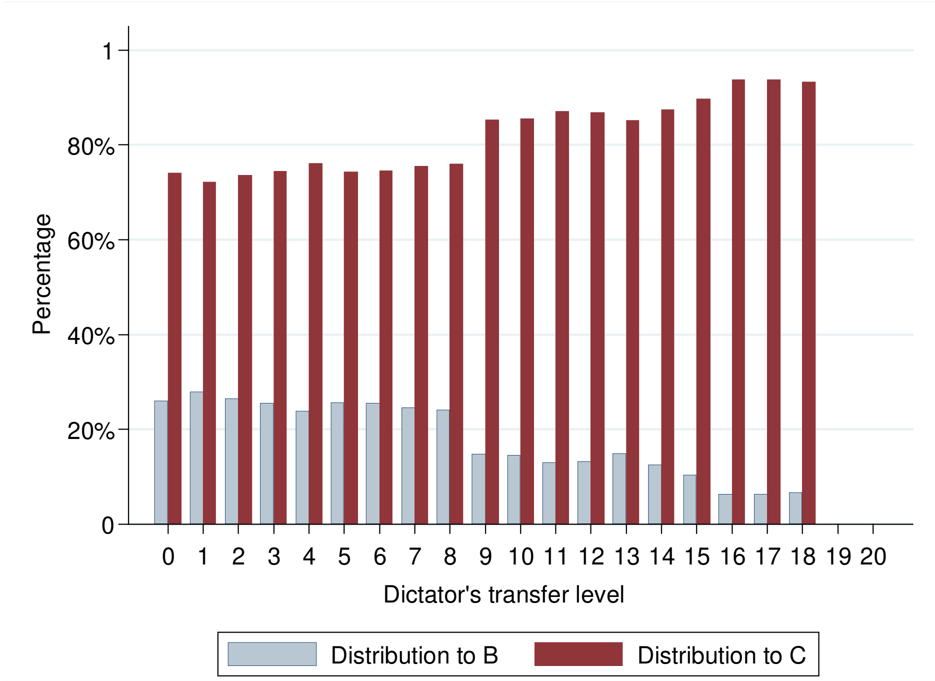


Figure 3.3. Share of money taken distributed to recipient and third party

In Table 3.4, we provide the OLS regression analyses of the determinants of the third party's decision to split money between the recipient and the third party. The dependent variable is the ratio of the money distributed to the recipient to the total amount of money available. The explanatory variables are, among others: 1) the ratio of punishment spending amount to the maximum allowable punishment amount (*Spending ratio*), which measures the extent to which the third party exploits the opportunity to enlarge the amount of money available for redistribution; 2) the negative and positive deviations of the dictator's transfer amount from the egalitarian sharing amount of S$10 ($D_{neg}$ *and* D_{pos}); 3) the measure of how altruistic the third party is (*Transfer as dictator*); 4) control variables for rounds and the third party's socio-demographic factors.

It can be seen that the regression coefficient D_{neg} has a positive sign, suggesting that a wider negative deviation from the egalitarian sharing amount would result in an increase of the proportion of money redistributed to the recipient. However, the coefficient is only marginally statistically significant and becomes non-significant when we control for rounds and socio-demographic variables.

The variable *Spending ratio* is highly significant. It implies that the proportion of money distributed to the recipient, to a large extent, can be predicted by the extent to which the third party exploits the opportunity to enlarge the size of the punishment-induced windfall money. That is, the closer the punishment spending is to the maximum allowable punishment amount, the less money is redistributed to the recipient. Thus, *Spending ratio* captures the intention of the third party behind the punishment.

Table 3.4. The percentage of distribution to the recipient by the third party in TPTR

Dependent variable: Distribution to the recipient / Total money taken		
D_{neg}	0.010*	0.008
	(0.005)	(0.005)
D_{pos}	-0.001	0.001
	(0.005)	(0.005)
Transfer as dictator	-0.002	-0.007
	(0.007)	(0.011)
Spending ratio	-0.691***	-0.773***
	(0.118)	(0.147)
Round 2		0.090
		(0.079)
Round 3		0.018
		(0.055)
Socio-demographic variables	No	Yes
Constant	0.750***	0.883***
	(0.104)	(0.307)
Observations	390	373
R^2	0.470	0.610

Note: OLS regressions with clustering on individuals. Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

3.3.4. Individual Heterogeneity

Thus far we have discussed the average behavior of the third party. In this section, we extend this discussion by analyzing the third party's individual heterogeneity in their punishment behavior. For this purpose, we evaluate the individual punishment spending pattern across all treatments, find common patterns, and classify the third parties into several types.

Table 3.5 presents the resulting type classification of the third parties in all treatments. Specifically, we categorize the third parties into five types. The first is the *own payoff maximizer* type (the selfish type). Notice that the definition of this type in the TPP and TPTT treatments is different from the one in the TPTR treatment. That is, in the TPP and TPTT treatments, it refers to the third parties who spend zero on punishment regardless of the dictator's transfer amount. In the TPTR treatment, it refers to the third parties who punish the dictator to the maximum allowable punishment amount and divert all the punishment-induced windfall money to themselves.

The second type is the *altruistic norm enforcer* type, which refers to third parties who only punish the dictator when the dictator's transfer amount is below S$10 in all treatments. Their punishment spending is roughly monotonic in the dictator's transfer amount. In addition, in the TPTR treatment the resulting punishment-induced windfall money is fully distributed to the recipient.

The third type is the *partially altruistic norm enforcer* type, which refers to third parties in the TPTR treatment who demonstrate similar punishment patterns as the altruistic norm enforcer types, but distribute some portion of the punishment-induced windfall money to themselves. While these third parties care about the recipient's well-being to some degree, especially when the dictator's transfer negatively deviates from the egalitarian sharing amount of S$10, they could not resist the temptation to siphon some portion of the money to themselves.

The fourth type is the *altruistic punishment lover*, which refers to third parties who punish the dictator even if the dictator has been generous by giving more than S$10 with punishment spending that monotonically decreases with the dictator's transfer amount.

The fifth type is the *irregular* type, which refers to third parties who do not fall into any of the other categories. A similar classification was used in the study of second- and third-party punishment by Fehr and Fischbacher (2004).

Table 3.5. Categorization of third parties based on their decision pattern

Type of third party	TPP (n=21)	TPTT (n=24)	TPTR (n=24)
Own payoff maximizer	28.57%	33.33%	41.67%
Altruistic norm enforcer	52.38%	41.67%	4.17%
Partially altruistic norm enforcer	-	-	12.50%
Altruistic punishment lover	4.76%	16.67%	12.50%
Irregular	14.29%	8.33%	29.17%

As shown in Table 3.5, the proportions of the third parties who belong to the *own payoff maximizer* in the TPP and TPTT treatments are relatively similar. The largest type in these TPP and TPTT treatments is the altruistic norm enforcer. Thus, in general it can be said that the majority of third parties in the TPP and TPTT treatments, in which selfish temptation is absent, care about upholding egalitarian norms and would be willing to impose costly punishment on dictators who violate the egalitarian sharing norm. In contrast, in the TPTR treatment in which selfish temptation is present, the proportion of *own-payoff maximizer* (selfish) type is the greatest (41.67%). The size of the *altruistic norm enforcer* types in the TPTR treatment decreases drastically to 4%. The other third parties can be categorized into the other three types. Many do punish unfair dictators and transfer money to the recipients, but at the same time they also divert some money to themselves. All in all, the evidence shows that the dominant type changes from the *altruistic norm enforcer* type to the *own-payoff maximizer* (selfish) type in the presence of selfish temptation.

In what follows, we present the third parties' decision patterns in the TPTT and TPTR treatments. Figure 3.4 and Figure 3.5 depict the individual punishment spending patterns in the TPTT and TPTR treatments. Figure 3.6 shows the individual redistribution patterns in the TPTR treatment. It can be seen from Figure 3.4 that the amount of punishment spending for the majority of third parties in the TPTT treatment is far from the maximum allowable amount of punishment spending. We do, however, observe a clear kink at the egalitarian sharing point (S$10) in third parties' punishment behavior.

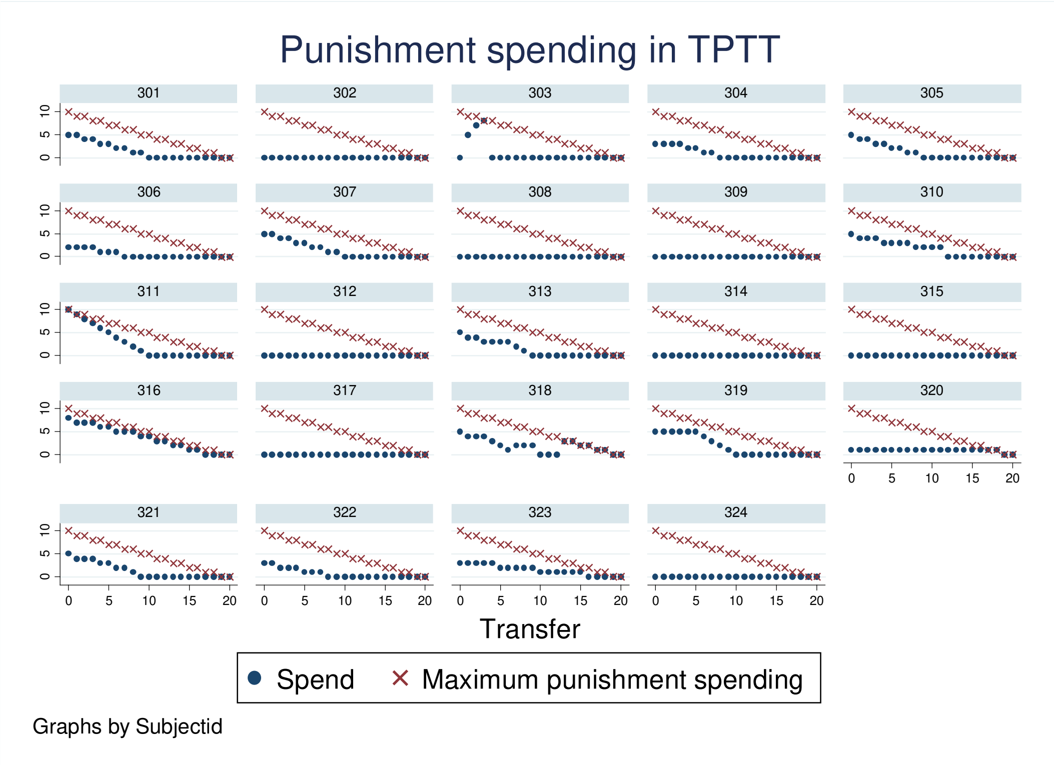


Figure 3.4. The punishment spending in TPTT by subject

The classification of the third parties in TPTR shown in Table 3.5 is based on punishment spending patterns and redistribution patterns. Since each third party is required to make two decisions, some intriguing questions arise. What does the pattern of redistribution look like? Do the third parties who impose the maximum allowable punishment always divert all of the punishment-induced windfall money to themselves? How severe are the third parties who care about the egalitarian sharing norm and at the same time are also interested in extracting some money during the redistribution process when they punish the dictator? How do they redistribute the money between the recipient and themselves? To answer these questions, we examine each third party's punishment spending and redistribution decision in the TPTR treatment in detail (see Figure 3.5 and Figure 3.6).

It can be seen from Figure 3.5 that a majority of third parties would punish the dictator to the maximum allowable punishment amount. The punishment line for these third parties coincides or almost coincides with the maximum allowable punishment line. Figure 3.6 shows that these third parties would always behave selfishly and redistribute all the punishment induced windfall money to themselves. The only exception is the third party with the ID of 415. She imposes the maximum allowable punishment on the dictator, but chooses to split the money accrued from the reduction in the dictator's payoff equally between the recipient and herself, rather than keeping it all for herself.

The third party with the ID of 405 is the only third party who can resist the selfish temptation completely. This third party imposes punishment when the dictator gives less than S$10 to the recipient and her punishment schedule has a negative slope. The money accrued from punishment would then be redistributed in full to the recipient. Once the dictator gives more than S$10, this third party does not impose any punishment.

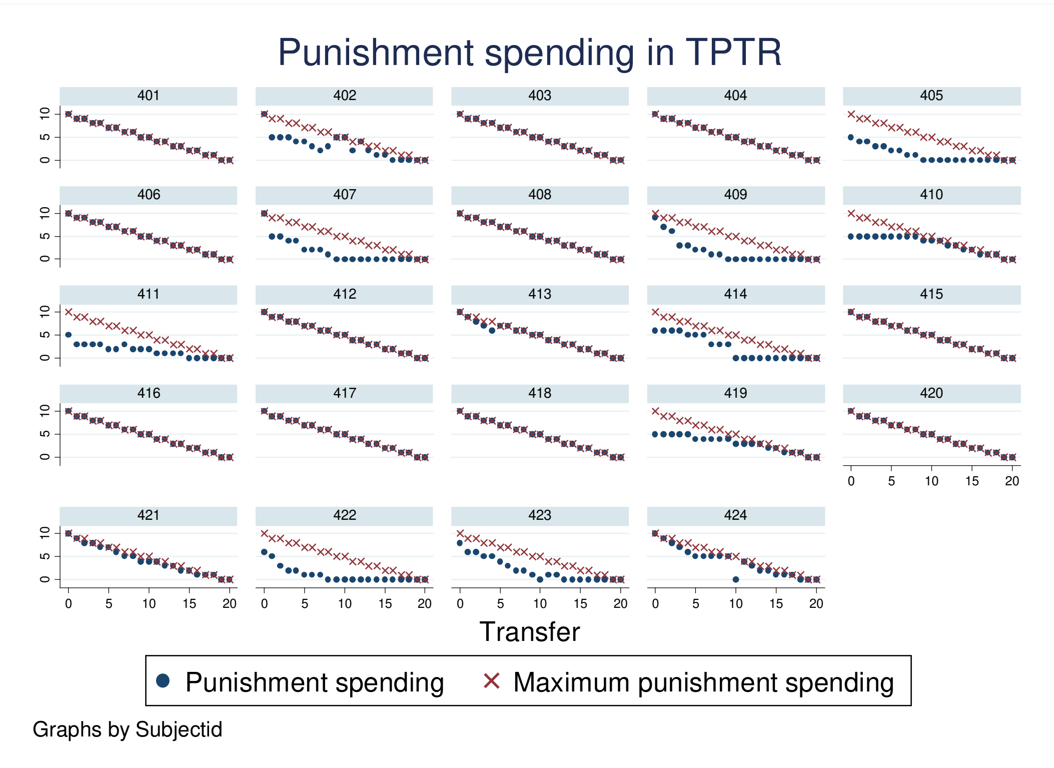


Figure 3.5. The punishment spending in TPTR by subject

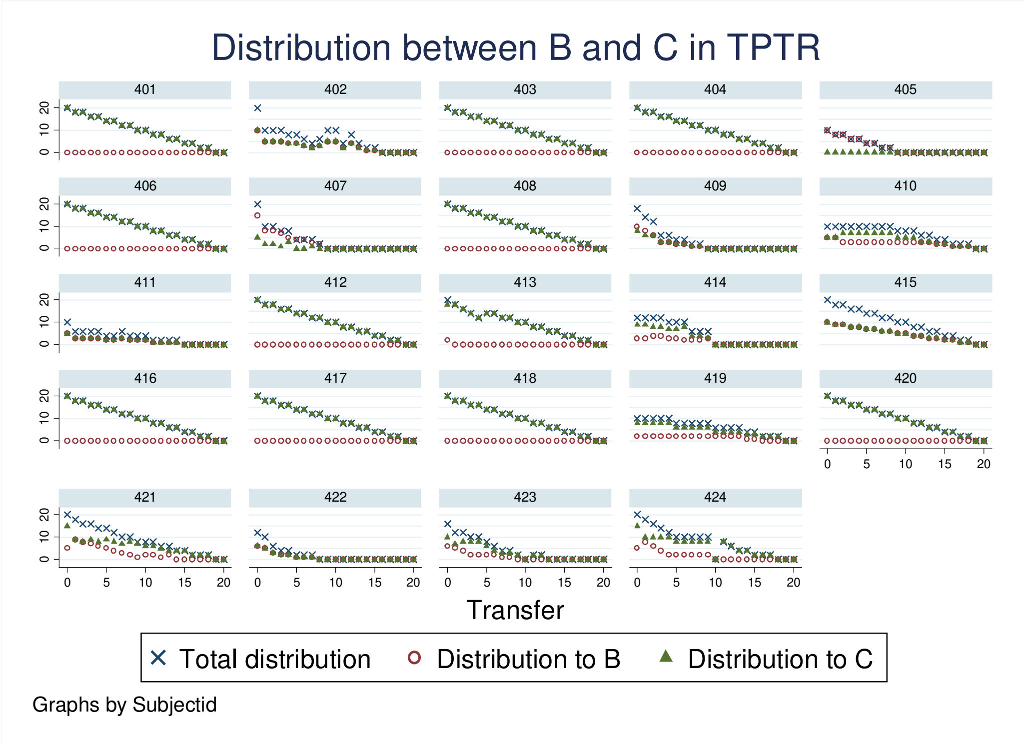


Figure 3.6. The redistribution in TPTR by subject

3.3.5. The Dictator's Transfer Amount

Though the dictator's behavior is not the main focus of this chapter, we briefly present it in this sub-section. Table 3.6 shows the dictator's average transfer amount by treatment. The average transfer is around S$ 5 in all treatments. It is highest in the TPP treatment and lowest in the TPTT treatment. However, none of the pairwise comparisons of the dictator's transfer described is statistically significant. The p-value from the two-sided Wilcoxon-Mann-Whitney test for the equality of transfer in the TPP and TPTT treatments is 0.488, and in the TPTT and TPTR treatments it is 0.525. All in all, this result shows that the dictator's transfer amounts in all treatments are not statistically significantly different from each other.

Table 3.6. Dictator's transfer by treatment

	TPP	TPTT	TPTR
Mean transfer (std. dev.)	5.29 (3.21)	4.42 (4.02)	4.96 (3.99)
No. of subjects transfer <10	19	19	18
Observations	21	24	24

3.4 Concluding Remarks

Rather than providing further evidence on the willingness of the third party to impose altruistic punishment, this chapter addresses the following two objectives. First, we are interested in investigating to what extent the third party's willingness to punish is motivated by kind intentions. We do so by giving the third party an opportunity to misappropriate the punishment-induced windfall money. Specifically, we compare a setting (the TPTT setting) in which the reduction in the dictator's payoff resulting from the punishment imposed by a third party is automatically transferred to the recipient, with another setting (the TPTR setting) in which the third party can split the dictator's payoff reduction money resulting from the punishment between herself and the recipient. Essentially, in the TPTR setting, we subject the third party to a selfish temptation. If the third party wants to be completely selfish and cares only about maximizing her own payoff, then she can punish the dictator to the maximum allowable punishment amount and then divert all of the punishment-induced windfall money to herself. We find that the punishment spending and the propensity to punish are much higher in the treatment in which the selfish temptation is present.

By analyzing the redistribution decision, we find that punishment in the TPTR setting is mostly driven by selfishness because the majority of the money taken is redirected to the third party. The egalitarian sharing norm has a rather limited effect in influencing the third party's redistribution decision on how to divide the punishment-induced windfall money between the recipient and herself. However, the sharing norm is influential in affecting the third party's decision on whether or not to distribute some positive amount to the recipient. All in all, our result shows that the incentive to uphold norms is weak and can be undermined by the presence of wicked temptation to enrich oneself. We find that a significant proportion of the third parties succumb to this temptation. Interestingly, more altruistic third parties impose lesser punishments, suggesting that they are aware of the temptation and want to pre-commit to lessening the misappropriation by reducing the punishment-induced windfall money available.

Our second objective is to explore the underlying motivation behind altruistic punishment. Two broad motives are examined: the retributive motive and the distributive motive. The retributive motive refers to punishment behavior that is motivated by and focuses on the dictator's action. The distributive motive refers to punishment behavior that is motivated by the income distributional impact of the dictator's action. To do so, we compare the standard TPP setting whereby the reduction in the dictator's payoff resulting from punishment simply evaporates with the TPTT setting we described in the earlier paragraph. In contrast to the TPTT setting, in the TPP setting the third party can adjust the relative income distribution between the dictator and the recipient only by changing the dictator's payoff. If the third party is motivated only by the retributive motive, then we should expect that the punishment spending amounts in the two settings should not be statistically significantly different.

Our results show that the punishment spending in the TPTT setting is not statistically significantly different from that in the TPP setting, although the introduction of the direct externality of punishment on the recipient does have a significantly positive effect on the third party's propensity to punish. This result seems to be in line with the retributive justice motive for altruistic punishment.

3.5 Appendix: Experimental Instructions

Treatment TPTR

General Instructions

Welcome to this study!

This is a study on decision making and it will last for about 1 hour. The study will be conducted in an anonymous fashion. Neither administrators nor other participants will be able to link your decisions to your personal identity. Please rest assured that your anonymity as a decision maker will be strictly preserved.

If you have questions at any time during the study, please raise your hand and we will attend to you privately. Communication between participants is strictly prohibited. Furthermore, once the study starts, please do not use the computer for any other purposes than participating in this study. Please turn your mobile into silence mode as well.

In this study, you will make a series of decisions in a game. Your payoffs from participating in this study are composed of two parts. First, you will receive a show-up fee regardless of your decisions. Second, your earnings in the game are determined by your own decisions and the decision of others. It is crucial for you to understand how to play the game.

The Game-Specific Instructions

The game will be played in groups of three. In each group, there is a Player A, a Player B and a Player C.

Player A will split **$20** with Player B. That is, Player A will decide how to split $20, and Player A and Player B will be paid according to Player A's decision. Player B has no say in the decision and has to accept Player A's decision.

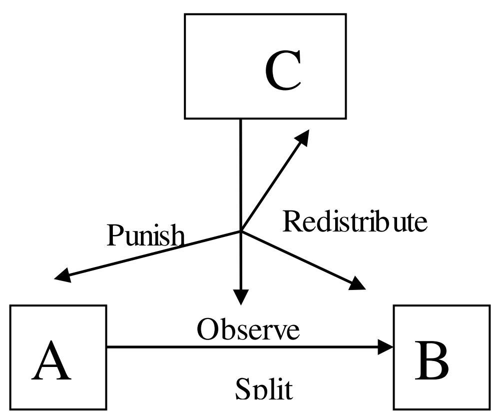
Player C observes Player A's decision and can choose to punish Player A. However, doing so is costly for Player C. Player C has an initial endowment of

$15. Player C can spend X dollar(s) to punish Player A. Every X dollar(s) punishment imposed on Player A will reduce Player A's payoff by 2X dollars. Player C can then decide on how to redistribute 2X dollars between him (her) self and Player B.

You will play the game for three rounds. In each round, you will be randomly grouped with two other participants in the lab. In every round you will be randomly re-matched. That means your group composition will change from round to round. You will be assigned to one role (A, or B, or C) for each round, and in the next two rounds you will play different roles. So you will assume each of these three roles only once.

At the end of each period, you will not be informed of the other players' decisions. You will know what happens in each round when all three rounds have been played. By then, the computer will randomly choose one round as a binding round. Your payoff will be determined by the decisions made by you and your group members in that particular round. All three rounds are equally likely to be the binding round, so make thoughtful decision in all three rounds.

The following graph describes the roles of each player.

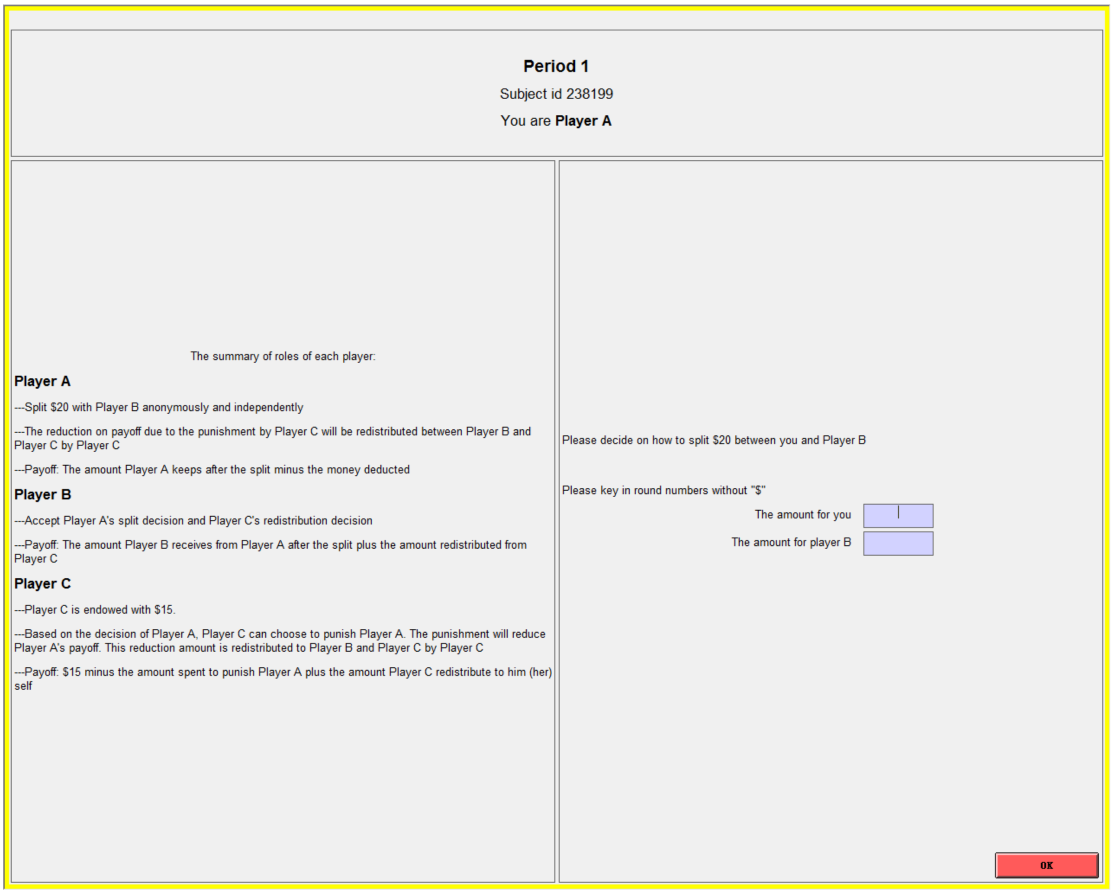


The summary of roles of each Player:

Player A

- Split $20 with Player B anonymously and independently
- The reduction on payoff due to the punishment by Player C will be redistributed between Player B and Player C by Player C
- Payoff: The amount Player A keeps after the split minus the money deducted

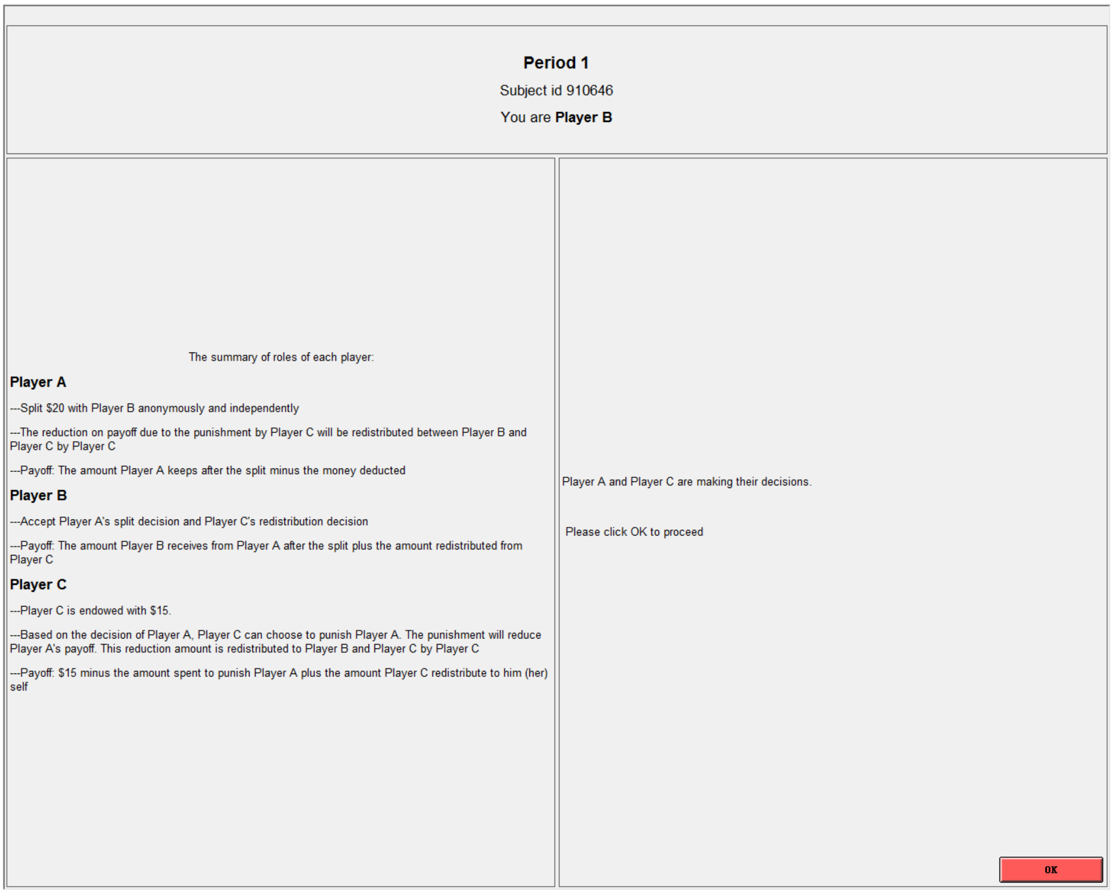
A sample screen of Player A:



Player B

- Accept Player A's split decision and Player C's redistribution decision
- Payoff: The amount Player B receives from Player A after the split plus the amount redistributed from Player C

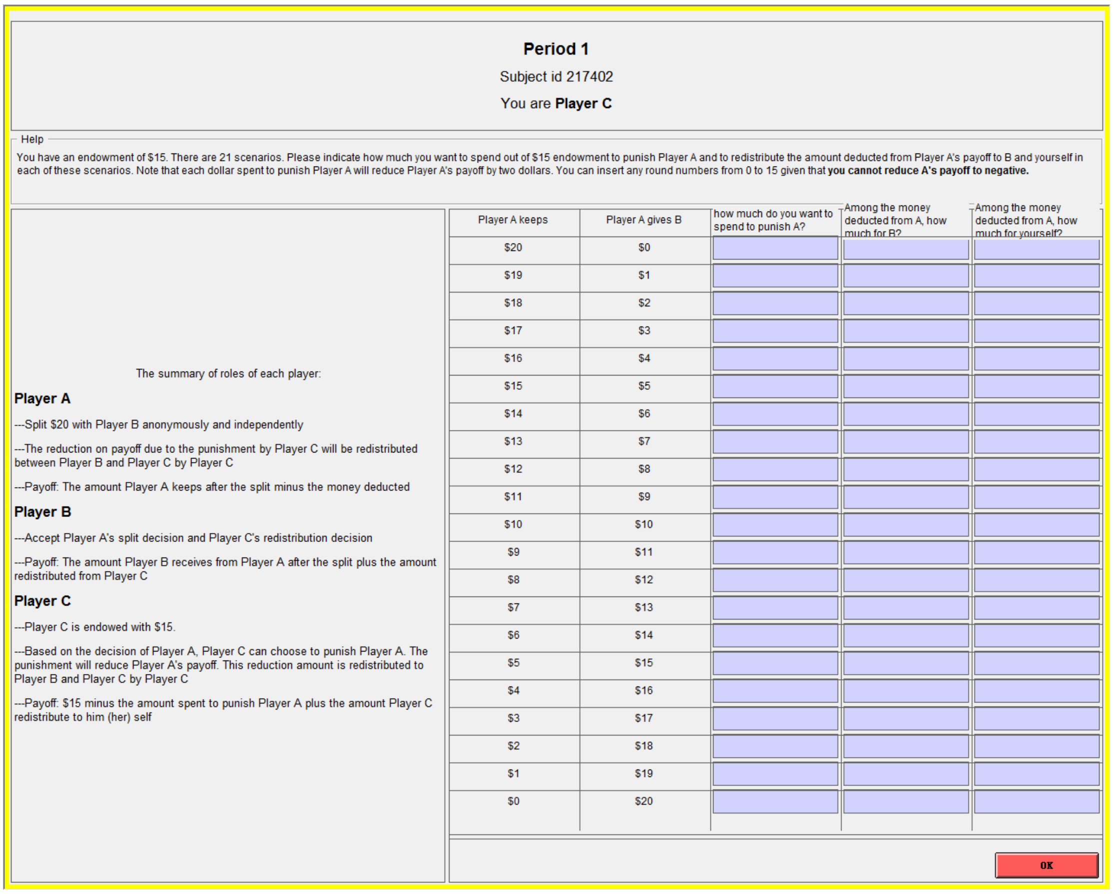
A sample screen of Player B:



Player C

- Player C is endowed with $15
- Based on the decision of Player A, Player C can choose to punish Player A. The punishment will reduce Player A's payoff. This reduction amount is redistributed to Player B and Player C by Player C.
- Payoff: $15 minus the amount spent to punish Player A plus the amount Player C redistribute to him (her) self.

A sample screen of Player C:



Period 1

Subject id 217402

You are **Player C**

Help

You have an endowment of $15. There are 21 scenarios. Please indicate how much you want to spend out of $15 endowment to punish Player A and to redistribute the amount deducted from Player A's payoff to B and yourself in each of these scenarios. Note that each dollar spent to punish Player A will reduce Player A's payoff by two dollars. You can insert any round numbers from 0 to 15 given that **you cannot reduce A's payoff to negative.**

The summary of roles of each player:

Player A

---Split $20 with Player B anonymously and independently

---The reduction on payoff due to the punishment by Player C will be redistributed between Player B and Player C by Player C

---Payoff: The amount Player A keeps after the split minus the money deducted

Player B

---Accept Player A's split decision and Player C's redistribution decision

---Payoff: The amount Player B receives from Player A after the split plus the amount redistributed from Player C

Player C

---Player C is endowed with $15.

---Based on the decision of Player A, Player C can choose to punish Player A. The punishment will reduce Player A's payoff. This reduction amount is redistributed to Player B and Player C by Player C

---Payoff: $15 minus the amount spent to punish Player A plus the amount Player C redistribute to him (her) self

Player A keeps	Player A gives B	how much do you want to spend to punish A?	Among the money deducted from A, how much for B?	Among the money deducted from A, how much for yourself?
$20	$0			
$19	$1			
$18	$2			
$17	$3			
$16	$4			
$15	$5			
$14	$6			
$13	$7			
$12	$8			
$11	$9			
$10	$10			
$9	$11			
$8	$12			
$7	$13			
$6	$14			
$5	$15			
$4	$16			
$3	$17			
$2	$18			
$1	$19			
$0	$20			

OK

For player C, there are **21 scenarios**. Please indicate how much you want to spend out of $15 endowment to punish Player A and to redistribute the amount deducted from Player A's payoff to B and him (her) self **in each of these scenarios**.

Note that each dollar spent to punish Player A will reduce Player A's payoff by two dollars. You can insert any whole number from 0 to 15 given that **you cannot reduce player A's payoff to negative.** After you made your decisions in all scenarios shown above, your decisions will be matched against Player A's actual allocation to Player B. Your payoff as Player C would be $15 minus the amount you spent on punishing Player A plus the amount Player C keeps in the redistribution. You don't know which scenario is going to happen; **hence you should make your decisions in all scenarios carefully.**

Chapter 4 Information Transparency for Equilibrium Selection in Coordination Games: An Experimental Study

4.1 Introduction

Humans are social beings; they interact constantly. They influence others and are influenced by others in their social surroundings. In order to survive, they need to cooperate with others and learn how to balance their individual interests with collective interests. To achieve a more socially desirable outcome they must learn how to coordinate their actions with other people to arrive at mutually consistent actions. Coordination is often hampered by the failure to develop an implicit understanding of others' intention and the inability to trust others' inclinations to take a mutually desirable action. In such situations, people often prefer to take a safer alternative action that yields a smaller payoff. If everybody behaves in the same way, society is stuck with a less socially desirable outcome.

To illustrate this further, consider the example of an airline company whose workers must prepare an airplane for departure. A timely departure requires the successful coordination of effort by multiple parties such as flight attendants, gate agents, mechanics, caterers, etc. If any party underperforms, the other departments' endeavors to achieve on-time departure are wasted. Overall performance is thus dragged down by the underperforming party and the flight is delayed. If one party is unsure about the commitment of other parties, and is sufficiently risk averse, that party would respond by underperforming too. The desirable outcome can only be achieved if all parties are able to coordinate their efforts and are willing to trust others' willingness to choose a mutually consistent action. They should be able to communicate seamlessly with others; however, communication is often hampered by location separation and hierarchical organization structure. How to make agents coordinate tacitly for an efficient outcome is thus an important question.

In game theory, a coordination setting like the one described above is usually depicted as coordination games. These are a class of games with pure strategy Nash equilibria that can be Pareto ranked. Equilibrium analysis of such games lacks the predictive power to foresee which equilibrium the players might end up with. Harsanyi and Selten (1988) propose the refinement concepts of payoff-dominance and risk-dominance. An equilibrium is said to be payoff-dominant if it is Pareto-superior relative to other equilibria. An equilibrium whose deviation losses are greatest is said to be risk-dominant. In other words, strategies constituting the risk-dominant equilibrium are relatively safe under strategic uncertainty. Harsanyi and Selten (1988) argue that payoff dominance should serve as the equilibrium-selection criterion in coordination games.

Coordination games with multiple Pareto-ranked equilibria have received the lion's share of attention in the experimental economics literature (see Devetag and Ortmann, 2007). However, ample experimental evidence has shown that people often fail to coordinate on the payoff-dominant equilibrium (Cooper et al., 1990), especially when the group size is large (Van Huyck et al., 1990). The secure and inefficient risk-dominant equilibrium is more likely to be chosen, leading to coordination failure.[27]

Numerous experimental studies have explored the determinants of coordination outcome. Various factors that affect the ability of subjects to overcome coordination failure have been examined. These factors can roughly be classified into two categories. The first category comprises those factors that are related to the payoff structure (i.e., the magnitudes of payoffs obtained from coordination); the second comprises contextual factors, such as the subject matching protocol, subject experience, availability of information, and the presence of external advice.

Some studies have examined the role of differences in payoff structure. Battalio et al. (2001) show that there is a positive correlation between the occurrence of risk-dominant equilibrium and the optimization premium, the latter defined as the pecuniary incentive accrued from the difference between the payoff from the best response and the payoff from the inferior response to a

[27]To be consistent with the literature, we refer to cases where people coordinate on the inefficient equilibrium instead of the Pareto-superior equilibrium as coordination failures.

partner's strategy. Battalio et al. vary the size of the optimization premium across experimental treatments. Brandts and Cooper (2006a) show that increasing the bonus rate for successful coordination effectively reduces coordination failure even in the presence of a history of coordination failure. The effect sustains regardless of the magnitude and the duration of the bonus. This suggests that the presence of financial incentives that enhance the payoffs from coordination, even if they are only offered temporarily, can achieve a more efficient outcome than that achieved without financial incentives. Crawford et al. (2008) show that in coordination games, where the games are made realistic by describing them using salience labels (focal points),[28] even a small asymmetry in payoffs accrued to players is enough to soften the effectiveness of a salience label in enhancing coordination. Relative to a treatment where payoffs are symmetric, the presence of a small payoff asymmetry would increase the incidence of coordination failures by around 30 percent.

The above studies belonging to the first category have one thing in common, namely varying magnitudes of payoffs across experimental treatments. Their focus is the effect of differences in the magnitudes of payoffs in mitigating coordination failure.

In contrast, in the second category the magnitudes of payoffs across treatments are identical. Treatments are only distinguished by the underlying environment of the coordination games. Changing the contextual factors is often a more economical way to facilitate coordination than changing the magnitudes of payoffs. Its objective is to reduce uncertainty about opponents' behavior and to facilitate better communication between players in order to develop an implicit mutual understanding and to provide assurance to players that others would likely play the payoff-dominant action.

Among studies belonging to the second category, Cooper et al. (1992) study the role of one-way and two-way communication between players in mitigating

[28] In their experiment, the coordination games are depicted as a hypothetical setting where a pair of subjects agree to meet up but cannot confirm beforehand their meeting takes place. Two alternative places are given: one is made salient by representing it as a landmark building (e.g., the Chicago Sears Tower) while the other is a small, unknown building (e.g., the AT&T building).

coordination failure. They show that two-way communication is the more effective. Van Huyck et al. (1992), Bangun et al. (2006), and Chaudhuri et al. (2009) study the role of non-binding external advice given to players to encourage them to adopt a payoff-dominant action. They show that the presence of external advice strengthens players' belief in other players' willingness to adopt payoff-dominant action and thereby facilitate coordination. Berninghaus and Ehrhart (2001) and Brandts and Cooper (2006b) investigate the effect of information on opponents' strategy in overcoming coordination failure. They find that revealing opponents' previous decisions (either the distribution of group members' decisions or each individual group member's decision) is effective in overcoming coordination failure.

However, these decentralized methods usually require stringent execution since a slight deviation from mutually optimistic beliefs may lead to coordination failure. For instance, two-way communication (where both players send out messages) is effective while one-way (only one player sends out messages) is not (Cooper et al., 1992); arbiters' assignments are credible only when they do not violate payoff dominance and symmetry (Van Huyck et al., 1992); advice from predecessors has to be "common knowledge" as a slight deviation from it (i.e., advice that is "almost common knowledge") may not work (Chaudhuri et al., 2009).

Our study follows the line of research exploring the use of contextual factors to overcome coordination failure. Using 2x2 stag hunt coordination games, we study the effectiveness of the disclosure of information on the governance mechanism determining the payoff from every pair of possible strategies that subjects may choose. More specifically, the governance mechanisms we examine in this chapter are the centralized reward and punishment schemes. Note that the focus of this chapter is not on the use of reward or punishment themselves in facilitating coordination, but rather on the information of the underlying mechanism determining payoffs. The games played in the control and experimental treatments are identical. The payoff structures used in these treatments are exactly the same. However, in the control treatment, we only present the final payoffs from coordination, while in the other treatment we provide detailed information on how those final payoffs are derived using the centralized reward and punishment schemes. Essentially, when these details are provided, subjects can see that there is a reward scheme behind the payoffs accrued from the payoff-dominant equilibrium and a punishment scheme behind the payoffs accrued from the risk-dominant equilibrium. We also elicit subjects' beliefs about their opponent's behavior, which may shed some light on the channels through which the mechanism works.

Our study contributes to the literature in the following ways. Firstly, to the best of our knowledge, this is the first study to use information on how the payoffs from coordination are derived to facilitate coordination. This is closely related to the literature on the perception problem (Camerer et al., 2004), which suggests that subjects may view the game played in experiments differently compared to what the experimenters think they are playing. Our study delves into this discrepancy in the perception problem. The information revelation mechanism develops the shared understanding of the game that subjects play and thus facilitates coordination. It shares the spirit of communication and is an implicit way. Secondly, we show that the effect of information on players' incentive to coordinate on the payoff-dominant equilibrium differs depending on the nature of the information provided.

The main findings are as follows. We find that revealing information about institutional rules regardless of the mechanism effectively increases payoff-dominant action, thus substantially reducing coordination failure. Information about the reward mechanism helps sustain the play of the payoff-dominant strategy over time, and information about the punishment mechanism even slightly increases the play during the course of the experiment. Both types of information increase beliefs while the latter also shows direct positive effects on actions. In addition, we posit that the presence of punishment or reward may make the payoff-dominant equilibrium more salient rather than changing people's preferences.

We proceed as follows. Section 4.2 introduces the experimental design. Section 4.3 discusses experimental results, followed by a conclusion in Section 4.4.

4.2 Experimental Design

We wished to study the effectiveness of a specific type of information sharing on equilibrium selection in stag hunt coordination games. In our coordination game setting, the shared information is information on how the stag hunt game is developed from a prisoner's dilemma game by introducing punishment or reward. We focused on information about two types of institutional rules (i.e., rewarding mutual cooperation and punishing unilateral defection). We aimed to investigate whether and to what extent revealing the institutional rules helps people coordinate on the payoff-dominant equilibrium. In addition, we were interested in whether revealing information on the punishment and reward mechanisms has differential effects on equilibrium selection.

There were two types of stag hunt game in our experiment. Each was played in both experimental and control conditions. The experimental and control treatments differed only in the information revealed. In other words, games played in these two treatment conditions were strategically equivalent. We will give details of the treatment conditions in the next subsection. Before subjects made their decisions, we asked them to predict the likelihood of their opponent making a cooperative decision. Beliefs have been found to be closely related to decisions in the literature (see, e.g., Charness and Dufwenberg, 2006; Croson, 2007). It has been suggested that contextual factors affect behavior through beliefs (Dufwenberg et al., 2011; Ellingsen et al., 2012). Belief elicitation allows us to explore how information about institutional rules shapes beliefs, which in turn spells action. Note that the revealed information about institutional rules may have a hybrid quality (i.e., shaping behavior through beliefs and shaping action directly). It would be interesting to see whether information about institutional rules regarding punishment and reward functions through different channels.

In addition to beliefs, we also elicited subjects' risk attitude. It has been shown in the literature that risk preferences and decisions under strategic uncertainty are closely related. For instance, risk preferences relate to trust (Bohnet and Zeckhauser, 2004; Schechter, 2007), coordination (Heinemann et al., 2009), behavioral patterns deviating from Nash in matching pennies games (Goeree et al., 2003), and contribution in public goods games (Teyssier, 2012). We used the multiple price list method similar to the one used by Holt and Laury (2002). Subjects were presented with 10 paired choices, one of which (option A) generated a deterministic payoff and the other (option B) generated two possible payoffs with certain probabilities. Table 4.1 presents the paired choices used in this risk elicitation. A risk neutral person would switch from option A to option B at line 5. The later the switch, the more risk averse the individual. The switching point measures one's risk attitude.

Table 4.1. The ten paired lotteries

Line	Option A	Option B	Expected payoff difference
1	$1	3 of 0%, 0 of 100%	$1
2	$1	3 of 10%, 0 of 90%	$0.70
3	$1	3 of 20%, 0 of 80%	$0.40
4	$1	3 of 30%, 0 of 70%	$0.10
5	$1	3 of 40%, 0 of 60%	-$0.20
6	$1	3 of 50%, 0 of 50%	-$0.50
7	$1	3 of 60%, 0 of 40%	-$0.80
8	$1	3 of 70%, 0 of 30%	-$1.10
9	$1	3 of 80%, 0 of 20%	-$1.40
10	$1	3 of 90%, 0 of 10%	-$1.70

4.2.1. Treatments

There were four treatments altogether. Figure 4.1 presents all the games involved in our experiment.[29] Subjects in all treatments played either stag hunt game 1 or stag hunt game 2. In the two experimental treatments, subjects were given information on how the stag hunt game is developed from the prisoner's dilemma game by introducing punishment or reward. The other two treatments served as baselines where subjects played stag hunt games without any information on the transformation process. Comparison between the experimental and baseline treatments sheds light on how revealed information

[29] In the instructions, we refer to players as "ROW" or "COLUMN" player. Their strategies "C" and "D" are referred to "Up" and "Down" for the row player, and "Left" and "Right" for the column player. We use "C" and "D" hereafter for convenience.

affects equilibrium selection. In what follows, we explain the treatment conditions in detail.

Prisoner's dilemma game

	C	D
C	3, 3	0, 4
D	4, 0	1.5, 1.5

Stag hunt game 1

	C	D
C	3, 3	0, 2
D	2, 0	1.5, 1.5

Stag hunt game 2

	C	D
C	5, 5	0, 4
D	4, 0	1.5, 1.5

Figure 4.1. The games in our experiment

The punish_stag1 treatment

In this treatment, subjects played the stag hunt game 1 shown in Figure 4.1 and were informed of the following transformation process. The basic game is the prisoner's dilemma game (pd game) shown in Figure 4.1. Then a S$2 punishment for the unilateral defector is introduced. The implementation of a punishment scheme transforms the pd game into stag hunt game 1. The revealed information on punishment suggests that defection is a discouraged behavior. Compared to the original pd game, defection becomes a less attractive strategy, too. Subjects were told that the final game played was the stag hunt game 1.

The reward_stag2 treatment

Subjects played the stag hunt game 2 indicated in Figure 4.1. Again, the transformation process was revealed to the subjects. The starting basic game is the same prisoner's dilemma game as that in the punish_stag1 treatment. A S$2 reward for mutual cooperation is introduced, which transforms the game into stag hunt game 2. The revelation of the reward mechanism makes cooperation a more attractive strategy. It also suggests that cooperation is encouraged by the central planner. It was underlined that the final game played was the stag hunt game 2.

The baseline treatments (stag1 & stag2)

Since subjects in the two experimental treatments above played stag hunt games with different payoffs, we conducted two control treatments. In the two control treatments, subjects simply played stag hunt game 1 and 2 without any information about institutional rules, which serve as the baseline for punishment_stag1 treatment and reward_stag2 treatment. These two baseline treatments are referred as the stag1 treatment and the stag2 treatment.

4.2.2. The Procedures

The experiment was conducted at Nanyang Technological University and was programmed by z-Tree (Fischbacher 2007). We conducted two sessions each for the punishment_stag1 and the reward_stag2 treatments and four sessions each for the stag1 and the stag2 treatments.[30] The number of subjects in each session ranged from 20 to 26. In total, 292 subjects participated in the experiment. Each session lasted around 70 minutes on average. The average earnings were about S$20 (roughly US$16), including a S$2 show-up fee. All subjects were recruited through a university-level email system. They came from various academic backgrounds, including science, engineering, social science, and business. We had a between-subject design so that each subject only participated in one session. No one had participated in a similar experiment before.

The experiment consisted of two stages. The first was the main game stage, followed by risk preference elicitation in the second. We provided hard-copy instructions on paper as well as on screen. The paper instructions were read

[30]We initially had two sessions for each baseline treatment. There was some divergence in the results for the two sessions so we ran an additional two sessions for each control to enlarge the sample size.

aloud by an experimenter before the experiment started.[31] All questions were answered in private. Subjects played two practice periods and then proceeded to play 25 real periods. Each player's role (row or column player) was randomly drawn every period. The pair composition was reshuffled from period to period, too.

Before subjects made their decision in every period, they were asked to make a prediction about their opponent's propensity to cooperate. The prediction was incentivized to elicit true beliefs. One out of 25 predictions was randomly selected as the payment foundation for belief elicitation. If the prediction fell into the correct range, an extra S$4 would be added to payment. No feedback on beliefs was given until subjects finished the 25 periods of play. However, subjects were informed of their opponent's decision at the end of each period. Among the 25 real periods, five were randomly selected as the payment for the decision part. After playing the game for 25 periods, subjects entered into the risk preference elicitation stage. Their choice in one out of ten lines was randomly selected as the payment for the second stage. This completed the experiment. Subjects were asked to complete a questionnaire regarding demographics after the experiment.

4.3 Experimental Results

In this section, we start with a descriptive summary of the experimental results, followed by some regression analyses. As a robustness check and also an extension of the current study, we also briefly present results from three additional treatments regarding cooperation in the prisoner's dilemma games.

4.3.1. Data Summary

Figure 4.2 presents the mean cooperation rate over time in all treatments. Note that games played in the punish_stag1 and stag1 treatments, and the reward_stag2 and stag2 treatments were strategically equivalent, respectively. The only difference is that subjects in the former treatment were informed of the original prisoner's dilemma game and the transformation process involving punishment and reward. The starting cooperation rates in all treatments are

[31]The instructions used in the experiment can be found in the appendix.

almost identical and they remain relatively close in the first five periods. Cooperation rates start to diverge after period 5.

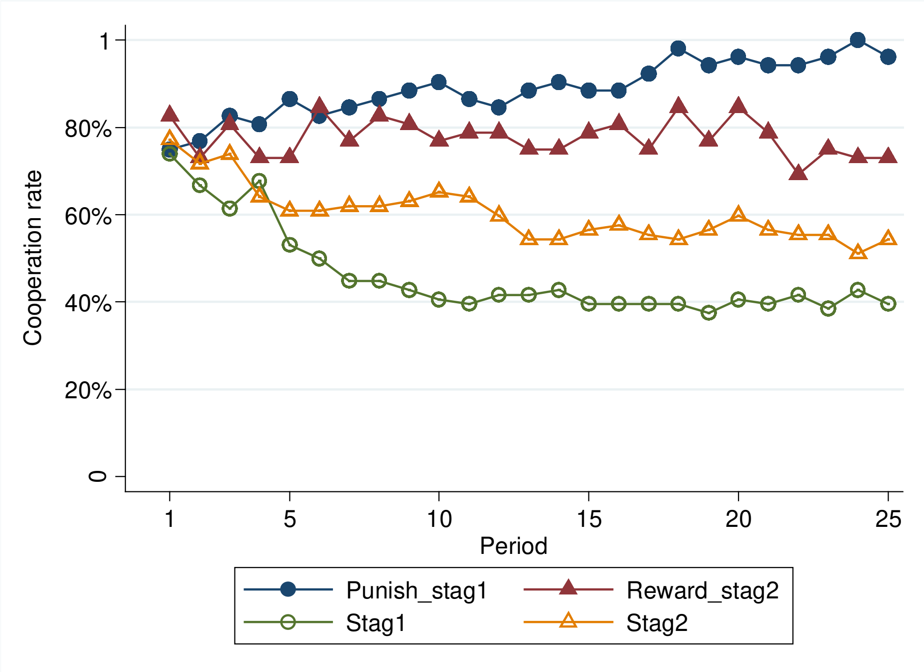


Figure 4.2. Mean cooperation rate over time in the coordination games

Cooperation rates in the punish_stag1 and reward_stag2 treatments where institutional rules were revealed were generally higher than in the baseline treatments. The differences between punish_stag1 and stag1, and reward_stag2 and stag2 are both statistically significant (two-sided Mann-Whitney test, punish_stag1 vs stag1, p-value < 0.01; reward_stag2 vs stag2, p-value = 0.03).[32] The difference between punish_stag1 and stag1 is larger, which suggests that revealing information about the punishment system might work better in terms of promoting cooperation.

There seem to be different evolutionary patterns over time across treatments. It appears that the cooperation rate decays over time in the two baseline treatments. It increases slightly over time and reaches almost full cooperation at the end in the punish_stag1 treatment. The cooperation rate remains relatively stable over time in the reward_stag2 treatment. The observation trends are verified by non-parametric tests. Using the average cooperation rate in the first five periods and the last five periods as units of observation, we find that the

[32]We used subject averages across periods as units of observation.

cooperation rate in the late periods is significantly higher in the punish_stag1 treatment (two-sided Wilcoxon matched-pairs signed-ranks test, p-value < 0.01), significantly lower in the two baseline treatments (p-value < 0.01 in both the stag1 and stag2 treatments), and not significantly different from that in early periods in the reward_stag2 treatment (p-value = 0.72).

Beliefs are found to be closely related to cooperation (Spearman rank correlation tests, p-value < 0.01 in all treatments). The distribution of average beliefs is similar to the distribution of cooperation rates presented in Figure 4.2. Likewise, we find significant differences of belief between punish_stag1 and stag1, and reward_stag2 and stag2. The evolutionary patterns of beliefs are consistent with the trend of cooperation rates, too.

Figure 4.2 indicates that people have a higher tendency to cooperate if institutional rules are revealed to them. Figure 4.3 delineates the extent to which such information sharing helps solve the coordination problem. It shows the distribution of mutual cooperation (the payoff-dominant equilibrium), mutual defection (the risk-dominant equilibrium), and disequilibrium outcomes by treatment.

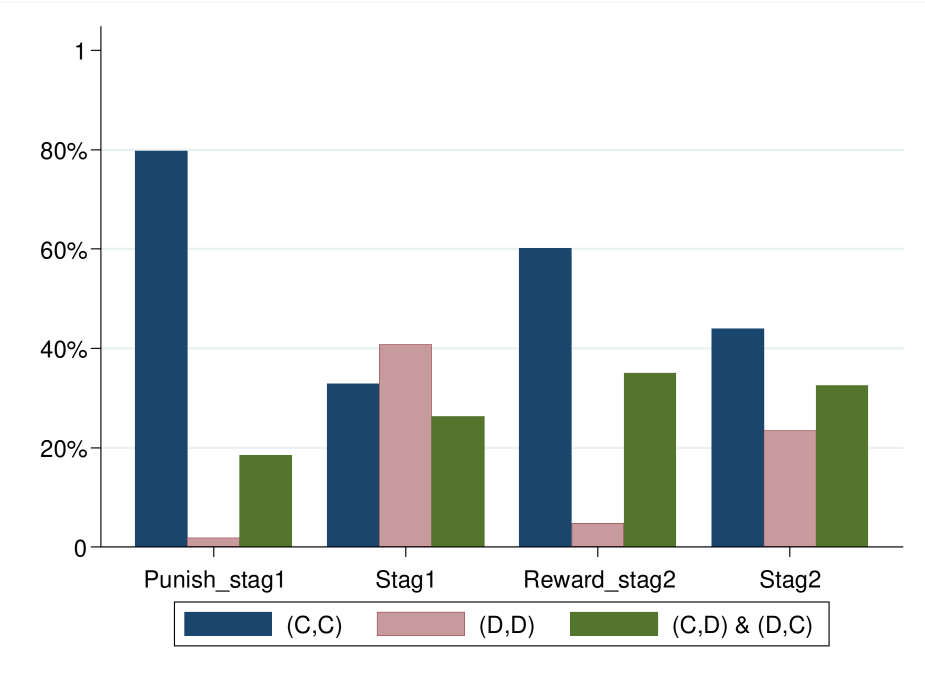


Figure 4.3. Distributions of decision pair type in the coordination games

Compared with the control stag1 treatment, there is a substantial decrease in mutual defection and a massive increase in mutual cooperation in the punish_stag1 treatment. In addition, disequilibrium outcomes decrease. Revealing the punishment mechanism effectively solves the coordination problem. People settle with the payoff-dominant equilibrium much more frequently than the risk-dominant equilibrium and end up with fewer disequilibrium outcomes. Information about the reward mechanism helps coordinate to some degree. Likewise, there is more mutual cooperation and less mutual defection in the reward_stag2 treatment than in the stag2 treatment. However, the improvement in coordination is of a smaller magnitude than that in the punish_stag1 treatment. In contrast to the decreased disequilibrium outcomes in the punish_stag1 treatment, disequilibrium outcomes slightly increase in the reward_stag2 treatment. This suggests that sharing information on mechanisms involving punishment and reward has different effects and may possibly work in different ways. We will explore this issue in more detail in later sections.

4.3.2. Regression Analyses

With the aid of several econometric models, we explore the formation of beliefs and the decision to cooperate.

Table 4.2 presents OLS estimates of the determinants of subjects' beliefs about their opponent's decision. This belief is expressed as the likelihood that the opponent will cooperate. Regressors include Period ($Period$), the subject's belief in the previous period (*Belief (t-1)*), the opponent's decision in the last three periods (*Others' Decision (t-1), Others' Decision (t-2), Others' Decision (t-3)*), [33] and the treatment dummy for the punish_stag1 treatment (*Punish_stag1*), the treatment dummy for the reward_stag2 treatment (*Reward_stag2*).

The belief formation process follows that used in Fischbacher and Gächter (2010), who find that in the context of public goods games, a subject's belief in

[33] As we used a random matching protocol, the opponent is likely to be different in each of these three periods. The regressor refers to the decision of the opponent in that a particular period.

period t is a weighted average of her belief in period $t-1$ and other group members' behavior in period $t-1$. In contrast with their findings that other group members' behavior in earlier periods has no significant effects on the belief formation in the current period, these variables do show significant effects in our study and thus three lags are included in the model.

Table 4.2. Belief formation in the coordination games

Dependent variable: Belief about the opponent's decision		
	Punish_stag1 vs Stag1	Reward_stag2 vs Stag2
Period	0.000	0.000
	(0.000)	(0.000)
Belief (t-1)	0.732***	0.813***
	(0.023)	(0.024)
Others' Decision (t-1)	0.101***	0.070***
	(0.008)	(0.007)
Others' Decision (t-2)	0.030***	0.013**
	(0.007)	(0.006)
Others' Decision (t-3)	0.031***	0.008
	(0.006)	(0.006)
Punish_stag1	0.018***	
	(0.006)	
Reward_stag2		0.012**
		(0.006)
Constant	0.047***	0.048***
	(0.009)	(0.011)
Observations	3,256	3,168
R^2	0.845	0.774

Note: OLS regressions with clustering on individuals. Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

The left-hand panel estimates the belief formation in the stag hunt game 1, including the punish_stag1 and stag1 treatments. The right-hand panel estimates the belief formation process in the other stag hunt game. Belief in the previous period always has significantly positive effects on belief in the current period. The coefficient is of a substantial magnitude in both panels and is the main factor affecting belief in the current period. The opponent's decisions in the last three periods are all significantly positive in the left-hand panel. The third lag becomes insignificant in the right-hand panel. In the both panels, the significance of the opponent's decision decreases substantially after the first lag. The treatment dummies for punish_stag1 and reward_stag2 are both positively significant. That is to say, controlling for other variables, revealing information on the punishment or reward mechanism increases subjects' belief in their opponent's propensity to cooperate.[34]

So far we have shown that revealing institutional rules increases subjects' beliefs in their opponent's cooperative behavior. In what follows, we explore the determinants of decisions. Table 4.3 shows probit estimates of the determinants of the cooperative decision. The dependent variable is a binary variable taking the value of 1 if the subject decides to cooperate and 0 otherwise. Explanatory variables are the subject's belief about their opponent's propensity to cooperate (*Belief*),[35] Period (*Period*), the number of safe options taken in the lottery (*No. of safe options*), and treatment dummies for the punish_stag1 and reward_stag2 treatments (*punish_stag1, reward_stag2*). The table also reports the marginal probability change at the sample mean of regressors while for binary variables, the marginal effects report the probability change when the indicator variable changes from 0 to 1.[36]

One's belief apparently carries a lot of weight in decision making. The coefficient is always positively significant and of a substantial magnitude. One is much more likely to cooperate if one believes one's opponent will do so too.

---

[34]Since a lagged dependent variable is used as a regressor, we also tried the "difference" and "system" generalized method-of-moments (GMM) dynamic panel estimation method for belief formation in individual treatments (Roodman, 2009). However, the long panel T and relatively small N lead to an explosive number of instruments, which may generate bias in estimates as indicated by a perfect Hansen statistic of 1.000.

[35]One may worry about the endogeneity of beliefs. On the one hand, it is not uncommon for belief to be used as a regressor in the literature (see, e.g., Charness and Dufwenberg, 2006; Croson, 2007; Fischbacher and Gächter, 2010; Dufwenberg et al., 2011); on the other, we applied the two-stage least squares estimation method, treating belief as an endogenous variable. Our conclusion remains the same.

[36]As for belief formation and cooperation, we also applied the random effects model. Since the estimation results are very similar to OLS, only OLS results are reported.

It has been shown earlier that revealing institutional rules helps increase subjects' belief in others' propensity to cooperate. This increased belief then leads to more cooperation. The effect of beliefs on behavior seems to be universal for both mechanisms. The risk attitude has mixed effects depending on the payoff structure of the coordination game. It has no effect in stag hunt game 1, where the difference in the cooperator's monetary payoff between the payoff-dominant equilibrium and disequilibrium is relatively small ($3\ vs\ 0$). The effect is marginally significant if the difference increases ($5\ vs\ 0$). The negative sign in the right-hand panel suggests that the more risk-averse a person is, the less likely he/she is to cooperate in stag hunt game 2. It might be the case that the more risk-averse person is, the less willing he/she is to take risks under strategic uncertainty if the cost of being a "sucker" is relatively high (i.e., the "sucker" gets 0 unless coordination is successful, in which case, he/she receives S$5). The treatment dummy for punish_stag1 is positively significant. This suggests that, controlling for other factors, revealing the punishment mechanism increases the likelihood of cooperation by 17 percentage. However, we do not find similar effects regarding the reward mechanism, as indicated by the insignificant coefficient in the right-hand panel.

In summary, we find that revealing institutional rules helps increase subjects' belief in their opponent's propensity to cooperate, which improves their own cooperation. This channel is universal for both types of information (i.e., information on both punishment and reward mechanisms). In addition to affecting beliefs, revealing the punishment mechanism directly improves cooperation too. It seems to have a hybrid quality (i.e., affecting behavior both through beliefs and directly). However, we do not find a similar hybrid quality for information about the reward mechanism. There seems to be a stronger effect of information on the punishment mechanism compared to that on the reward mechanism. It is in line with findings in voluntary contribution games that punishment works better than reward in terms of promoting cooperation (Sefton et al., 2007; Sigmund et al., 2001).

Table 4.3. Cooperation in the coordination games

Dependent variable: Cooperation decision = 1				
	Punish_stag1 vs Stag1		Reward_stag2 vs Stag2	
	Probit	Marginal effects	Probit	Marginal effects
Belief	5.291***	1.922***	4.100***	1.384***
	(0.276)	(0.113)	(0.423)	(0.153)
Period	-0.003	-0.001	-0.007	-0.002
	(0.005)	(0.002)	(0.004)	(0.001)
No. of Safe Options	-0.026	-0.010	-0.075*	-0.025*
	(0.039)	(0.014)	(0.042)	(0.014)
Punish_stag1	0.485***	0.169***		
	(0.180)	(0.059)		
Reward_stag2			0.163	0.054
			(0.189)	(0.062)
Constant	-2.716***		-1.597***	
	(0.234)		(0.391)	
Observations	3,700	3,700	3,600	3,600

Note: Probit regressions with clustering on individuals. Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

4.3.3. Institutional Rules in Prisoner's Dilemma Games

Thus far, we have analyzed the effects of revealing institutional rules on cooperation in stag hunt coordination games. It is noteworthy that cooperation is an equilibrium strategy. When cooperation increases, we are not sure whether it is because the revealed institutional rules make the mutual cooperation equilibrium more salient or because the revealed punishment or reward mechanism changes people's preferences. People might perceive the institutional rules as a signal from the central authority to encourage cooperation, making them more willing to cooperate.

To shed light on the issue, we ran three additional treatments using prisoner's dilemma games. In the current study, we have focused on stag hunt coordination games that were developed from prisoner's dilemma games by introducing punishment or reward. Coordination games were our baselines. To isolate equilibrium saliency from changed preferences, we employed a set of prisoner's dilemma games where equilibrium saliency is absent as cooperation is no longer an equilibrium strategy. We used the prisoner's dilemma game in Figure 4.1 as the baseline in the added treatments, and implemented a punishment or reward mechanism. We chose the punishment or reward amount so that the game remained a prisoner's dilemma game. Since cooperation is not an equilibrium strategy, more cooperation, if any, may only be the result of changed preferences due to the punishment or reward mechanism. Mechanisms in prisoner's dilemma games were supposed to be more powerful in shaping preferences than those used in stag hunt games. Compared with the baseline, not only was the signal delivered that cooperation was encouraged and that defection was discouraged, but the actual payoff also changed because of the use of punishment and reward. However, information about punishment and reward was revealed by signal delivery and the payoff remained the same as the baseline in the stag hunt games.

We had two sessions for each treatment. The number of subjects in each session was either 24 or 26. The experimental procedure was similar to other treatments. In the baseline treatment (*pd treatment*), subjects played the prisoner's dilemma game shown in Figure 4.1. In the treatment with punishment (*punish_pd treatment*), subjects were shown the pd game and were told that there was a S$0.50 punishment for unilateral defection. In the treatment with reward (*reward_pd*), subjects were informed of the pd game and a S$0.50 reward for mutual cooperation. [37] The payoff structure after punishment or reward was also displayed. Note that the game remained a prisoner's dilemma after punishment or reward was implemented.

Figure 4.4 delineates the average cooperation rate over time in the three treatments. The typical decaying trend over time appears in all three. We do not see much difference between the experimental and baseline treatments. Using subjects' average cooperate rates over time as units of observation, none of the differences is statistically significant (two-sided Mann-Whitney test, pd vs

[37] The punishment and reward amounts were set to 0.5 because we didn't want to transfer the prisoner's dilemma game to coordination names.

punish_pd, p-value = 0.85; pd vs reward_pd, p-value = 0.11). The use of punishment and reward does not seem to increase cooperation.

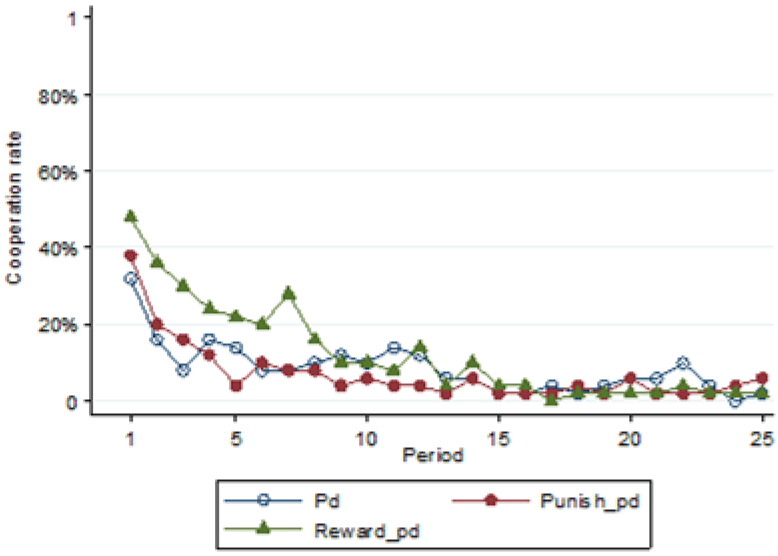


Figure 4.4. Mean cooperation rate over time in the prisoner's dilemma

Figure 4.5 shows the distribution of types of decision pairs. The unique Nash equilibrium (i.e., mutual defection) is clearly the dominant type in all treatments. The proportion of the socially efficient outcome (i.e., mutual cooperation) is close to zero. We do not observe much difference between treatments. The use of punishment or reward in the prisoner's dilemma game does not make people cooperate more even though they are encouraged to do so. Therefore, preferences do not seem to change in the presence of punishment or reward.

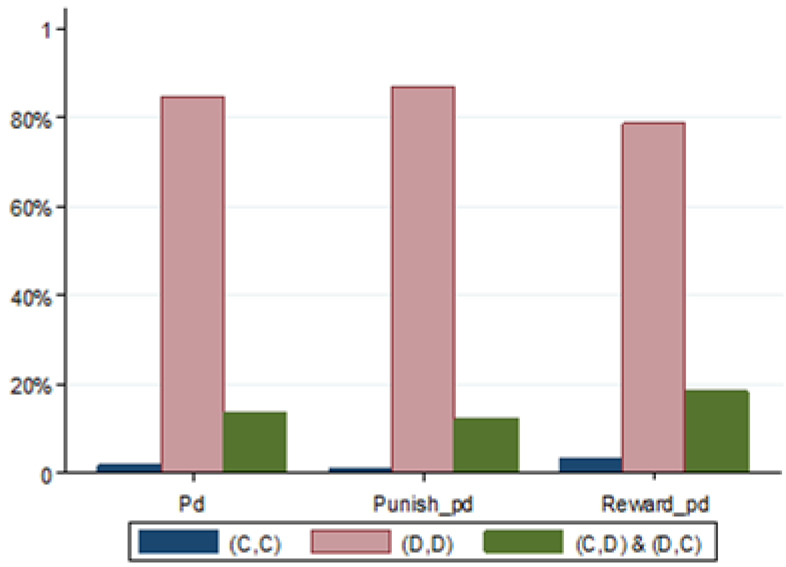


Figure 4.5. Type of decision pair in the prisoner's dilemma games

4.4 Conclusion

The goal of this chapter was to study the role of information on institutional rules regarding the underlying reward and punishment mechanisms for equilibrium selection in stag hunt games. We had two experimental treatments with full information and two control treatments. In the full information treatment, subjects were informed how the prisoner's dilemma game was transformed into a stag hunt game by introducing reward or punishment. In the control treatment, this information was absent. We elicited subjects' beliefs about their opponent's behavior before each round of play. We were also interested to know how the revealed information shapes subjects' behavior by investigating the dynamics of their beliefs and decisions. To find out whether the presence of reward and punishment changes preferences, we added three additional treatments using the prisoner's dilemma game, where cooperation is not an equilibrium strategy.

Our results indicate that sharing information on institutional rules is effective in inducing cooperation. The occurrence of coordination failure is substantially reduced. Revealing information about the reward and punishment mechanisms increased subjects' belief in their opponent's propensity to cooperate, which in turn spells action. Besides working through beliefs, information about punishment mechanisms has direct positive effects on decisions. We do not find similar direct effects on action for information about reward mechanisms. Moreover, the results from the prisoner's dilemma games suggest that the use of reward and punishment does not make people more willing to cooperate when cooperation is not an equilibrium strategy. Thus, we posit that the revelation of information about the reward and punishment mechanism makes the payoff-dominant equilibrium more salient rather than changing people's preferences.

4.5 Appendix: Experimental instructions

A. Section 1

A.1 The punish_stag1 treatment

General Information

Welcome to all of you! You are now taking part in an interactive study on decision making. **Please pay attention to the information provided here and make your decisions carefully. If at any time you have questions to ask, please raise your hand and we will attend to you in private.**

Please note that **unauthorized communication is prohibited**. Failure to adhere to this rule would force us to stop this study and you may be held liable for the cost incurred in this simulation.

Your participation in this study is voluntary. You will receive **2 S$** show-up fee for participating in this study. You may decide to leave the study at any time. Unfortunately, if you withdraw before you complete the study, we can only pay you for the decisions that you have made up to the time of withdrawal, which could be substantially less than you will earn if you complete the entire study.

The amount of your earnings from this study depends on the decisions you and others make. At the end of this session, your earnings will be paid to you privately and in cash. It would be contained in an envelope (indicated with your unique user ID). You will need to sign a claim card given to you and exchange your claim card with your payment.

General Instructions

Each of you will be given **a unique user ID** and it **will be clearly stated on your computer screen**. At the end of the study, you will be asked to fill in your user ID and other information pertaining to your earnings from this study in the claim card. **Please fill in the correct user ID to make sure that you will get the correct amount of payment**.

Rest assured that your **anonymity will be preserved** throughout the study. You will **never be aware of** the personal identities of other participants **during or after** the study. Similarly, other participants will also **never be aware** of your personal identities **during or after** the study. You will only be identified by your user ID in our data collection. All information collected will **strictly be kept confidential** for the sole purpose of this study.

Specific Instructions

The total duration of this study is approximately 1 hour. You will have to go through two sections.

Your total payment = earnings from decisions you made in section 1
+ earnings from predictions you made in section 1
+ earnings from choices you made in section 2
+ show up fee

Section 1

In section 1, you will form a pair with another participant and interact anonymously through computer interface for several periods. Each participant will only be identified by his (her) unique and randomly generated numerical ID displayed on his (her) computer screen during the experiment. The person with whom you are matched will be randomly re-drawn after every period. You are paired anonymously, which means that you will never learn the identity of the other person in any of the periods. Similarly, no one will know your identification in any period.

One of you will be assigned as player ROW and the other person will be assigned as player COLUMN. Your role will also be randomly re-drawn in each period, so that sometimes you will be player ROW and sometimes you will be player COLUMN.

Here is the basic game:

		COLUMN Left	COLUMN Right
ROW	Up	3, 3	0, 4
	Down	4, 0	1.5, 1.5

Player ROW and Player COLUMN make choices independently and simultaneously. Player ROW chooses Up or Down; player COLUMN chooses Left or Right.

The **1st number in each cell** refers to the payoff (in SGD) for **Player ROW**, while the **2nd number in each cell** refers to the payoff (in SGD) for **Player COLUMN**. Thus, for example, if player ROW chooses Up and player COLUMN chooses Left, player ROW would receive 3 dollars and player COLUMN would receive 3 dollars.

Now, we introduce punishment to this game. We'll **punish player COLUMN** by deducting 2 dollars from his/her payoff **only if** player COLUMN chooses **Right** when player ROW chooses **Up**, namely, no punishment on player COLUMN for choosing Right if player ROW chooses Down. Similarly, we'll **punish player ROW** by deducting 2 dollars from his/her payoff **only if** player ROW chooses **Down** when player COLUMN chooses **Left**. Thus, for example, if player ROW chooses Down and player COLUMN chooses Left, player ROW would receive 2 dollars (4 -2 = 2) and player COLUMN would receive 0. With punishment, the **final game payoff** becomes:

		COLUMN	
		Left	Right
ROW	Up	3, 3	0, 2
	Down	2, 0	1.5, 1.5

However, you will be asked to **make a prediction** about the actions of other participants who could potentially be your partner **prior to** making your own decision. If your prediction falls within the **correct range** (that is, **it diverges from the actual outcome by +/-2**), you will receive **S$ 4 (four) on top of your earnings from this part**.

You will then learn your payoff and the strategy taken by the person with whom you are paired at the end of each period. This completes one period of play. You'll do several periods, out of all decisions you made in real periods, 5 of them will be randomly chosen to determine your payoff from this part. In addition, one of all predictions you made in real periods will be randomly chosen to determine your extra earning obtained from the prediction part in section 1. **That's to say, every period could possibly decide your payment, so do make your prediction and decision carefully!** You'll see your total payment at the end of the study. We'll pay you individually and privately in cash at the end of the experiment.

Things to Remember: Please click on the OK button as soon as you have read and understood the information on screen because the system will wait until everyone has clicked OK before it proceeds. You are playing for real money.

This is the end of the instructions for section 1; you will be given the instruction for the next section once you have completed section 1.

Are there any questions? Please feel free to ask by raising your hand. We'll attend to you privately.

===***===

A.2 The reward_stag2 treatment

General Information

Welcome to all of you! You are now taking part in an interactive study on decision making. **Please pay attention to the information provided here and make your decisions carefully. If at any time you have questions to ask, please raise your hand and we will attend to you in private.**

Please note that **unauthorized communication is prohibited**. Failure to adhere to this rule would force us to stop this study and you may be held liable for the cost incurred in this simulation.

Your participation in this study is voluntary. You will receive **2 S$** show-up fee for participating in this study. You may decide to leave the study at any time. Unfortunately, if you withdraw before you complete the study, we can only pay you for the decisions that you have made up to the time of withdrawal, which could be substantially less than you will earn if you complete the entire study.

The amount of your earnings from this study depends on the decisions you and others make. At the end of this session, your earnings will be paid to you privately and in cash. It would be contained in an envelope (indicated with your unique user ID). You will need to sign a claim card given to you and exchange your claim card with your payment.

General Instructions

Each of you will be given **a unique user ID** and it **will be clearly stated on your computer screen**. At the end of the study, you will be asked to fill in your user ID and other information pertaining to your earnings from this study in the claim card. **Please fill in the correct user ID to make sure that you will get the correct amount of payment**.

Rest assured that your **anonymity will be preserved** throughout the study. You will **never be aware of** the personal identities of other participants **during or after** the study. Similarly, other participants will also **never be aware** of your personal identities **during or after** the study. You will only be identified by your user ID in our data collection. All information collected will **strictly be**

kept confidential for the sole purpose of this study.

Specific Instructions

The total duration of this study is approximately 1 hour. You will have to go through two sections.

Your total payment = earnings from decisions you made in section 1
+ earnings from predictions you made in section 1
+ earnings from choices you made in section 2
+ show up fee

Section 1

In section 1, you will form a pair with another participant and interact anonymously through computer interface for several periods. Each participant will only be identified by his (her) unique and randomly generated numerical ID displayed on his (her) computer screen during the experiment. The person with whom you are matched will be randomly re-drawn after every period. You are paired anonymously, which means that you will never learn the identity of the other person in any of the periods. Similarly, no one will know your identification in any period.

One of you will be assigned as player ROW and the other person will be assigned as player COLUMN. Your role will also be randomly re-drawn in each period, so that sometimes you will be player ROW and sometimes you will be player COLUMN.

Here is the basic game:

		COLUMN	
		Left	Right
ROW	Up	3, 3	0, 4
	Down	4, 0	1.5, 1.5

Player ROW and Player COLUMN make choices independently and simultaneously. Player ROW chooses Up or Down; player COLUMN chooses Left or Right.

The **1st number in each cell** refers to the payoff (in SGD) for **Player ROW**, while the **2nd number in each cell** refers to the payoff (in SGD) for **Player COLUMN**. Thus, for example, if player ROW chooses Up and player COLUMN chooses Left, player ROW would receive 3 dollars and player COLUMN would receive 3 dollars.

Now, we introduce reward to this game. We'll **reward both players** by adding 2 dollars to each player's payoff only **if** player COLUMN chooses **Left** and player ROW chooses **Up** simultaneously. There are no rewards in other choice combinations. Thus, for example, if player ROW chooses Up and player COLUMN chooses Left, player ROW would receive 5 dollars (3 +2 = 5) and player COLUMN would receive 5 dollars (3 +2 = 5). With reward, the **final game payoff** becomes:

		COLUMN Left	Right
ROW	Up	5, 5	0, 4
	Down	4, 0	1.5, 1.5

However, you will be asked to **make a prediction** about the actions of other participants who could potentially be your partner **prior to** making your own decision. If your prediction falls within the **correct range** (that is, **it diverges from the actual outcome by +/-2**), you will receive **S$ 4 (four) on top of your earnings from this part**.

You will then learn your payoff and the strategy taken by the person with whom you are paired at the end of each period. This completes one period of play. You'll do several periods, out of all decisions you made in real periods, 5 of them will be randomly chosen to determine your payoff from this part. In addition, one of all predictions you made in real periods will be randomly chosen to determine your extra earning obtained from the prediction part in section 1. **That's to say, every period could possibly decide your payment, so do make your prediction and decision carefully!** You'll see your total payment at the end of the study. We'll pay you individually and privately in cash at the end of the experiment.

Things to Remember: Please click on the OK button as soon as you have read and understood the information on screen because the system will wait until everyone has clicked OK before it proceeds. You are playing for real money.

This is the end of the instructions for section 1; you will be given the instruction for the next section once you have completed section 1.

Are there any questions? Please feel free to ask by raising your hand. We'll attend to you privately.

===***===

B. Section 2

Section 2

In this part of the experiment you will be asked to make a series of choices. How much you receive will depend partly on **chance** and partly on the **choices** you make. The decision problems are not designed to test you. What we want to know is what choices you would make in them. The only right answer is what you really would choose.

For each line in the table in the next page, please state whether you prefer option A or option B. Notice that there are a total of **10 lines** in the table but just one line will be randomly selected for payment. **You do not know which line will be paid when you make your choices. Hence you should pay attention to the choice you make in every line.** After you have completed all your choices, the computer will randomly generate a number, which determines which line is going to be paid.

Your earnings for the selected line depend on which option you chose: If you chose option A in that line, you will receive **$1**. If you chose option B in that line, you will receive either **$3** or **$0**. To determine your earnings in the case you chose option B there will be second random draw. The computer will randomly determine if your payoff is 0 or $3, with the chances stated in Option B.

You earnings from section 2 will be revealed at the end of the study after you have completed a short questionnaire that will be shown to you on your computer screen.

Are there any questions? Please feel free to ask by raising your hand. We'll attend to you privately

===END===

Bibliography

Adair, J. G., T. W. Dushenko, and R. Lindsay (1985). Ethical regulations and their impact on research practice. *American Psychologist 40*(1), 59.

Aghion, P. and P. Bolton (1992). An incomplete contracts approach to financial contracting. *Review of Economic Studies 59*(3), 473–494.

Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters 80*(1), 123–129.

Almenberg, J., A. Dreber, C. L. Apicella, and D. G. Rand (2011). *Third party reward and punishment: group size, efficiency and public goods, in "Psychology of Punishment"*. Nova Science Publishers.

Altmann, S. and A. Falk (2009). The impact of cooperation defaults on voluntary contributions to public goods. Technical report, Mimeo.

Andreoni, J. (1995). Cooperation in public-goods experiments: kindness or confusion? *American Economic Review 85*(4), 891–904.

Bangun, L., A. Chaudhuri, P. Prak, and C. Zhou (2006). Common and almost common knowledge of credible assignments in a coordination game. *Economics Bulletin 3*(1), 1–10.

Battalio, R., L. Samuelson, and J. Van Huyck (2001). Optimization incentives and coordination failure in laboratory stag hunt games. *Econometrica 69*(3), 749–764.

Baumrind, D. (1985). Research using intentional deception: Ethical issues revisited. *American Psychologist 40*(2), 165.

Bennardo, A., M. Pagano, and S. Piccolo (2009). Multiple-bank lending, creditor rights and information sharing. *CEPR Discussion Paper No. DP7186*.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity, and social history. *Games and Economic Behavior 10*(1), 122–142.

Berninghaus, S. K. and K.-M. Ehrhart (2001). Coordination and information: Recent experimental evidence. *Economics Letters 73*(3), 345–351.

Beshears, J., J. J. Choi, D. Laibson, and B. C. Madrian (2009). The importance of default options for retirement saving outcomes: Evidence from the united states. In *Social Security Policy in a Changing Environment*, pp. 167–195. University of Chicago Press.

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes 63*(2), 131–144.

Bohnet, I. and R. Zeckhauser (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization 55*(4), 467–484.

Bolton, G. E. and A. Ockenfels (2000). Erc: A theory of equity, reciprocity, and competition. *American economic review 90*(1), 166–193.

Bolton, P. and D. S. Scharfstein (1996). Optimal debt structure and the number of creditors. *Journal of Political Economy 104*(1), 1–25.

Brandts, J. and G. Charness (2011). The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics 14*(3), 375–398.

Brandts, J. and D. J. Cooper (2006a). A change would do you good.... an experimental study on how to overcome coordination failure in organizations. *American Economic Review 96*(3), 669–693.

Brandts, J. and D. J. Cooper (2006b). Observability and overcoming coordination failure in organizations: An experimental study. *Experimental Economics 9*(4), 407–423.

Broseta, B., E. Fatas, and T. Neugebauer (2003). Asset markets and equilibrium selection in public goods games with provision points: An experimental study. *Economic Inquiry 41*(4), 574–591.

Brown, M., T. Jappelli, and M. Pagano (2009). Information sharing and credit: Firm-level evidence from transition countries. *Journal of Financial Intermediation 18*(2), 151–172.

Brown, M. and C. Zehnder (2007). Credit reporting, relationship banking, and loan repayment. *Journal of Money, Credit and Banking 39*(8), 1883–1918.

Carpenter, J. P. and P. H. Matthews (2012). Norm enforcement: Anger, indignation, or reciprocity? *Journal of the European Economic Association 10*(3), 555–572.

Carpenter, J. P., P. H. Matthews, et al. (2004). Why punish? social reciprocity and the enforcement of prosocial norms. *Journal of Evolutionary Economics 14*(4), 407–429.

Charness, G., R. Cobo-Reyes, and N. Jiménez (2008). An investment game with third-party intervention. *Journal of Economic Behavior & Organization 68*(1), 18–28.

Charness, G. and M. Dufwenberg (2006). Promises and partnership. *Econometrica 74*(6), 1579–1601.

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

Charness, G. and M. Rabin (2005). Expressed preferences and behavior in experimental games. *Games and Economic Behavior 53*(2), 151–169.

Chaudhuri, A., A. Schotter, and B. Sopher (2009). Talking ourselves to efficiency: Coordination in inter-generational minimum effort games with private, almost common and common knowledge of advice. *Economic Journal 119*(534), 91–122.

Chavez, A. K. and C. Bicchieri (2013). Third-party sanctioning and compensation behavior: Findings from the ultimatum game. *Journal of Economic Psychology 39*, 268 – 277.

Cheng, X. and H. Degryse (2010). Information sharing and credit rationing: Evidence from the introduction of a public credit registry. *CentER Discussion Paper Series No. 2010-34S*.

Choi, J., D. Laibson, and B. Madrian (2004). Plan design and 401 (k) savings outcomes. Technical report, National Bureau of Economic Research.

Coffman, L. C. (2011). Intermediation reduces punishment (and reward). *American Economic Journal: Microeconomics 3*(4), 77–106.

Colin F. Camerer, George Loewenstein, . M. R. E. (2004). *Advances in Behavioral Economics*. Princeton University Press.

Cooper, D. and J. H. Kagel (2009). *Other regarding preferences: a selective survey of experimental results*, Volume 2. Princeton University Press.

Cooper, R., D. V. DeJong, R. Forsythe, and T. Ross (1990). Selection criteria in coordination games: Some experimental results. *American Economic Review 80*(1), 218–233.

Cooper, R., D. V. DeJong, R. Forsythe, and T. W. Ross (1992). Communication in coordination games. *Quarterly Journal of Economics 107*(2), 739–771.

Crawford, V. P., U. Gneezy, and Y. Rottenstreich (2008). The power of focal points is limited: even minute payoff asymmetry may yield large coordination failures. *American Economic Review 98*(4), 1443–1458.

Croson, R. T. (2007). Theories of commitment, altruism and reciprocity: Evidence from linear public goods games. *Economic Inquiry 45*(2), 199–216.

De Janvry, A., C. McIntosh, and E. Sadoulet (2010). The supply-and demand-side impacts of credit market information. *Journal of Development Economics 93*(2), 173–188.

Degryse, H., V. Ioannidou, and E. Von Schedvin (2012). On the non-exclusivity of loan contracts: An empirical investigation. *Working Paper*, Available at SSRN 2028933.

Devetag, G. and A. Ortmann (2007). When and why? a critical survey on coordination failure in the laboratory. *Experimental Economics 10*(3), 331–344.

Dewatripont, M. and J. Tirole (1994). A theory of debt and equity: Diversity of securities and manager-shareholder congruence. *Quarterly Journal of Economics 109*(4), 1027–1054.

Diamond, D. W. (1989). Reputation acquisition in debt markets. *Journal of Political Economy 97*(4), 828–862.

Djankov, S., C. McLiesh, and A. Shleifer (2007). Private credit in 129 countries. *Journal of Financial Economics 84*(2), 299–329.

Doblas-Madrid, A. and R. Minetti (2013). Sharing information in the credit market: Contract-level evidence from us firms. *Journal of Financial Economics 109*(1), 198?23.

Dufwenberg, M., S. Gächter, and H. Hennig-Schmidt (2011). The framing of games and the psychology of play. *Games and Economic Behavior 73*(2), 459–478.

Ellingsen, T., M. Johannesson, J. Mollerstrom, and S. Munkhammar (2012). Social framing effects: Preferences or beliefs? *Games and Economic Behavior 76*(1), 117–130.

Falkinger, J., E. Fehr, S. Gächter, and R. Winter-Ebmer (2000). A simple mechanism for the efficient provision of public goods: Experimental evidence. *American Economic Review 90*(1), 247–264.

Fehr, E. and U. Fischbacher (2004). Third-party punishment and social norms. *Evolution and Human Behavior 25*(2), 63–87.

Fehr, E. and S. Gächter (2000). Cooperation and punishment in public goods experiments. *American Economic Review 90*, 980–994.

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

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Fischbacher, U. and S. Gächter (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review 100*(1), 541–556.

Fitzgerald, M. (February 17, 2014). 'we were too far in:' citizen heroes chase down, tackle defiant crash suspect. *The Record*, Retrieved from http://www.recordnet.com/.

Fredrik, C., J. Olof, and K. Pham (2011). Funding a new bridge in rural vietnam: A field experiment on conditional cooperation and default contributions. *Working Papers in Economics 503*.

Gneezy, U. and A. Rustichini (2000). A fine is a price. *The Journal of Legal Studies 29*(1), pp. 1–17.

Goeree, J. K., C. A. Holt, and T. R. Palfrey (2003). Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior 45*(1), 97–113.

Gromb, D. (1994). Renegotiation in debt contracts. *Working Paper, INSEAD.*

Gronberg, T., A. Luccasen, and J. Van Huyck (2003). Contributions and crowd-out of public goods: Competing models and experimental evidence. *Unpublished manuscript*.

Grossman, S. J. and O. D. Hart (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy 94*(4), 691–719.

Harsanyi, J. C. and R. Selten (1988). *A general theory of equilibrium selection in games*, Volume 1. The MIT Press.

Hart, O. (2001). Financial contracting. *Journal of Economic Literature 39*(4), 1079–1100.

Hart, O. and J. Moore (1990). Property rights and the nature of the firm. *Journal of Political Economy 98*(6), 1119–1158.

Hart, O. and J. Moore (1998). Default and renegotiation: A dynamic model of debt. *Quarterly Journal of Economics 113*(1), 1–41.

Heinemann, F., R. Nagel, and P. Ockenfels (2009). Measuring strategic uncertainty in coordination games. *The Review of Economic Studies 76*(1), 181–221.

Henrich, J., J. Ensminger, R. McElreath, A. Barr, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science 327*(5972), 1480–1484.

Hey, J. D. (1998). Experimental economics and deception: A comment. *Journal of Economic Psychology 19*(3), 397–401.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review 92*(5), 1644–1655.

Hoppe, T. and A. Sadrieh (2007). An experimental assessment of confederate reserve price bids in online auction. Technical report, University Magdeburg, Faculty of Economics and Management.

Hotchkiss, E. S., K. John, R. M. Mooradian, and K. S. Thorburn (2007). Chapter 14 - bankruptcy and the resolution of financial distress. In B. E. Eckbo

(Ed.), *Handbook of Empirical Corporate Finance*, Volume 2, pp. 235 – 287. Elsevier.

Houser, D. and R. Kurzban (2002). Revisiting kindness and confusion in public goods experiments. *American Economic Review 92*(4), 1062–1069.

Houser, D., E. Xiao, K. McCabe, and V. Smith (2008). When punishment fails: Research on sanctions, intentions and non-cooperation. *Games and Economic Behavior 62*(2), 509–532.

Isaac, R. M., D. Schmidtz, and J. M. Walker (1989). The assurance problem in a laboratory market. *Public Choice 62*(3), 217–236.

Jappelli, T. and M. Pagano (2002). Information sharing, lending and defaults: Cross-country evidence. *Journal of Banking & Finance 26*(10), 2017–2045.

Jappelli, T. and M. Pagano (2006). The role and effects of credit information sharing. In G. Bertola, R. Disney, and C. B. Grant (Eds.), *The Economics of Consumer Credit*, Chapter 10, pp. 347. MIT Press.

Johnson, E., J. Hershey, J. Meszaros, and H. Kunreuther (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty 7*(1), 35–51.

Johnson, N. D. and A. A. Mislin (2011). Trust games: A meta-analysis. *Journal of Economic Psychology 32*(5), 865–889.

Kahneman, D., J. Knetsch, and R. Thaler (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives 5*(1), 193–206.

Kim, Y. (1996). Equilibrium selection in n-person coordination games. *Games and Economic Behavior 15*(2), 203–227.

Kurzban, R., P. DeScioli, and E. O'Brien (2007). Audience effects on moralistic punishment. *Evolution and Human Behavior 28*(2), 75–84.

Laury, S. K. and C. A. Holt (2008). *Voluntary provision of public goods: experimental results with interior Nash equilibria*, Volume 1. Elsevier.

Leibbrandt, A. and R. López-Pérez (2011). The dark side of altruistic third-party punishment. *Journal of Conflict Resolution 55*(5), 761–784.

Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics 1*(3), 593–622.

Li, J., E. Xiao, D. Houser, and P. R. Montague (2009). Neural responses to sanction threats in two-party economic exchange. *Proceedings of the National Academy of Sciences 106*(39), 16835–16840.

List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy 115*(3), 482–493.

Madrian, B. and D. Shea (2002). The power of suggestion: Inertia in 401 (k) participation and savings behavior: Erratum. *Quarterly Journal of Economics 117*(1), 377.

Marlowe, F. W., J. C. Berbesque, A. Barr, C. Barrett, A. Bolyanatz, J. C. Cardenas, J. Ensminger, M. Gurven, E. Gwako, J. Henrich, et al. (2008). More 'altruistic' punishment in larger societies. *Proceedings of the Royal Society B: Biological Sciences 275*(1634), 587–592.

Messer, K. D., H. Zarghamee, H. M. Kaiser, and W. D. Schulze (2007). New hope for the voluntary contributions mechanism: The effects of context. *Journal of Public Economics 91*(9), 1783–1799.

Nelissen, R. M. and M. Zeelenberg (2009). Moral emotions as determinants of third-party punishment: Anger, guilt, and the functions of altruistic sanctions. *Judgment and Decision Making 4*(7), 543–553.

Nicks, S. D., J. H. Korn, and T. Mainieri (1997). The rise and fall of deception in social psychology and personality research, 1921 to 1994. *Ethics & Behavior 7*(1), 69–77.

Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics 92*(1-2), 91–112.

Nikiforakis, N. (2010). Feedback, punishment and cooperation in public good experiments. *Games and Economic Behavior 68*(2), 689–702.

Nikiforakis, N. and H. Mitchell (2013). Mixing the carrots with the sticks: third party punishment and reward. *Experimental Economics 17*(1), 1–23.

Norton, E. C., H. Wang, and C. Ai (2004). Computing interaction effects and standard errors in logit and probit models. *Stata Journal 4*(2), 154–167.

Ortmann, A. and R. Hertwig (2002). The costs of deception: Evidence from psychology. *Experimental Economics 5*(2), 111–131.

Padilla, A. J. and M. Pagano (1997). Endogenous communication among lenders and entrepreneurial incentives. *Review of Financial Studies 10*(1), 205–236.

Padilla, A. J. and M. Pagano (2000). Sharing default information as a borrower discipline device. *European Economic Review 44*(10), 1951–1980.

Pagano, M. and T. Jappelli (1993). Information sharing in credit markets. *Journal of Finance 48*(5), 1693–1718.

Palfrey, T. R. and J. E. Prisbrey (1997). Anomalous behavior in public goods experiments: How much and why? *American Economic Review 87*(5), 829–846.

Pratchett, T. and N. Gaiman (2006). *Good Omens: The Nice and Accurate Prophecies of Agnes Nutter, Witch*. Harper Torch.

Price, M. E., L. Cosmides, and J. Tooby (2002). Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior 23*(3), 203–231.

Rankin, F. W., J. B. Van Huyck, and R. C. Battalio (2000). Strategic similarity and emergent conventions: Evidence from similar stag hunt games. *Games and Economic Behavior 32*(2), 315–337.

Rauchdobler, J., R. Sausgruber, and J.-R. Tyran (2010). Voting on thresholds for public goods: Experimental evidence. *FinanzArchiv: Public Finance Analysis 66*(1), 34–64.

Riedl, K., K. Jensen, J. Call, and M. Tomasello (2012). No third-party punishment in chimpanzees. *Proceedings of the National Academy of Sciences 109*(37), 14824–14829.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system gmm in stata. *Stata Journal 9*(1), 86.

Samuelson, W. and R. Zeckhauser (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty 1*(1), 7–59.

Sanfey, A. G., J. K. Rilling, J. A. Aronson, L. E. Nystrom, and J. D. Cohen (2003). The neural basis of economic decision-making in the ultimatum game. *Science 300*(5626), 1755–1758.

Scharlemann, J. P., C. C. Eckel, A. Kacelnik, and R. K. Wilson (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology 22*(5), 617–640.

Schechter, L. (2007). Traditional trust measurement and the risk confound: An experiment in rural paraguay. *Journal of Economic Behavior & Organization 62*(2), 272–292.

Schmidt, D., R. Shupp, J. M. Walker, and E. Ostrom (2003). Playing safe in coordination games: the roles of risk dominance, payoff dominance, and history of play. *Games and Economic Behavior 42*(2), 281–299.

Sefton, M., R. Shupp, and J. M. Walker (2007). The effect of rewards and sanctions in provision of public goods. *Economic Inquiry 45*(4), 671–690.

Sigmund, K., C. Hauert, and M. Nowak (2001). Reward and punishment. *Proceedings of the National Academy of Sciences 98*(19), 10757.

Smith, V. L. (1991). *Papers in experimental economics*. Cambridge University Press.

Spier, K. E. (1992). Incomplete contracts and signalling. *RAND Journal of Economics 23*(3), 432–443.

Teyssier, S. (2012). Inequity and risk aversion in sequential public good games. *Public Choice 151*(1-2), 91–119.

Thaler, R. and C. Sunstein (2003). Libertarian paternalism. *American Economic Review 93*(2), 175–179.

Thaler, R. and C. Sunstein (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory 21*(2), 265–293.

Twain, M. (1897). *Following the Equator, Chapter LXVI*. American Publishing Company.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review 80*(1), 234–248.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil (1991). Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *Quarterly Journal of Economics 106*(3), 885–910.

Van Huyck, J. B., A. B. Gillette, and R. C. Battalio (1992). Credible assignments in coordination games. *Games and Economic Behavior 4*(4), 606–626.

Weimann, J. (1994). Individual behaviour in a free riding experiment. *Journal of Public Economics 54*(2), 185–200.

Xiao, E. (2013). Profit-seeking punishment corrupts norm obedience. *Games and Economic Behavior 77*(1), 321–344.

Zhou, D. (January, 9, 2014). Passersby picked almost all 6000 yuan scattered on the street from a migrant worker when cleaning glasses. *fjsen.com*, Retrieved from http://www.fjsen.com/.