

Three Essays on Experimental Economics

By

Sookie Xue Zhang

THESIS

Submitted to the University of Adelaide
in partial fulfillment of the
requirement for the degree of

Doctor of Philosophy
in
Economics

July 2017

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution in my name and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signature of Author

Abstract

This thesis consists of three essays using experimental economics to empirically study human behaviors in different economic contexts. Each essay is a self-contained paper.

In the first paper, we try to address a puzzle of an unanticipated stoppage observed during houses auctions in Australia. Although no new information is revealed during the suspension, sellers perhaps intend by suspending the auction to trigger some psychological process which would lead to more aggressive bidding and therefore higher revenues. The stoppage allows bidders the time to imagine how they would live in their future home as if they were owning the house. The feeling of having the house can potentially trigger endowment effects, which generate additional attachment value to the object. In order to test this conjecture, we computerize an English auction for a real good in the laboratory with and without a stoppage. When the auction was stopped, we targeted the highest bidders by placing the object in front of them and informing them that they could keep the good if they won the auction. Unexpectedly, we observe a similar average auction price between the control treatment and the treatment with the stoppage. A deeper exploration shows that the targeted subjects won less frequently in the stop treatment than their counterparts in the control treatment. We conclude that there must be two opposite effects taking place in the stop treatment such that the same average auction price is observed as in the control treatment. A cooling-off effect makes the targeted subjects less aggressive in bidding while a heating-up effect induces the waiting subjects to bid more aggressively.

In the second paper, we study experimentally how informative cheap talk is in a delegation game where information is asymmetric and incentives are misaligned. We are particularly interested in the efficiency of delegation when we alter the cardinality of the message space. This paper contributes to the cheap talk literature by a novel delegation scenario that studies how different forms of messages affect the degree of information transmission. The one-shot

three-person delegation game is based on a repeated real-effort task. Two players can simultaneously send a costless message about their past performance along with their avatar to the delegator of their group. A delegator then can choose a player and delegate. Delegation replaces the delegator's performance in the profit function by the future performance of the chosen person. In order to misalign preferences, the delegator has to pay a fixed bonus to the person she chooses. In the baseline treatment, we adopt a structured message space which consists of integers to represent how well a player has performed in the addition task (i.e. Precise Message Treatment, PMT). Then, we introduce noise by partitioning the message space into intervals (i.e. Fuzzy Message Treatment, FMT). Lastly, we implement free text communication to allow subjects any message they want (i.e. Free Communication Treatment, FCT). In line with the lie-aversion literature, truthful reports and moderate lies are observed across all treatments. Surprisingly, information is transmitted in both the FMT and the FCT but not in the PMT. We find that on average delegators ignored messages in the PMT, but increased the frequency of delegation when they received messages indicating better performances in both the FMT and the FCT. Compared to the situation where no delegation options are allowed, the joint channel of cheap talk and delegation improve social welfare to some degree. The highest efficiency is obtained in the FMT, where players can express freely how competent they are. An important reason is that delegators are able to extract information contained in messages of different styles.

In the third paper, we investigate the social welfare enhancing effect of costly contracts used to resolve future distributional conflicts. A recent study by Bayer (2016) shows that subjects still cooperate to a certain extent in social dilemma situations, but welfare losses from competition in distributional contests destroy welfare gains from voluntary cooperation. We extend this study by providing a costly contract option before the two-stage cooperation and contest game. If a mutual agreement is made to implement the contract, the second stage distributional contest is avoided. As the baseline treatment, we adopt a simple equal split sharing rule and calibrate the contract cost to be the average effort incurred in Bayer's

contest game. Interestingly, we find that the costly equal split contract can stabilize individual contributions among those who opt in. Moreover, we find a significant improvement in the average profit compared to the control treatment where no contract option exists. However, the frequency of contracting declines dramatically in early periods. We further vary the type and the cost of a contract in two dimensions. Along the first dimension, we change the sharing rule to a proportional split conditional on individual contributions. This removes the social dilemma dimension of the cooperation game and theoretically allows for the implementation of the first-best. As expected, the majority of subjects opted for the contract with full cooperation following in most cases. Along the second dimension, we decrease the cost of an equal split contract. The lower contracting cost helps to delay the decline of the average contracting frequency. It seems that an equal split contract selects subjects who are more cooperative into signing the contract, which increases average welfare.

Acknowledgments

My PhD journey feels like a confusing long psychological experiment that I have ever expected. Lucky for me, I have met the best life mentor to guide me through this daunting journey and encountered many interesting people who make lonely research activities more fun. I am foremost indebted to my principal supervisor Ralph-Christopher Bayer for his excellent guidance and generous support. Without his feedback and encouragement through numerous inspiring discussions over the past four (plus) years this thesis would not have been possible. Thank you for always being there for me every time I am about to panic. Also, I appreciate delightful conversations and comments on different parts of the thesis from the seminar speakers who have visited our school (Ben Greiner, Andreas Lange, Martin Kocher, David Byrne, Jean-Robert Tyran, Tom Wilkening, Werner Güth, Alan Kirman, Bettina Klaus, Ronald Stauber). My sincere gratitude goes to my two co-supervisors, Paul Pezani-Christou and Duygu Yengin, as well for their valuable advice and support when I was on the job market.

I have benefited a lot from the courses and workshops taught by Jacob Wong, Nicholas Sim, Seungmoon Choi and Tatyana Chesnokova. In particular, I want to thank Dmitriy Kvasov for enjoyable interactions and his contagious passion in teaching and talking game theory and mathematics. I thank all professional staffs in the School of Economics, especially to Allison Stokes and Sandra Elborough for their warmhearted help. I appreciate the training and support of conducting experiments from Mickey Chan. I also thank my laboratory fellows Kim Wu, David O'Callaghan and Hamish Gamble for running experiments together. Especially, I would like to thank Robert Garrard for his kindness in helping me out along the way.

Last but not the least I am very grateful to my friends for their companionship and comfort, especially during the tough time. I am forever indebted to my father for his unconditional love and support to whom I dedicate this thesis.

Contents

1	Introduction	1
2	The Psychological Effects During a Suspension in an English Auction	7
2.1	Introduction	7
2.2	Related Literature	10
2.2.1	English Auctions	10
2.2.2	Endowment Effect	12
2.3	Experiment Design and Procedure	14
2.4	Results	17
2.4.1	An overview of group bidding behaviors	18
2.4.2	Revenue	22
2.4.3	Individual Bidding Behavior	26
2.4.4	Discussion	29
2.5	Conclusion	30
3	Cheap Talk Delegation Experiments	32
3.1	Introduction	32
3.2	Design of Experiments	37
3.2.1	Cheap Talk Delegation Game	37
3.2.2	Gender Effect	38

3.2.3	Treatments and Parameter Values	39
3.2.4	Experimental Procedure	42
3.3	Results	44
3.3.1	Basic Data Summary	44
3.3.1.1	Performance in the Adding-up Task	44
3.3.1.2	Improvements in the Real Effort Task	46
3.3.1.3	Is There a Gender Gap in Confidence Level?	49
3.3.2	Delegation Frequency and Efficiency of Delegation	52
3.3.3	Individual Behavior of Delegates	56
3.3.3.1	How Players Message in the PMT?	57
3.3.3.2	How Players Message in the FMT?	59
3.3.3.3	Comparison of Messages between the Two Treatments	63
3.3.3.4	How Players Message in the FCT?	64
3.3.4	Behavior of Delegators	66
3.3.4.1	Performance of Delegators	67
3.3.4.2	Comparison of Messages Received with the Own Performance	68
3.3.4.3	Revisiting the Role of Initial beliefs	69
3.3.4.4	Probit Regressions	70
3.3.4.5	Is There a Gender Gap in Delegation in the PMT and the FMT?	74
3.3.4.6	How Delegators Delegate in the CFT?	75
3.4	Conclusion	76

4	Cooperation and Distributional Conflict with a Costly Contract Option	79
4.1	Introduction	79
4.2	Experimental Design:	84
4.2.1	Basic Model and Predictions	84
4.2.2	Parameters of Treatments	88
4.2.3	Experiment Procedure	91
4.3	Results	92
4.3.1	Social Welfare	92
4.3.2	Investment Stage: Contributions	95
4.3.3	How does an equal split contract affect contributions?	95
4.3.4	Contest Stage: Efforts	100
4.3.5	Frequency of Contracting	103
4.3.6	What drives people into a distributive fight?	104
4.4	Conclusion	106

List of Tables

2.4.1 Auction prices across the two treatments	23
3.3.1 Summary of Performance in the Adding-up Task	45
3.3.2 Performance in the Adding-up Task by Roles and Delegation Outcomes . . .	47
3.3.3 OLS Regression of Improvements	48
3.3.4 An OLS Regression of the Elicited Beliefs	51
3.3.5 An OLS Regression of Messaged Individual Performance	58
3.3.6 A Probit Regression of the Propensity for Truthful Reports	62
3.3.7 An OLS Regression of the Size of a Lie	63
3.3.8 A Probit Regression of Text Messages	66
3.3.9 Performance in <i>Task 1</i> of Delegators Conditional on Their Decision	68
3.3.10A Probit Model (1) of Delegation Decision	71
3.3.11A Probit model (2) of Delegation Decision	72
3.3.12A Probit Regression of Delegation Decision in the FCT	75
4.2.1 Session information	91
4.3.1 Average profit and average contribution	94
4.3.2 Pooled OLS regressions for individual contributions	100
4.3.3 An Pooled OLS Regression for Individual Efforts	102
4.3.4 A random-effects Probit model of entering a contest	106

List of Figures

2.4.1	Dynamics of individual bid across groups in the control treatment	19
2.4.2	Dynamics of individual bid across groups in the suspension treatment	21
2.4.3	CDF of winning bids by treatments	24
2.4.4	Box-plot of auction prices by different treatments	26
2.4.5	A summary of the winning proportion by treatments	27
3.2.1	Distribution of potential gain or loss from delegation	40
3.3.1	CDF of Number of Correctly Solved Questions by Gender	46
3.3.2	Cumulative Distribution of Self-reported Percentage Beliefs by Gender	50
3.3.3	Cumulative Distribution of Over-confidence by Gender	52
3.3.4	Comparisons across Treatments	53
3.3.5	Average Efficiency Level in Four Scenarios	55
3.3.6	Scatter-plot of Messages against Actual Performance in <i>Task 1</i>	57
3.3.7	Message Screen in the FMT	60
3.3.8	Discrete Histogram of Lie by Gender in the FMT	61
3.3.9	Histogram of the Size of Lies by Treatments	64
3.3.10	Performance of Delegators by Decision	68
3.3.11	Lowess Graph by Treatments	69
3.3.12	Cumulative Distribution of Initial Belief of Delegators	70
3.3.13	Delegation Frequency by Message Categories	73

3.3.14	Frequency of Being Chosen for Delegation	74
4.2.1	Game structure	85
4.2.2	Average effort in unconstrained low treatment	90
4.3.1	Average profit by period and treatment	93
4.3.2	Average contribution by treatments	96
4.3.3	Average contribution conditional on the outcome of an equal split contract	97
4.3.4	CDF of individual contributions conditional on the outcome of contracting	98
4.3.5	Actual efforts vs. Nash efforts in the ESC treatments	101
4.3.6	Frequency of contracting over periods by treatments	103
4.3.7	CDF of relative uneven degree in contributions by outcome of contracting	105

Chapter 1

Introduction

This dissertation exploits the tool of controlled laboratory experiments for the study of human behaviors in different economic contexts. In the real world, the potential contributing factors to a phenomenon of interest are usually correlated or sometimes even unobservable. Therefore, it is hard to conduct a clean analysis that allows for causal inference. One direct way to tackle this issue is to conduct an experiment to control the data generating process. Ideally, one wants to carry out the experiment in the real world. However, the cost of a field experiment can be prohibitive. In the extreme case, a failed social experiment may lead to disasters such as a riot or a war. A safer way, on the other hand, is to use logic and deduction by specifying a mathematical model to perform thought experiments. Nonetheless, abstract theories are conditionally true on a set of assumptions which are usually a compromise between realism and technical tractability. Therefore, laboratory experiments serve as a bridge to link the real world and an abstract theory of human behavior in different conditions with a bearable cost. Although the environment induced in the laboratory is artificial to a certain degree, we are willing to sacrifice some external validity for better internal validity to implement clean casual inference. Three independent chapters are included in the thesis.

In chapter two we study possible psychological processes that can take place during the suspension of an English auction and how the psychological effects can change people's

bidding behaviors. When the value of a good is unknown, auctions are an important mechanism to determine the market price. Among the four basic auction formats (i.e. the English, Dutch, first and second price sealed-bid auctions) outlined by Vickrey (1961), the English auction is the most familiar in real life. This format is particularly common in the real estate industry. A common structure for the sale in Australia usually goes as follows. Interested buyers gather in front of the house on a Sunday afternoon. Each buyer can raise the arm to increase their bid. A man in a suit hosts the auction and announces the ongoing bid aloud every time a buyer raises their arm. The bidding price goes up until no one calls a higher bid. The highest bidder receives the house by paying the last bid. Oddly, we observe that in quite a few the open bidding process is interrupted by an unexpected suspension. It seems that the auctioneer deliberately allows some time for bidders to ponder. Thus, we wonder if the suspension of an English auction can increase the seller's revenue?

We computerize the English auction in the laboratory for the same real good, namely a box of chocolates, with and without a stoppage. Computerization is vital to exclude auctioneer effects. Our initial conjecture is that an endowment effect can be triggered during the suspension which then leads to more aggressive bidding behaviors and therefore higher revenues for sellers. When the auction is stopped, bidders have time to imagine their consumption of the good and get a feeling for owning the good. The feeling of enjoying and owning the good can introduce extra attachment value to the good. If so, bidders would bid more aggressively to guarantee the win. In theory the auction price is determined by the bidder who has the second highest valuation. Ideally, we want stop the auction when the second highest bidder is holding the highest bid to maximize the effect on average prices. Since we can not control individual valuations for a box chocolates, we adopt groups of three bidders only to maximum ex ante the chance of targeting the right bidder. In the stop treatment, we suspend the auction after 90 seconds for all groups and place the box of chocolates in front of the highest bidders.

Unexpectedly, we find no significant difference in terms of the average auction price between the control treatments and the stop treatments. However, we find that the winning probability of the targeted subjects in the stop treatment groups is significantly lower than their counterparts in the control treatment. Thus, we infer that there must be two effects going on in different directions to cancel out each other such that the overall auction prices remain the same on average. The targeted subjects experience a cooling-off effect rather than an endowment effect and bid less aggressively later on. In contrast, the waiting subjects who have not been targeted experience a heating-up effect and bid more aggressively when the auction resumes.

In chapter three we study how much efficiency can be achieved in the context of task allocations when information is asymmetric but delegation and cheap talk are allowed. We introduce a novel three-person delegation game based on repeated real effort tasks. Additionally, we are interested in the gender gap in being chosen for delegation. Therefore, we choose the canonical addition task used in the study of Niederle & Vesterlund (2007) where they find that no performance differences across genders, but also find that female shy away from competition. We introduce the gender dimension by asking subjects to choose one avatar out of eight, i.e. four different ethnic background with two genders each, and match one female and one male as the two available subjects in a group. Everyone performs the adding-up task and is incentivised to guess their relative ranking among all subjects as precisely as possible. Then a third subject (male or female) is grouped with two available subjects. Before the repetition of the adding-up task, the two available delegates can send a costless message about their performance to the delegator. The delegator can read the two messages and see their avatars at the same time to make a delegation decision. Delegators can decide either to use their own performance or the chosen player's future performance to calculate their payout for their earnings in the first period real effort task. We adopt a piece rate salary for a randomly selected payout of one real effort task. In this way, we are able to have a clean delegation motive for efficiency since everyone has to perform the repeated

task.

The focus of the study is how the coarseness and the size of the message space influences the degree of information transmission and consequently efficiency of delegation. We misalign the preferences between principal and agent by requiring a bonus fee to be transferred from the delegator to the chosen player if delegation happens. On the equilibrium path, we expect that players exaggerate their performance and delegators would disregard the messages. In other words, we implement a setting such that only babbling equilibria exist in theory. Nonetheless, experimental research of lie aversion (Gneezy et al. 2013, López-Pérez & Spiegelman 2013, Lundquist et al. 2009) shows that subjects tend to reveal more information than the most informative equilibrium predicted by a Bayesian Nash Equilibrium. Then we vary the message space in two dimensions. In the baseline treatment (i.e. Precise Message Treatment, PMT) players chose any number they want to represent the number of solved problems in the first adding-up task. Along one dimension we introduce noise by partitioning the message space into intervals (i.e. Fuzzy Message Treatment, FMT). Players in the FMT choose an interval which indicates where their performance belongs. Along the other dimension we increase the size of the message space by allowing free text communication (i.e. Free Communication Treatment, FCT). In the FCT players can message anything which leads to an infinitely rich message space.

We find a significant efficiency improvement across treatments compared to the scenario where delegation is not allowed. The highest efficiency is reached in the FCT where more information is transmitted. Approximately, one third of players reported their true performance across treatments and the size of the average lie is moderate. Surprisingly, information transmission is found in the FMT and the FCT but not in the baseline treatment. When messages are in the form of intervals or in free-form text, delegators respond to the favorable message by increasing the frequency of delegation to the corresponding player. In contrast, delegators in the baseline treatment are less driven by the messages than by

their own performance and their subjective belief on their relative ranking among all subjects. We further identify other regarding preferences beyond lie aversion using the interval setting in the FMT. Players whose actual performances locate to the left end of an interval tend to lie less in terms of both frequency and magnitude to reduce the potential harm to the payoff of the delegator. We find male players are more frequently chosen than female players in the PMT. Unfortunately, we are not able to identify whether the effect is coming from the supply side or demand side given the limited number of observation we have got.

In the last chapter we study how effective an option to contract in advance is in resolving future distributional conflicts. We extend a two-stage cooperation and contest game of Bayer (2016b) by providing a contract option in advance. When property rights of common assets are not well defined, people have to invest wasteful resources to fight over the distribution. Rational players who foresee the severity of distributional fights may refrain themselves from cooperation in the first place. This creates a further hold-up problem in the cooperation game where a social dilemma already exists due to individual incentive to free-ride. Bayer finds that subjects are able to overcome both the social dilemma and the hold up problem in the cooperation game but social welfare is worse than the equilibrium prediction. To remedy the worsened social welfare, we introduce a contract option in advance to specify a sharing rule to split the proceeds. Implementation of a contract requires a mutual agreement from both players in a group and incurs a fixed cost for each.

Bayer's two stage game becomes a natural control treatment where a contract option is absent. The welfare enhancing effect of a contract option depends on both the type and cost of the contract. In the baseline treatment we keep the social dilemma in the cooperation game by adopting an equal split social norm. Then we calibrate the implementation cost to be the average effort incurred in Bayer's *unconstrained_low* treatment. On the equilibrium path, we expect that no players would take the contract and would invest nothing in the cooperation game. We would also expect no welfare improvement at all if all subjects,

on average, take the contract but invest in the same fashion as in the control treatment. Interestingly, we observe a significant treatment effect in terms of average profit. We find that an equal split contract has a positive effect on individual contributions. However, only approximately 30% of subjects manage to have a mutual agreement on the contract and the contracting frequency drops dramatically in early rounds. Therefore, we decrease the implementation cost as low as possible keeping the sharing rule the same. Lower cost postpones the decline of contracting frequency and increases the average percentage up to 50%. The positive effect of the equal split contract on individual contribution remains significant, but lower contracting cost attracts more free riders. Since the contest game provides subjects a chance to fight over their fair share, the contracting frequency falls eventually in spite of the low contracting cost. Consistent with this conjecture we also find evidence that subjects fight harder if they contribute more in the low cost treatment. In a further treatment, we keep the cost the same but adopt a proportional split of the proceeds conditional on individual contribution to remove the social dilemma in the cooperation game. As expected, a majority of subjects can reach a mutual agreement to stay on the contract and fully invest in the cooperation game.

Chapter 2

The Psychological Effects During a Suspension in an English Auction

2.1 Introduction

Since the first regular wool auction was conducted in Melbourne's Bourke Street in 1850, auctions have become a popular method of selling commodities in Australia. The general perception is that auctions are likely to generate relatively more revenue than private negotiations. This view is particularly strongly held in the real-estate industry. Real estate auctions typically follow the format of an English auction. One particular phenomenon which has caught our attention is that the auctioneers often stop the bidding process for a while to "consult" with the sellers of a house. The halt usually happens if the bidding dries up. The auctioneer may say they are 'going inside' or 'seeking advice or instructions' from the seller. This phenomenon seems more prevailing that one should know before entering an housing auction as mentioned in a recent post by an state's consumer affair regulator.¹ This paper investigates if and how the suspension during the process of the auction changes people's

¹See: <http://www.realestate.com.au/advice/need-know-auction-hammer-falls/>

bidding behavior. We are particularly interested in the question if suspending an auction triggers some psychological processes that lead to more aggressive bidding and therefore higher revenues.

In an English auction bidders have to submit increasing bids. The bidder bidding the highest when nobody else wants to bid anymore is the winner of the auction, receives the object and pays his or her bid. Hence, some information on a bidder's valuation will be revealed in the course of the auction. Each bid represents the minimum willingness to pay of the bidder at that time, assuming independent private values². When the auction is suspended, no new information is revealed. The implication is that from a purely game-theoretic point of view, the suspension should not have any effect on behavior. Hence, any difference in behavior induced by the suspension of an auction has to have other psychological causes. One possible scenario may be that the bidder holding the highest bid during the stoppage might develop an attachment to the auctioned good. Such a psychological ownership can emerge when the target is visible, attractive and can be experienced by the individual (Pierce et al. 2003). English auctions for selling houses are commonly held at the property with bidders inspecting the house before the auction and submitting their bids from the road in front of the house. Australian property auctions therefore meet the conditions which favor the formation of psychological ownership. Hence our major conjecture is that psychological ownership might positively affect a bidder's assessment of the valuation of an auctioned good and generates an effect similar to the endowment effect (Kahneman et al. 1990, Knetsch 1989).

Real-world property auctions are unfortunately not suitable for studying the effect of stoppages. One would require many identical properties auctioned off to bidders from the same bidder pool with and without a stoppage in the auctions. There are no two identical prop-

²We aware that values may be affiliated due to the prevailing market price in real estate auctions. However, we focus on the consumption value of houses and if extra attachment value can be added through some psychological process. Therefore, we deliberately chose an auctioned good for our experiments to meet the independent private value assumption.

erties. Moreover, it is impractical for experimenter to control which auctions will have stoppages and which not. Using data from auctions, where auctioneers determine if the auctions are suspended for some time or not, would be problematic because of endogeneity. For this reason experiments are the methodology of choice for our question. Here we can auction off the same object (a specific box of chocolates) with and without stopping the auction. Moreover, by using a computerized English auction we exclude any auctioneer effects. Since we randomly assign treatments to sessions (i.e. auctions with stoppage and without), we can draw inferences about the impact of the stoppage once we have sufficiently many auctions.

In the treatment with a stoppage, we suspend the auction after 90 seconds for 180 seconds. For those who hold the highest bid when the auction is suspended, we place the object, a box of chocolates, in front of them and inform them privately that they can keep it if they win the auction. We call these subjects the targeted subjects. For the rest of the subjects (i.e. waiting subjects), we just remind them privately that they are not the highest bidder and that the auction resumes after 180 seconds.

Based on our initial conjecture of a stoppage leading to an endowment effect we would expect the following: 1) Auction prices are higher in the treatment group (i.e. in the auctions with stoppage) than in the control group. 2) Targeted subjects in the treatment group have higher probabilities of winning the auction than their counterpart in the control group. If we want to maximize the effect on average prices, then it would be ideal to target the bidder with the second highest valuation of the object, as the equilibrium winning bid in an English auction is equal to the second highest valuation. However we can not control the valuation subjects have for a box of *Lindt* chocolates. Chance will play a role in which subject is targeted. As we expect the endowment effect to be triggered in the subject that is targeted only, we require a reasonably large number of observations. To see this recall that an increased willingness to pay of an individual bidder only influences the sales price

in an English auction if it increases the second highest valuation. Similarly the endowment effect will only have an influence on who wins the auction if it leads to the targeted person's willingness to pay after the stoppage to be the highest in the among all bidders while it was not before.

Unexpectedly, we find that there is no significant difference in terms of the auction prices in suspended versus continuous auctions. Our initial conjecture was that our study was under-powered due to not enough relevant subjects being targeted. However, this conjecture proved incorrect, as we do find a significant effect of stoppages on the winning probability. Surprisingly, the targeted subjects in the experimental treatment have a lower probability of winning the auction than their counterparts in the control treatment. After the suspension of the auction, the targeted subjects bid less aggressively. Since the overall auction prices remain the same on average, we conclude that the targeted subjects bid less aggressively than their counterparts in the continuous auction. Instead of an endowment effect the targeted subjects cooled off, while the subjects that did not have the highest bid during the suspension became more aggressive once the auction was opened again.

2.2 Related Literature

There is no literature directly addressing this or a similar question. Nonetheless we can relate our study to two branches of literature (i.e. the English auction and the endowment effect), which provide the background to our experimental design.

2.2.1 English Auctions

Auctions in practice have a long history, dating back to 500 B.C., when women were auctioned off for marriage in ancient Babylon. Among various formats of auctions, the English

auction is the most prevalent auction format. In the English auction, the price is successively raised until only one bidder remains. This can be done by having an auctioneer announce prices, or by having bidders call the bids themselves, or by having bids submitted electronically with the current best bid posted (McAfee & McMillan 1987). Vickrey (1961) in his pioneering work showed that the outcome of an English auction coincides with that of a second-price sealed bid auction if the bidders play dominant strategies. In a second-price sealed bid auction bidders simultaneously submit their bids. Then the highest bidder wins the object and pays the bid of the second-highest bidder. Later, Milgrom and Weber(1982) provided a characterization of equilibria in an auction where prices increase continuously and bidders decide when to quit. Such an auction is equivalent to the standard English auction but much easier to analyze.

English auctions are particularly difficult to analyze if values are correlated (see Kamecke and Izmalkov for examples). For our purpose we want to get as close to the auction format used in real estate auctions and therefore use a continuous open outcry English auction. If, as we assume the valuations of *Lindt* chocolates are purely private, then one would expect subjects to incrementally increase their bids up to their private value. If every bidder follows this bidding strategy, then the winner of the good will be the bidder with the highest value and the selling price equals to the second highest value. Thus we expect the resulting prices of the English auction to provide us with information on the valuation of the second highest bidder.

Various experimental studies on bidding behavior using induced value in different auction formats suggests significant overbidding in first and second sealed bid auctions but not in ascending auctions(Cooper & Fang 2008, Kagel & Levin 1993) Another strand of empirical literature(Malmendier & Lee 2011, Jones 2011) identifies overbidding using eBay online auction data in a dynamic second price auction format. This format requires bidders to submit their maximum willingness to pay and an automated proxy system increases their bids up

to that amount as competing bids come in. Numerous followed-up research tests various possible explanations including risk aversion, regret, bounded rationality and non-standard preferences such as “spite” or “joy of winning” (Kagel & Roth 2016). However, those factors identified in the literature play no role in our experiment since we conjecture that endowment effect may be relevant in our context.

A few recent experimental studies (Ariely & Simonson 2003, Heyman et al. 2004, Ehrhart et al. 2015) also test if pseudo-endowment effect attributes overbidding. Ariely & Simonson (2003) and Heyman et al. (2004) speculate that a longer duration of being the lead bidder may trigger the pseudo-endowment effect. Ehrhart et al. (2015) further hypothesize that different auction formats differs in the strength of the pseudo-endowment they evoke. In contrast, our design aims to address a puzzle observed in real estate auctions rather than test the pseudo-endowment effect directly in an auction setting. We are interested in whether suspension of an English auction may trigger some psychological effects that would increase the seller’s revenue.

2.2.2 Endowment Effect

The endowment effect coined by Thaler (1980) is an observation from experiments that people seem to attach additional value to things they own simply because the goods belong to them. There are two predominant ways of conducting experiments to test if endowment effect exist. The exchange paradigm proposed by Knetsch (1989) involves a choice between two goods. By manipulating the initial endowment condition, Knetsch observed that subjects are unwilling to trade their current position for an alternative (i.e. exchange asymmetries exist). Following this line of paradigm, substantial disparities between assessments of valuations have been reported for a number of consumer goods. However, the disparities disappear in cases of transparent choices, such as induced-value experiments where specific values are assigned by experimenters and subjects act on the basis of these valuations (can

not cite here and convert thesis). Thus we decided to choose a real good for our auction experiment instead of induced values. In order to quantify the magnitude of the endowment effect, the valuation paradigm is adopted by Kahneman et al. (1990), who conduct experiments in a market setting with opportunities to learn and find out that on average selling prices are significantly higher than buying prices. Selling prices reflect subjects' willingness to pay (i.e. WTP) while buying prices reflect subjects' willingness to accept (i.e. WTA). In their studies WTP is substantially larger than (i.e. more than twice) WTA. The subsequent literature (Plott & Zeiler 2005) points out the endowment effect is sensitive to the experimental procedures such as an incentive-compatible elicitation device, training, paid practice, and anonymity. We believe that our subject of study, the English auction has the advantage that we avoid most of these issues. One merit of the English auction is that its rules are simple and intuitively easy to understand for subjects, especially with a commonly known real object such as chocolates. The other merit is that the auction mechanism provides an alternative way of eliciting valuations to the complicated Becker-DeGroot-Marschak mechanism (Becker et al. 1964), which is harder to understand for subjects. A disadvantage is that only the second-highest valuation is elicited perfectly, while only lower or upper bounds for the other valuations can be deduced. In order to avoid the house money effect (i.e. people tend to make a more risky investment with the money that they are provided with by the experimenter (Thaler & Johnson 1990)), we decided to couple our experiment with another unrelated one. Our subjects were able to use the payment they received from another unrelated experiment to bid in the chocolate auction. However subjects are not informed with their precise earnings from previous to minimize the income effect on their biddings. An important paper by List (2003) points out that real market experiences can eliminate the endowment effect. Our experimental settings is as close as possible to that of a real English auction. Our subjects might not have a lot of experience with auctions but we also do not expect bidders in real auctions to have a lot of experience. It remains to mention that in our case the effect we expect to be in operation is not a true endowment effect, as during the

stop the bidders are not actually owning the object. We call the effect we expect to observe a pseudo endowment effect.

2.3 Experiment Design and Procedure

Our initial conjecture of how the suspension during an English auction in real life can change people's bidding behavior goes as follows: the down-time and visual contact with the object allows for reflections. The bidder with the highest bid during the pause develops a feeling of being on top and develops an attachment to the object. This attachment shift the reference point from the status quo, i.e. not owning the object to owning the object. Taking owning the object as the new reference point, dependent on the preferences the bidder's willingness to pay has changed. A loss averse bidder (Kahneman & Tversky 1979, Tversky & Kahneman 1991) is now willing to bid more than her initial valuation of the object to avoid the painful sensation of losing it. In short we expect higher auction price and higher winning probability if the endowment effect takes place.

We test this hypothesis by simply bringing an English auction for chocolates to the laboratory and vary if there is a stoppage or not. In the control group, we conduct an English auction without any intervention. In the treatment group, we conduct the auction in the same fashion but with a short suspension. The object to be auctioned off is a box of *Lindor* chocolates. This chocolate is a valued and well-known brand in Australia. The use of a real object, instead of using an induced value is necessary as it is unlikely that an induced value (i.e. the winner of the auction will be paid a previously determined privately known amount of money less his bid) can generate an endowment effect(Lange & Ratan 2010). The disadvantage of using a real object is the loss of control over valuations. Further, we believe that a box of *Lindor* chocolates captures the private value aspect of properties to some degree. The assessment of the valuation of chocolates depends primarily on subjects'

preference for chocolates. Additionally, chocolates can be consumed immediately unlike like lottery tickets or movie tickets. The choice also aims to alleviate the issue of uncontrolled risk preferences. We allow subjects to have a taste of the chocolates before they can bid for them.

We are aware that buying a house is a much more important decision than buying some chocolates. The stakes are much higher. Despite this clear difference, we still believe that our experiment is useful for testing if the observation that house auctions are often suspended could be caused by real-estate agents capitalizing on the endowment effect.

Our design is guided by the objective to test if 1) the auction price is higher on average in the treatment group and if 2) the probability of winning the auction is higher for the targeted subject in the treatment group than for the person that would have been targeted in the control condition?

We use the software package z-Tree (Fischbacher 2007) to program an English auction to be run in the laboratory. In real life property auctions an auctioneer announces increasing bids communicated via hand signals by bidders and terminates the auction once no new bids are received anymore. We use a table with all the past bids and a clock to emulate the auctioneer by a computer. The clock indicates the remaining time of the auction and restarts once a subject enters a new bid. The time limit for every bid is 30 seconds. A new bid has to be higher than the current highest bid and can be any non-negative one decimal number. The auction ends when the clock runs out and the highest bid wins. Once the auction is finished all subjects are informed about the outcome. The real transaction happens in the end of the experiment. The subject who won the auction receives the box of chocolates and paid her bid out of the payment for an unrelated experiment that had taken place before the auction.³ Note that liquidity constraint did not come into play as the payment for the previous experiment was higher by magnitudes than the expected of the chocolates. The

³The unrelated experiment is a variation of contest games.

purpose of coupling with an unrelated experiment is to avoid any welfare effect.

There are quite a few critical design choices when it come to implementing the stop treatment. Theoretically, the price in the English auction is determined by the bidder with the second highest valuation if every bidder follows the optimal bidding strategy. In this case the bidders quit the auction only when the ongoing price surpass their valuations. We ideally would like to stop the auction when the second highest valuation bidder is holding the highest bid⁴. If the leading bid during the suspension gives this subject a feeling of owning and enjoying the good, the pseudo endowment effect increases the sales price through the increased valuation of the second highest bidder. Unfortunately, ex ante we do not know the subjects' valuation of chocolates. To optimize the chance of targeting the "right" subject, we choose small auction groups of only three bidders. Then we have a third chance to hit the second highest valuation subject. This gives us a good chance of hitting the second highest bidder.

In order to keep the treatment auctions comparable we decided to suspend all the auctions in the stop treatment at the same time and for the same amount of time. We run the control group first to get an idea about the typical auction speed and duration. The quickest auction group ended after 36 seconds while the slowest auction group ended after 398 seconds. To guarantee the auction is still on for about 70% of groups we choose to stop the auction after 90 seconds. Two types of messages were sent to subjects when the auction was suspended. For those who held the highest bid at that time, we put a box of chocolates in front of them and informed them that they could keep the chocolates if they won the auction. Alternatively, for those who did not hold the highest bid, we just told them that the auction would resume after a short suspension. To allow us enough time to place the box of chocolates in front of to the right subjects, we suspended the auction for 180 seconds. After the auction resumed, we collected all boxes of chocolates. Compared to the average auction duration,

⁴If the second and third values are close together, then even pausing when the third-highest-value bidder had the last bid could affect the final price. However, we can not control for individual values. Eventually, we want to avoid the highest valuation bidder is targeted during the stoppage.

180 seconds of suspension can be considered as a long time span.

We ran six sessions with 18, 24, or 27 subjects respectively and had a total of 144 subjects. Subjects were randomly matched into a group of three members. We had 22 groups in the control and 26 groups in the stop treatment⁵. The procedure was as follows: after all subjects had finish the previous experiment, we announced that an auction for the purchase of a box of chocolates was about to be held. In order to provide common knowledge of the quality of the chocolates, we allow all subjects the option to a tiny piece of the chocolates. After everyone had a taste of the chocolates and instructions were read aloud. Before the auction began, we confirmed that no one had any remaining questions about the English auction that was about to start. In the treatment groups the English auction took place in the same fashion as in the control groups, with one exception. We suspend the auction for 180 seconds after 90 seconds. The auctions was only conducted once. The winner of the auction gets the chocolate as bidding finished in the end of the experiment. His or her winning bid is automatically deducted from the payment of the previous experiment.

2.4 Results

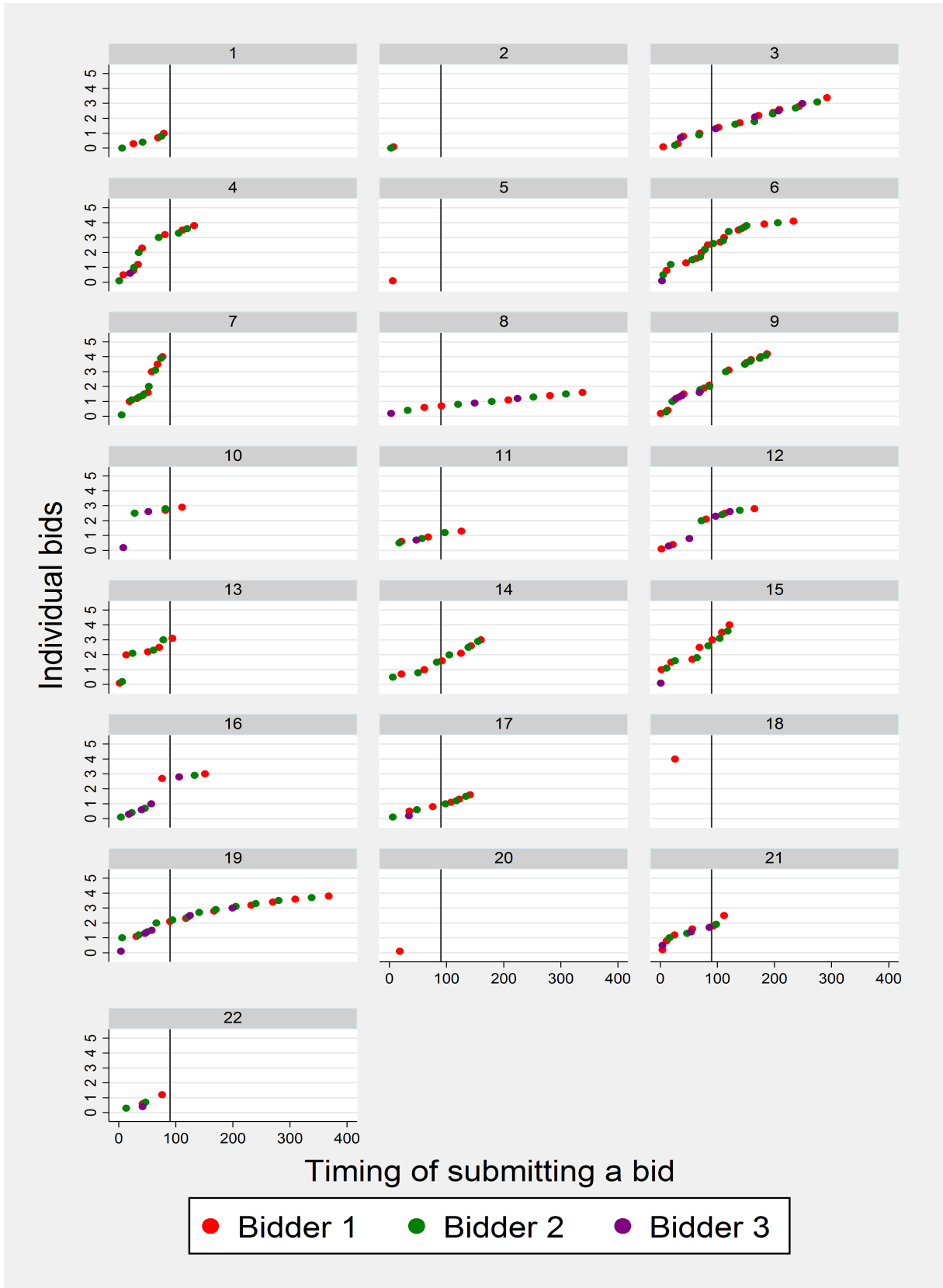
In this session we present our results in the following sequence. We start by an overview of group bidding behaviors to get an intuitive idea of the bidding dynamics. Then we test the two main hypotheses in terms of 1) auction prices and 2) winning probabilities. In the end we discuss possible explanations of our findings.

⁵A subject in the stop treatment made a mistake by entering 16 when she intended for raising a bid of 1.6. She informed us as immediately during the experiment. So we remove the data of her group all together.

2.4.1 An overview of group bidding behaviors

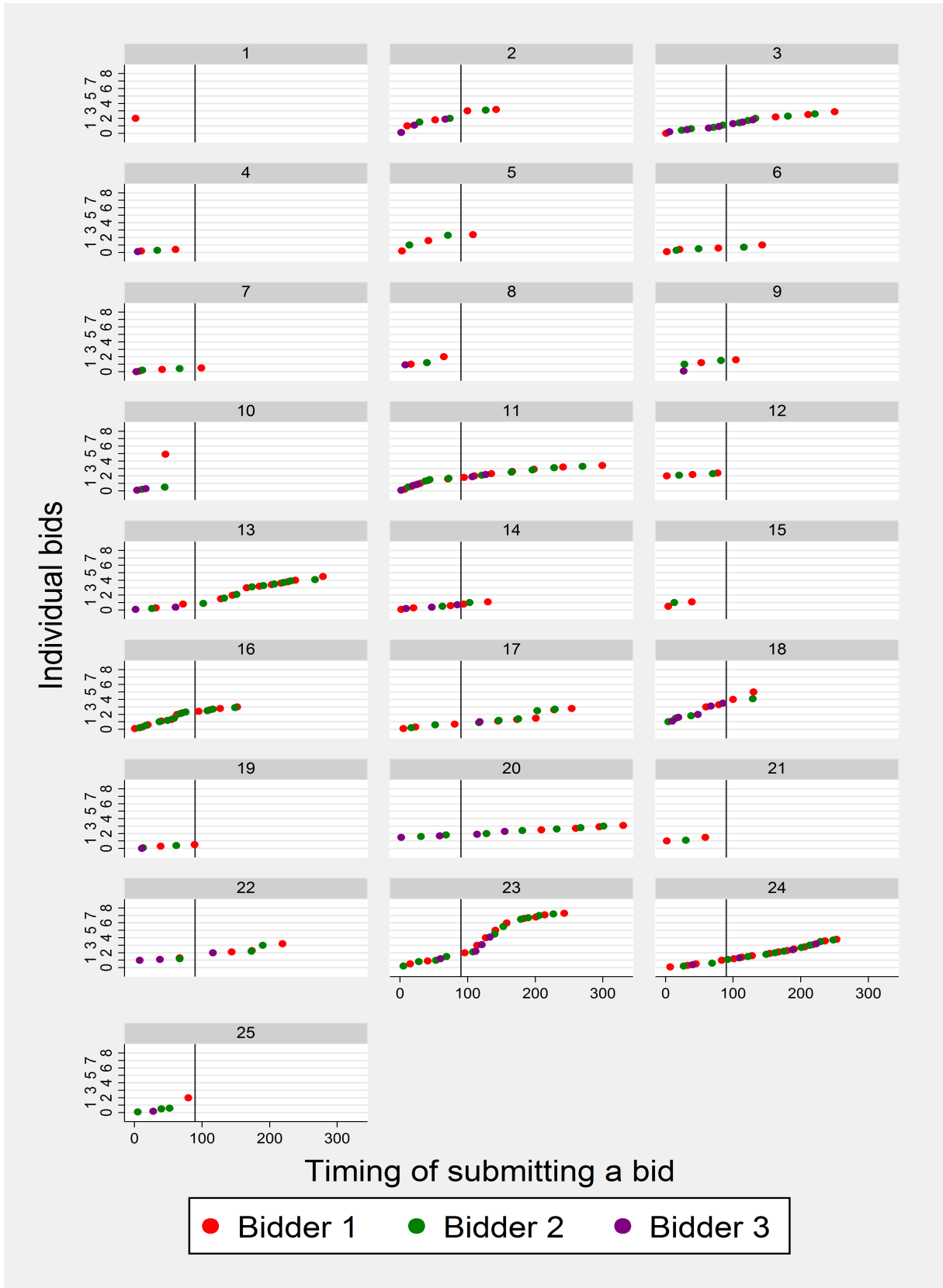
Unlike the theoretically equivalent second price sealed bid auction, the strategy space of an English auction is huge, as it contains a full contingent plan for all bidders for any possible bidding history. For this reason, bidding behavior can be highly heterogeneous within each group. In order to obtain an intuitive feeling for the data, we plot the bidding dynamics (i.e. bids along the time dimension) for both the control treatment and the stop treatment. Figure 2.4.1 and Figure 2.4.2 provide scatter plots of bids against the corresponding submission time, which captures the bidding dynamics. In the stop treatment we specify a vertical line at the 90th second to indicate the 180-second stoppage. In order to have an easy read of graphs, we relabel IDs of bidders using different colors to indicate types. Bids by the eventual winner (i.e. Bidder 1) of the auction are marked in red, the bids of the bidder who had the second highest bid (i.e. Bidder 2) are coded in green, while the bids of the bidder that dropped out first (i.e. Bidder 3) are depicted in purple.

Figure 2.4.1: Dynamics of individual bid across groups in the control treatment



In the control treatment we conducted an ordinary English auction without any interruptions. If subjects are not interested in chocolates at all, they can submit a zero bid and the auction will end soon. For instance, one can look at the group 2, 5 and 20. When subjects have a strong interest in chocolates, we would expect a higher auction price being realized. However, a higher auction price does not necessarily imply a longer auction. The duration of an auction depends on the particular bidding dynamics within each group. As shown in the group 4, 6, 7, 15, 18 and 19, for example, the duration of the auction varies in spite of the similar auction price. In extreme cases, the auction ended instantly in the group 18 with just one active bidder while the auction lasted almost 400 seconds in the group 19 with three active bidders raising bids consecutively. Recall in theory the optimal bidding strategy is to raise bids incrementally to the private value of individuals. Nonetheless, in reality bidding jumps and sniping behaviors are common. Here we also find distinct bidding jumps, such as in the group 13 and 17. Almost every one waited to the last second to raise a new bid and consequently the auction became lengthy in the group 8.

Figure 2.4.2: Dynamics of individual bid across groups in the suspension treatment



Given the observed bidding behavior in the control treatment, we decided to suspend the auction after the 90th second to allow enough groups getting the stop treatment. We implement the stoppage for 17 groups out of 25 as shown in Figure 2.4.2.. Compared to the maximum price (i.e. around 4 Dollars) achieved in the control treatment, we have three groups (i.e. 13, 18, 23) ending the auction with higher prices. In particular the bidding dynamics in the group 23 seems to fit our initial conjecture. Assume that the sequence of bidders who left the auction reflects the ranking of their valuations. Then Bidder 1 must hold the highest valuation while Bidder 2 hold the second highest valuation. In the group 23, we seemed to have targeted the right subject who competed more aggressively later on, which drives the auction price up to near 7.5 Dollars. However as one can see in the group 14 and 18, we also presumably have targeted the wrong subject (i.e. the least valuation bidder). In this case it is hard to observe any treatment effect, especially when the effect size is small.

2.4.2 Revenue

We have 22 groups in the control (i.e. auctions without stop) and 26 groups in the stop treatment. In the stop treatment we have 9 groups finishing before the 90 seconds including one group in which a subject made a typing mistake. Thus we end up with 17 observations where the event of interest, i.e. the suspension of the auction, occurred. For an unbiased comparison we also disregard the groups in the control treatment that finish before 90 seconds. Note that this is necessary for a comparison without selection issues. Also note that in the instructions subjects were not told that a suspension would happen. Therefore, the groups that did not make it into 90 seconds should had exactly the same information regardless of the treatment they were assigned to. So we have 15 observation in the control that comparable to the stop treatment. Table 2.4.1 gives an basic summary of auction prices across the control and the stop treatment.

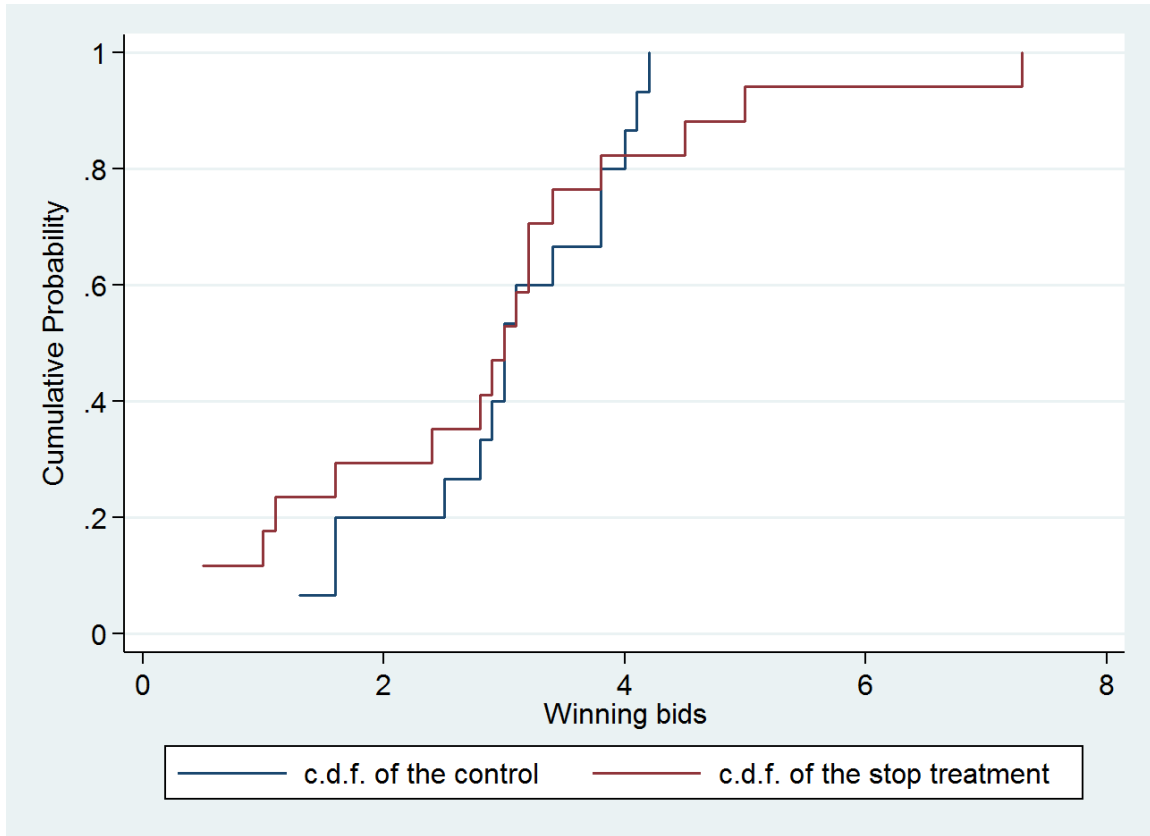
Table 2.4.1: Auction prices across the two treatments

Treatments	min	max	mean	p10	p25	p50	p75	p90	p95	sd
None stop	1.3	4.2	3.01	1.6	2.5	3	3.8	4.1	4.2	0.93
Stop	.5	7.3	2.95	.5	1.6	3	3.4	5	7.3	1.73

Assume the observed price data follows a normal distribution and use information of sample variance. We can simulate a detectable effect size of 1.63 dollars by choosing the significance level at 5% and the power at 90%. Compared to the market price, i.e. 15 dollars, our current design would allow an effect size of 11%. Now if we increase the sample size of the treatment (due to the concern of inability of targeting the “right” subject) gradually up to 34, the detectable effect size will correspondingly decrease to 1.26 dollars or 8% in terms of percentage. As one can see, even a doubled sample size can only improve the detectable effective size by 3%, which is quite marginal. Therefore, we believe the overall power of the design is not our major concern.

There is a large difference in the maximum price (i.e. 3.1 dollars) but a small difference in the minimum price (i.e. 0.8 dollars). Comparing the same percentile of the distributions reveals that the control groups have slightly higher prices below the 75% percentile, while the prices are higher in the right tail in the stop treatment. The centrality measures for the two distributions are quite close as both mean and median are around 3 Dollars. We have a larger (about twice as high a standard deviation) dispersion of auction prices in the treatment than in the control. Finally, we can look at the cumulative distributions of the winning bids by treatments as shown in Figure 2.4.3. The blue line represents the control and the red line the stop treatment. The two distributions are very close in the center and diverge somewhat in the tails.

Figure 2.4.3: CDF of winning bids by treatments



If our stop treatment invokes the endowment effect on the relevant subjects (i.e. the second highest valuation), we would expect that the auction prices in the stop treatment are larger on average than the prices in the control treatment. To test the treatment effect, we use the Mann-Whitney U-test, with the resulting auction prices as the independent observations. Suppose the auction prices in the two groups come from two independent populations T and C . We specify the null hypothesis as $H_0 : T = C$, suggesting the auction prices of the two groups are drawn from the same population. The alternative hypothesis is $H_1 : T \neq C$, stating the auction prices of the two groups come from different populations. Using the rank information of the data from both groups (15 observations in the control and 17 in the stop treatment), we find that we cannot reject the null hypothesis that the prices were drawn from the same distributions ($p=0.61$).

As a powerful alternative to the Wilcoxon-Mann-Whitney rank-sum test, we also conducted the Fisher-Pitman permutation test for two independent samples. The permutation test uses the information of the difference between any two data points and exhausts all the permutation possibilities to calculate the p-value. We can form the same hypotheses H_0 and H_1 as before. The null hypothesis H_0 states there is no difference between the distributions where the two groups of auction price are drawn from. The alternative hypothesis H_1 indicates that there is a difference with respect of the population distributions. If the null hypothesis H_0 were true, we would believe that both control and treatment auction prices are from the same distribution and should be observed with an equal chance. In our case we have a 17 out of 39 possible ways to permute the positions of data (i.e. 51021000000 different combinations). For each possibility, we can calculate the difference of the sums according to the theoretically possible distribution under H_0 and compare it with the same measure evaluated with the observed data ⁶.

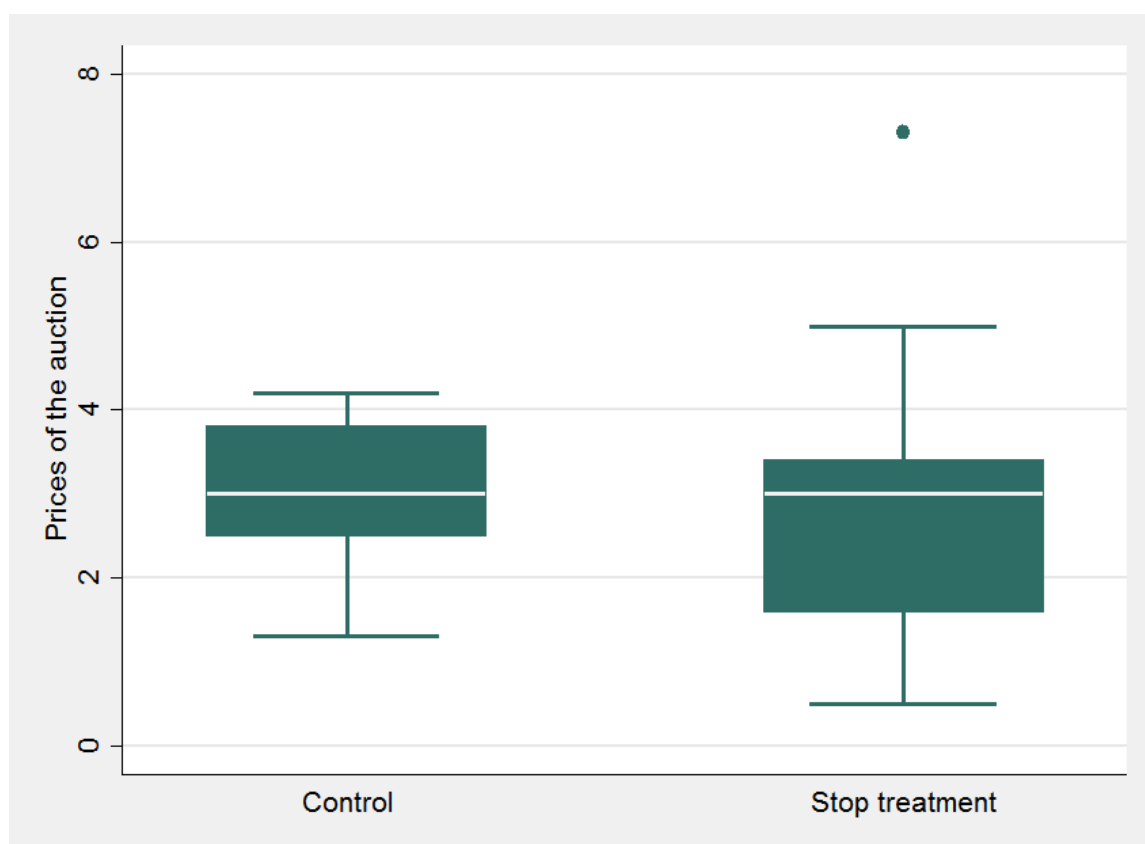
Given the computational complexity of the exact algorithm, we run a 200000 runs Monte-carlo simulation to approximate the p-value. We obtain a two-tailed p-value of 0.841. The probability of falsely rejecting the null hypothesis that the auction prices in control group equal to that in treatment groups is 84.1% which confirms the result obtained by using a rank-sum test.

In summary we can not reject the null hypothesis that there is no difference in terms of the auction prices between the control and the stop treatment. However we need to be cautious to draw the conclusion that our treatment has no effect. Not being able to reject the null hypothesis could also be due to a lack of power. The inability to detect the truth suggests the issue of the power of tests. We have tried two theoretical powerful non-parametric hypothesis tests, especially the imputation test which displays asymptotic power efficiency of 100% (Siegel 1956). It could be that our sample size is not large enough such that no test can be sufficiently powerful to detect a meaningful difference. A visualization of the

⁶For the technical details, refer to Kaiser (2007).

winning bid distributions in the two treatments via a box plot in Figure ?? suggests that the center of the two distributions do not differ at all. It appears though that the variance of prices is higher in the stop treatment. A Levine test for the equality of variance reveals that there is a weak significant difference ($p=0.08$) in variance across treatments. However the result is not robust among different statistics (i.e. constructed by using mean value or median value or trimmed mean value).

Figure 2.4.4: Box-plot of auction prices by different treatments

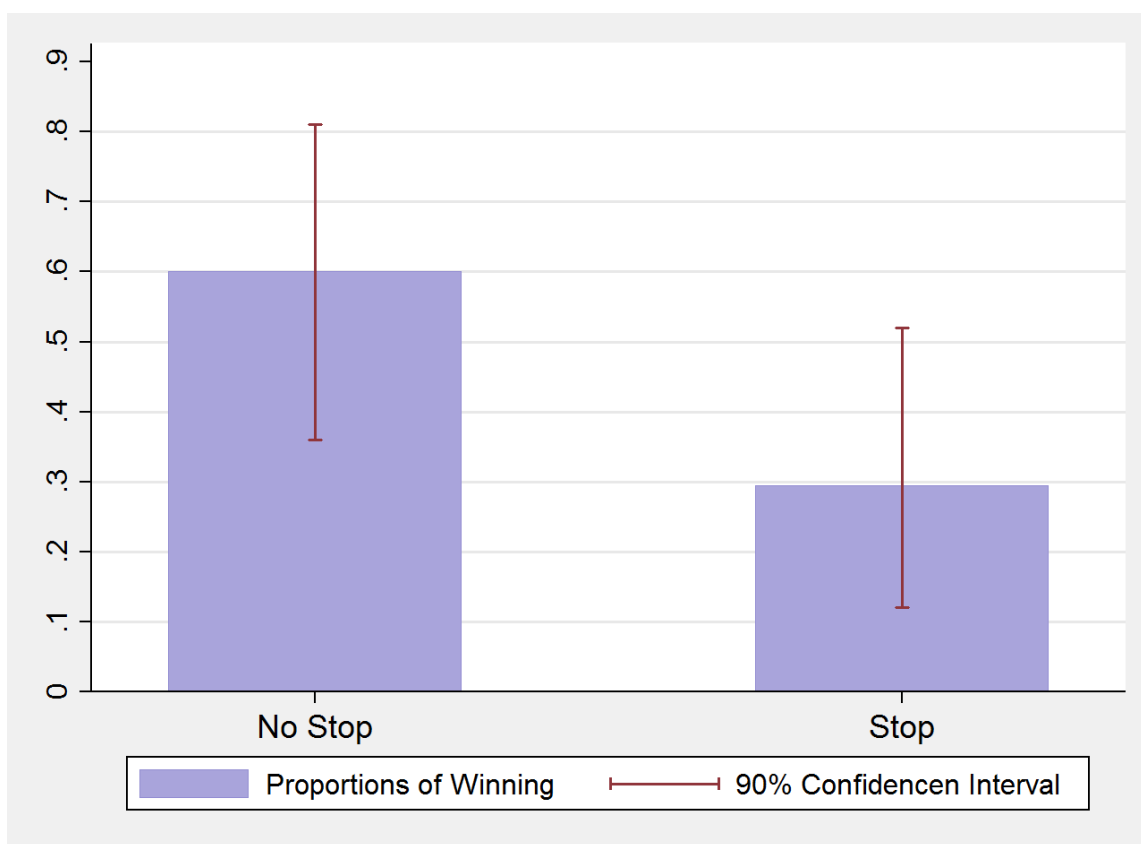


2.4.3 Individual Bidding Behavior

Our initial hypothesis was that the highest bidder during the suspension of an auction develops an attachment to the object which translates into an increased willingness to pay. If our

hypothesis is true, then we would expect the targeted subjects in the stop treatment to win their auctions more frequently than the subjects in the control treatment that would have been targeted if the auction were suspended. To test the conjecture we compare the winning probability of the targeted subjects with their counterparts who also held the highest bids at the 90-second mark of the auctions in the control treatment. We record the outcome of the targeted subjects and their counterparts as 1 for winning the auction and 0 for losing the auction. Figure 2.4.5 shows the fraction targeted winners in the stop treatment and the fractions of their counterparts in the control treatment.

Figure 2.4.5: A summary of the winning proportion by treatments



Note: [1] Sample size is 15 in the control treatment and 17 in the stop treatment respectively.
[2] Winning proportions are 0.60 in the control group and 0.29 for the stop treatment.

It is surprising that contrary to our initial conjecture the probability of winning the auction

for an average targeted subject is about half that of her for the counterpart in the control treatment. To confirm that the difference is unlikely to be purely caused by chance, we conduct a two sample test of proportions. Stata reports a 0.082 p-value under a two-tailed alternative hypothesis specification. Thus the winning probability is significantly smaller in the stop treatment than in the control treatment. In order to exclude any placebo effect, we repeat the same comparison as if the auction were stopped after 80 seconds. In this scenario, the average winning proportion remains the same in the control treatment but increases to 0.35 in the stop treatment. Now the difference in proportions between the two treatments is not no longer significant ($p=0.162$, two-sided).

We can further infer that the targeted subjects in the stop treatment bid less aggressively after the 180 seconds suspension which leads to winning less often than the counterparts in the control treatment. It seems that our targeted subjects experience the opposite of the conjectured endowment effect. We borrow the term cooling-off effect from the analysis of emotions in decision making in order to describe what happens to our targeted subjects. The main finding in the behavioral decision theory literature (Cardella & Chiu 2012, Grimm & Mengel 2011, ?, Bosman et al. 2001) is that delaying the decision allows time for emotional motivations to “cool-off” so that more deliberate and less emotional decision can be taken. A cooling-off period protocol implemented in a two stage experimental game (eg: an ultimatum game or a stackelberg duopoly game) varies from 10 minutes to 24 hours. However, no consensus has been reached on the effectiveness of cooling-off periods with respect to duration and the context of the decision making environment.

In our case we implement the three minutes suspension for the purpose of invoking the endowment effect in the context of an English auction. No evidence in regard to the auction prices supports our expected treatment effect. However we do observe that a significant cooling-off effect takes place among those targeted subjects. Suppose during the suspension of the auction only those targeted subjects exhibit the cooling-off effect while all other wait-

ing subjects are unaffected. Then we should expect lower auction prices in the stop treatment than in the control treatment. However, average prices are very similar and clearly not significantly different. Thus there must be an offsetting effect that leads to non-targeted subjects bidding more aggressively after the suspension. We call this the heating-up effect. The magnitude of the two effects must be appropriately equal to lead to prices that are not different from those in the control treatment.

2.4.4 Discussion

We summarize our findings as follows: 1) The overall selling prices in the two treatments (i.e. the auction without stop and with stop) do not differ significantly and 2) subjects that held the highest bid during the suspension win the auction less often than subjects that held highest bid at the same time in the treatment without suspension. We explain this by a mixture of a cooling-off effect and a heating-up effect triggered in the targeted and not targeted subjects. Finding this mix of effects is quite surprising. The fact that we were not able to properly target a bidder with a certain rank-order statistic in terms of valuation might have led to the two effects offsetting each other over all.

The targeted subjects during the suspension of the auction are the highest bidders at that time and will have the auction object put in front of them. Moreover, they are informed privately that they can keep the box of chocolates if they win the auction. Instead of leading to an endowment effect as hypothesized, cooling-off occurs. After the three minute these subjects bid more conservatively than their counterparts in the control treatment. It is possible that the subjects use the three minutes to reconsider if they really want the box of chocolates that much. Alternatively, and more likely, they are the highest bidder because they were bidding aggressively before the suspension and the suspension breaks their action fever they might have been in.

On the other hand, the waiting subjects who do not get a box of chocolates placed in front of them bid more aggressively later on than their counterparts in the control treatment. The same suspension time serves as a heating-up period for them. Seeing others having a box of chocolate being placed in front of them seems to arouse competitive emotions for the waiting subjects. The focus of the waiting subjects must have been drawn to the fact that they are currently not bidding aggressively enough to win the auction. The idea of winning the auction makes them emotionally more competitive once the auction restarts.

2.5 Conclusion

In this paper we report an experimental approach to explore why and how the suspension during the process of an auction might change people's bidding behavior. Contrary to our initial conjecture we do not find evidence for the pseudo-endowment effect in the context of an English auction. Instead, we identify two other psychological effects initiated by the suspension of an English auction. The cooling-off effect makes the targeted subjects less aggressive bidders while the heating-up effect induces the waiting subjects to bid more aggressively. Those two opposite effects cancel each other out and lead to the same mean selling prices in auction with and without suspension.

It seems that the auctioneer can hardly profit by merely suspending. However, recall that we were not able to target the subject with a specific rank-order statistic, if an auctioneer has additional information on the ranking of valuations, then it becomes possible to use the suspension to increase the revenue if the two effects are translating to the field. Then you would want to target a bidder with a valuation that is outside of the top two, as the second highest valuation determines the selling price. Then the heating-up effect for bidders with the initially highest and second-highest valuations will increase the revenue. This does not seem unrealistic, as real estate agents often ask for private offers before they decide to

sell the house by auction. In this way they get a reasonably good signal about the relative valuations of some of the bidders.

Chapter 3

Cheap Talk Delegation Experiments

3.1 Introduction

The optimal allocation of tasks is important for achieving efficient outcomes at work places. Take a research project managed by a group of economists for instance. When a principal investigator (i.e. PI) contemplates assigning tasks within the group, she naturally wonders who is capable of what. The PI knows her own ability for sure and can perform the task by herself. However, it will not be efficient to do so if there is someone else in the group who is able to better perform the same task. More often than not, the PI would initiate an informal talk with her group members in order to get to know their abilities. Yet, on reflection the PI may also wonder about the group members' incentive. In the case of cheap talk, if messages sent do not affect payoffs, then to what extent can the PI rely on the message sent by the group members when deciding on delegation. Moreover, how much efficiency can be achieved by delegation when cheap talk is allowed?

In theory, delegation (Holmström 1977) is viewed as a tool to exploit an information advantage or a special skill possessed by an agent. The theoretical benefit for a principal from

delegation depends crucially on the alignment between his own and the agent's preferences. In the extreme case, where both the principal and the agent have the same objective, truthful reporting and optimal delegation are predicted by theory. Nonetheless, recent experimental studies (Fehr et al. 2013, Owens et al. 2014, Bartling et al. 2014) show that principals may not want to lose control of the outcome and are reluctant to delegate in spite of the theoretical benefits. Other experimental studies are concerned with other functions than efficiency that delegation can have. These include 1) serving as a commitment device (Fershtman & Gneezy 2001, Katz 1991, Kockesen & Ok 2004), 2) as a tool to shift responsibility (Hamman et al. 2010, Bartling & Fischbacher 2011), and 3) as a motivation device (Charness et al. 2012, 2016). In these previous studies, delegation refers to transferring a decision right to implement different payoffs among players. In contrast, we focus on the role of delegation to allocate the task to the most able agent and therefore achieving efficiency. We study in particular a situation where information is asymmetric and costless one way communication is allowed.

Crawford & Sobel (1982) initiate the formal study of strategic communication by providing a canonical model of cheap talk. They demonstrate how cheap talk between a sender and a receiver can be informative depending on the extent to which the players' interests are aligned. Later, Dickhaut et al. (1995) and Cai & Wang (2006) conduct laboratory experiments to test the equilibrium predictions derived from the cheap talk model. Their results support the basic intuition that less information is transmitted when preferences diverge. Moreover, Cai & Wang observe over-communication. Subjects tend to reveal more information than what theory would allow to be transmitted in the most informative of all equilibria. Possible explanations investigated in similar sender-receiver games are bounded rationality (Sánchez-Pagés & Vorsatz 2007) and lie aversion (Gneezy 2005, Hurkens & Kartik 2009, Lundquist et al. 2009).

In this paper we provide an experimental study that investigates how much efficiency can

be achieved by delegation when the agent's message about her ability is cheap talk. We introduce a novel three-person delegation game based on a real effort task. There is one principal (called delegator) and two agents (called players) in a group. Everyone has to perform the same addition task for two periods. In the first period, we incentivise subjects by a piece rate payment (i.e. a fixed payment per correct answer). Every subject only learns her own performance. Before the repetition, the delegator can decide if he wants to replace his past performance by the future performance of one of the players. In order to help him to decide, the players can send messages about their past performances, which are costless. Everybody is paid the same piece rate for their performance. If the delegator delegates, then his first-period earnings are calculated based on the chosen player's second round performance. In order to misalign the preferences between principal and agent, delegators are required to pay a fee to the chosen player if they delegate. This in theory rules out any information transmission in equilibrium. Players have an incentive to exaggerate their past performance in order to improve their chance to be chosen. Our main research question is whether information transmission is still possible when the game is played by real humans despite the fact that our game only allows for babbling equilibria. Moreover, we investigate how the coarseness and the size of the message space influences the degree of information transmission.

Our experimental treatment variation consists of changing the precision of messages that can be sent by the players. As baseline treatment we adopt a message space in which players can send any number they want to represent the number of solved problems in the first adding-up task (i.e. Precision Message Treatment, PMT). Among other things, this message space allows us to quantify the degree to which players exaggerate their performance. In line with the research on lie aversion (Fischbacher & Föllmi-Heusi 2013), we also observe a surprisingly large portion of truthful reports (almost 50%) in spite of the incentive to lie. When subjects lie in our experiment, they on average exaggerate their performance in a modest way. The average monetary value of a lie, calculated as the number of tasks times

the piece rate, is about as large as the bonus a player receives if chosen. The modest lying implies that messages contain information that potentially can be extracted by the delegator. However, the information contained in the messages is not picked up by the delegators. Delegation is driven less by the messages than by the delegator's own performance and his subjective belief on his relative ability ranking compared to all subjects. In order to investigate if the specific message space hinders information transmission, we introduce two treatments, one with a coarser message space and one with an infinitely rich message space.

In one treatment we introduce noise by partitioning the message space into intervals of a fixed length to make the message space coarser (i.e. henceforth, Fuzzy Message Treatment, FMT). Intuition would suggest that communication with noise makes information transmission even more difficult. Unexpectedly, we find that players send messages in a similar fashion as in the PMT. Most often, we observe either truthful reports or moderate lies. Even more surprisingly, in contrast to the baseline treatment, delegators make good use of the information and delegation frequencies increase with favorable messages received. In addition, we are able to identify other-regarding preference beyond lie-aversion by making use of our interval structure. Given the same degree of exaggeration in the messages, a player whose performance sits on the left end of a message category can cause greater harm to the payoff of the delegators than a player whose performance is located on the right end. We find that players whose performances are located on the left end of an interval tend to lie less frequently about their performance. If they lie then the size of the lie is also smaller.

In the two treatments discussed so far, the message space was either of the same or lower cardinality than the type space. In our final treatment, we employ a very large message space by allowing free-text messages to be sent (i.e. Free Communication Treatment, FCT). We find that free-text messages lead to the highest level of efficiency and the most information being transmitted. After analyzing messages that led to delegation and messages

that did not, we find that truthful messages that indicate performance differ in style and structure from messages that either contain lies or are talking about other things than past performance. This allows delegators to better judge the information content of messages.

An additional concern of this study is the gender gap in demand for leadership as documented by Reuben et al. (2012) and Reuben et al. (2014). These papers find that the prevailing stereotypes deem male to be more capable of solving math problems and thus males are more likely than females to be chosen to perform tasks for a group. In order to subtly add the gender element in our context, we introduce avatars that differ both in gender and ethnic background. At the beginning of the experiment subjects are required to choose one avatar that best represents themselves. Further, our matching ensures that of the two subjects that can be chosen by the delegator one is male and one is female. In this way a delegator can tell the gender of the players who sent the message by seeing their avatars. While we observe that significantly more male subjects are chosen in the Precise Message Treatment, we can not identify whether the observed difference is coming from the demand side or supply side. Earlier experimental research (Croson & Gneezy 2009) on gender differences in preferences has pointed out that women are reluctant to initiate negotiations or enter competitions.

To our best knowledge we are the first to study delegation with cheap talk in a realistic performance context. The novel setting allows us to focus on how cheap talk can facilitate delegation to achieve efficiency. Most closely related to our study Dessein (2002) and Lai & Lim (2012) investigate the trade-off between a loss of control under delegation and a loss of information under communication. In their setting both delegation and cheap talk are used to exploit the private information held by agents.

3.2 Design of Experiments

Our cheap talk delegation game is built on a prototypical real-effort task of adding up five two-digit numbers, which was introduced to the literature by Niederle & Vesterlund (2007).¹ They provide evidence of a gender gap in preferences for competition. Conditional on performance, women choose the competitive payment scheme less often than men. Our setting provides a complementary setting to study the gender gap by looking at the demand side (i.e. whether one gender is more likely to be selected for the task). We adopt a piece rate payment scheme to incentivise subjects to solve as many questions as possible in seven minutes. All subjects perform the adding-up task first and then repeat the same task with the same piece rate payment. For the delegation game subjects are randomly assigned to groups of three, with the restriction that each group will have at least one male and female player.

3.2.1 Cheap Talk Delegation Game

We call the adding-up task that has been performed before the delegation game *Task 1* and the repetition that will occur later *Task 2*. There are one delegator and two players in a group. A delegator can delegate the adding-up task to one of the players who will then perform the task on his behalf. Delegators will have to perform the delegation task regardless if they delegate or not. Suppose delegators do not need to perform the task if they delegate. Having nothing to do and being bored could be an incentive for not to delegate. Therefore, in our design delegation of the task will result in a replacement of the past performance of the delegator by the future performance of the chosen player. This has also the advantage that when deciding to delegate the delegators own performance is fixed and beliefs about his own future performance are irrelevant.

¹There is an active experimental research line in the gender gap at work place from the supply side recent years. An interesting session discussion in 2016 can be found at AEA webcasts.

Individual performances in *Task 1* are private information. A player can signal his or her performance in *Task 1* by a costless message to the delegator. When delegation happens, a bonus is transferred from the delegator to the chosen player. The preferences of the delegators and the players are perfectly misaligned, as delegation induces a payment from the delegator to the chosen player. In theory only the babble equilibria exist in such a setting, if we assume that players are rational and have the objective to maximize their monetary payoff. In equilibrium messages are disregarded by the delegators and decisions are only based on their prior beliefs.

However, if subjects exhibit either social preferences, bounded rationality or aversion to lying, then information transmission might occur. Thus we investigate if information transmission is possible when the game is played by real humans. In addition, we wonder if the bonus can trigger positive reciprocity between the delegator and the chosen player such as in the gift-exchange literature (e.g.:Berg et al. (1995), Bolton & Ockenfels (2000), Falk & Fischbacher (2006) and Dufwenberg & Kirchsteiger (2004)). Since the bonus is coming from the delegator rather than from an outsider party, the chosen player who interprets the bonus as a signal of trust might reciprocate with a higher effort and therefore a better performance.

3.2.2 Gender Effect

In order to activate innate gender differences we adopt an experimental technique, called priming of natural social identities, which has been used extensively in social psychology. Research in the area has found that making different natural social identities salient subtly through priming can impact behavior and outcomes, such as test performance (Aronson et al. 1998) or subjects' perceptions (Bargh & Pietromonaco 1982). We use a set of questions as a stimulus to sensitize subjects' perception of gender differences. The questionnaire includes questions about general preferences over food, role models, life style, and so on.²

²The items used for priming were: Who are your role models? Please list some of your favorite food. What kind of movie you prefer to watch for a casual weekend? Please describe your ideal way of relaxation. What is

Male and female tend to have different answers to our questions either due to nature or culture. Priming works through the process of answering these questions and therefore reinforces the innate gender differences.

After the questionnaire stage, avatars, representing combinations of the two genders and four ethnic backgrounds, are shown on the screen. Subjects are required to choose one that best represent themselves without knowing what they will be used for. Groups are formed randomly with the restriction that they at least contain one female and male subject each. One female and one male subject are chosen to become Player 1 and Player 2, while the remaining subject (male or female) is given the role of delegator. This procedure ensures that there are male and female delegators that have the choice to delegate to either a male or a female player.

3.2.3 Treatments and Parameter Values

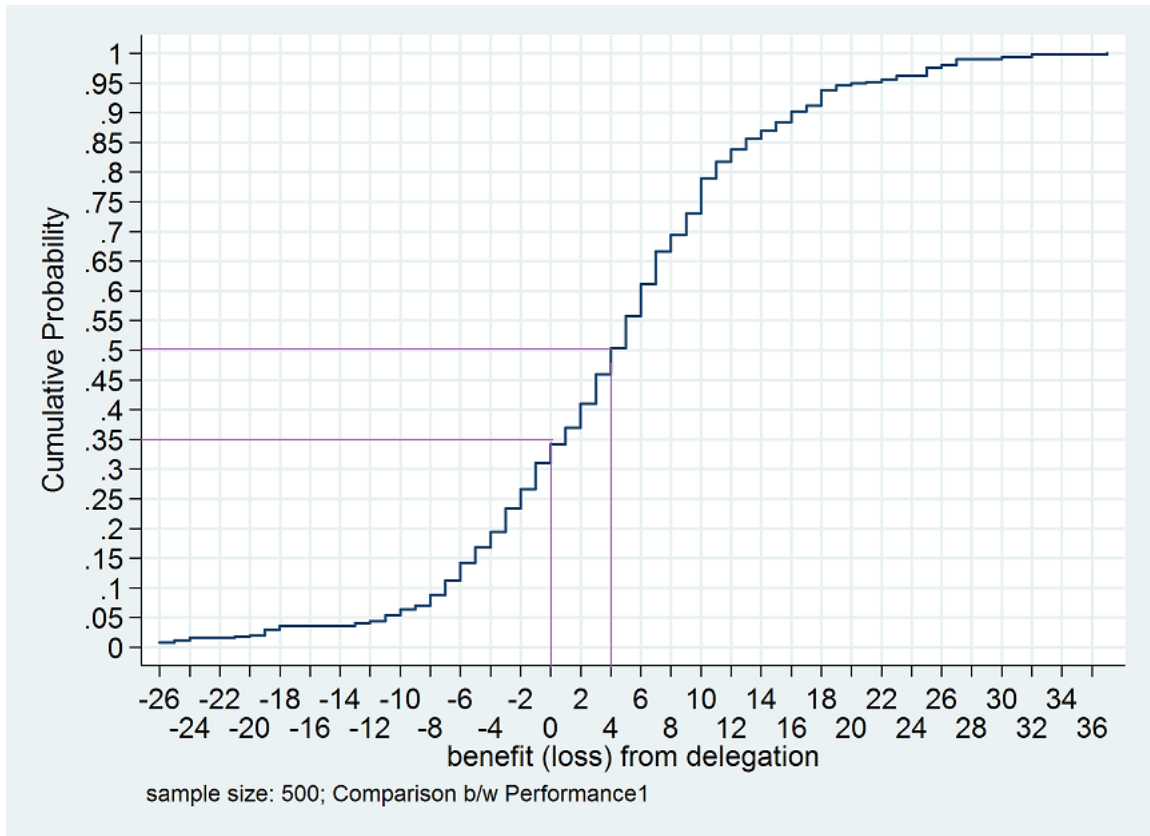
In order to determine a reasonable size of the bonus, we use some data on performance in the adding-up task that were generated in a pilot study. The size of the bonus determines the fraction of cases in which an informed delegator would find it individually rational to delegate. The higher the bonus the lower the fraction as the performance of the chosen player has to be high enough to compensate the delegator for paying the bonus. To keep things interesting the fraction should not be too small. On the other hand, we would like to have a sizable bonus causing a sufficiently strong incentive to lie. In the end we settled on a targeted fraction of rational delegation of 50%.

We draw from the performance data gathered in a pilot experiment with replacement and form 500 hypothetical groups. In these groups we calculate the maximum net gain delegators can make if they have public information about the performance and delegate to the

the major responsibility in your current life? What is your childhood dream? Please name some achievements that you are proud of.

best player. Figure 3.2.1 plots the resulting cumulative distribution of the potential gain or loss from delegation.

Figure 3.2.1: Distribution of potential gain or loss from delegation



If there is no cost for delegation, 65 percent of rational delegators can benefit from delegations. The figure reveals that for our target of 50 percent of cases being profitable, a bonus equivalent to the value of four questions should be imposed.

The targeted average earnings for our subjects was 20 Dollars per hour. With this target we arrived at a piece rate of one Dollar per correctly solved question and a bonus of four Dollars. One of the tasks was randomly chosen for payment.

In our experiments we implement three treatments. In the base line treatment (i.e. Precise

Message Treatment, PMT) players enter a number indicating their past performance as a message to the delegator. The message space in the PMT consists of non-negative integers. Then we vary the message space in two different directions with respect to size. In the Fuzzy Message Treatment (i.e. FMT) subjects could only indicate an interval in which their claimed performance falls. Here the message space is not rich enough to send a precise message. Our treatment variation into the other direction makes the messages space richer than what is required for precisely messaging the past performance. In the Free Communication Treatment (i.e. FCT) we allow players to send free-text messages.

In the FMT, a set of intervals are provided to players, from which they can choose one to indicate the range of their past performances. An important question is how to decide the interval length and position compared to the delegator's own performance. In the PMT delegators are able to decide whether delegation is profitable or not if they believe that messages are true. In order to keep such a property in the FMT, we locate the performance of a delegator on the right end point of an interval and set the length of intervals to be the size of the bonus (i.e. each interval contains 4 numbers). Thus delegation would be beneficial if the chosen player's message is true and at least one interval higher than the delegator's performance. A further advantage of this procedure is that the location of the own performance in the reporting intervals could be excluded as a reason for behavioral differences across delegators.

The restricted message space adopted in the FMT to a certain degree mirrors real life situations where performance is not objective and cannot be communicated precisely. However, in our experiments no mechanisms such as law or reputation exist to incentivise information revelation and therefore little can be told from a highly structured message. In the real world people can communicate in real language. Using real language allows for more than just sending a message that communicates a number. Meta communication is possible, which might be important in real life, while in standard game theory message spaces that

are at least as large as the type space are sufficient for maximum information transmission if the preferences are aligned. The FMT allows us to test if the theoretically superfluous size of the message space created by 400 characters free text helps or hinders information transmission.

3.2.4 Experimental Procedure

The cheap talk delegation game was programmed using Z-tree (Fischbacher 2007) and carried out at Adlab (i.e. Adelaide University Laboratory for Experimental Economics Web-page). Subjects were recruited with help of the online system ORSEE (Greiner 2015). Overall, 342 subjects, which were predominantly university students, participated the experiments. In our between-subject design every participant could only take part in one session. Subjects were paid in private at the end of the session according to their performance and earned on average 16 Dollars.

The experiments consisted of two parts. In the first part of the experiment subjects were required to complete a few simple tasks without any interaction with other participants. Instructions for these tasks were shown on the screen. Subjects did not know the content of the subsequent tasks at that time. Tasks in the first part of the experiment were: 1) answering the priming questionnaire; 2) choosing an avatars; 3) completing an adding up task, and 4) answering two belief elicitation questions. We deliberately withheld the information on the second task to remove strategic motives when selecting avatars.

Subjects sum up five two digit numbers in the adding up task. They were not allowed to use a calculator but could use scrap paper. The numbers to be added were randomly drawn in advance and two sets of 50 questions were generated for *Task 1* and *2* in each session.

On the screen, the five numbers to be added appear in a box, and the subject has an input box to submit an answer. Once a subject enters the answer and clicks the submit button, a new

problem appears together with feedback on whether the former answer was correct or false. A record of the number of correct and false answers is kept and displayed on the screen. Subjects have seven minutes to solve as many problems as they can. Their performance in the task is equal to the number of correctly solved problems.

In the belief elicitation stage we asked the following two questions: 1) What do you think, how was your relative performance in the adding-up task? Please indicate the percentage (0 to 100%) of other participants that solved more questions than you. 2) Suppose we repeat the same task and you try as hard as possible. How many correct questions can you achieve? Subjects were reminded of their actual performance in *Task 1* to rule out the concern of memory errors. We incentivised the first question by giving 5 Dollars to the person in the session closest to the real rank. The first question elicits subjects' beliefs about their relative performance. The second question elicits the expectation of subjects about their future performance.

In the second part of the experiment subjects received new written instructions about the cheap talk delegation game and the sequence of play. After the computer determined the groups and subjects received instructions, the players that might be chosen as delegates were asked if they would be happy to be considered for delegation. Again, the available players were reminded of their actual performance in *Task 1* to avoid memory lapses. Then they were asked to send a message about their past performance to the delegator. Depending on the treatment the message could be an integer, an interval or free text. Delegators had to make a delegation decision. Based on the messages and avatars of the players who decided to be available for delegation, the delegator had to decide if and if then to whom to delegate. Next, subjects were informed about the results of the delegation round and had to complete *Task 2* (i.e. the same adding-up task as before). Finally, the computer randomly chose either to pay *Task 1* or *Task 2*. Note that if *Task 1* was chosen and the delegator actually delegated, then the delegator was paid according to the performance of the delegatee and the bonus

was transferred to the delegatee. By the end of the experiment we conducted a routine questionnaire to collect basic information such as age, gender, math training and so on.

3.3 Results

In this section we start by presenting basic descriptive statistics from the first part of the experiment. Next, we focus on the efficiency of delegation by comparing across treatments. We find that both the FMT and the FCT obtain a significantly higher efficiency than the baseline treatment. In order to find the driving force of the observed treatment effect, we further look into the behavioral differences of players and delegators across treatments to investigate if and how information is transmitted.

3.3.1 Basic Data Summary

We ran five sessions for each treatment with between 15 and 27 subjects per session. Overall, we had 120 subjects in the PMT, 117 in the FMT, and 105 in the FCT. In each treatment, we required that each group to contain at least one male avatar and one female avatar. However, six female subjects and nine male subjects misrepresented their gender by choosing an opposite gender avatar. In this session we use the gender information collected in the routine questionnaire to check the gender gap.

3.3.1.1 Performance in the Adding-up Task

The adding-up task chosen requires both skill and effort, which allows for enough variance in performances across the population. Following Niederle & Vesterlund (2007), we anticipate no gender difference in performances under the piece rate payment scheme. A

pairwise rank-sum test across treatments confirms that there is no significant difference between genders. Table 3.3.1 reports the average number of correctly solved problems by gender using the pooled data.

Table 3.3.1: Summary of Performance in the Adding-up Task

Performance	Male	Female
In <i>Task 1</i>	13.68 (6.08)	12.51 (5.32)
2-sided rank-sum	p=0.1273	
In <i>Task 2</i>	16.80 (7.30)	15.12 (6.05)
2-sided rank-sum	p=0.1190	

Note: Average with standard deviation in parentheses.

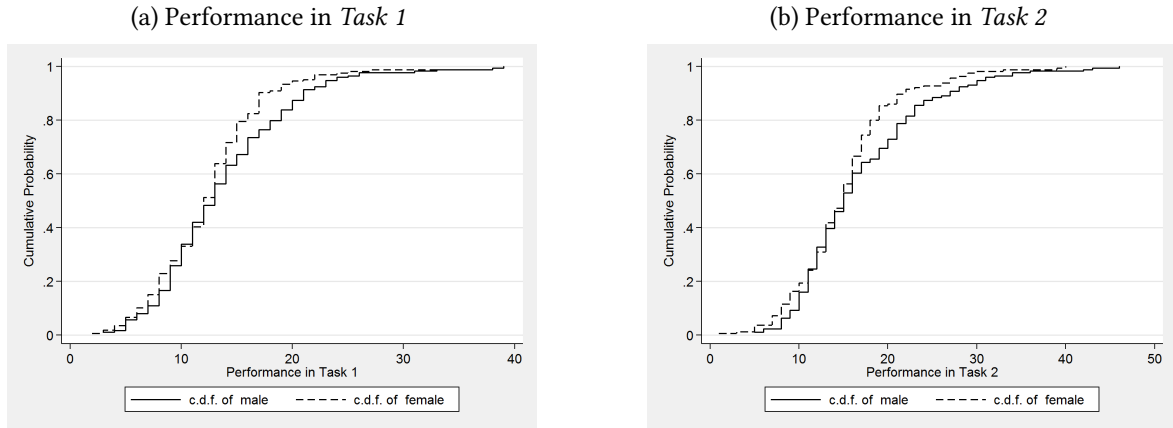
Sample size is 340 in *Task 1* and 339 in *Task 2*. We have removed two outliers who reached the 50th question from both tasks and one extra outlier who performed 0 in *Task 2*.

We observe one outlier in each treatment. One female subject in the FCT and one male subject in the FMT both reached the upper bound of the 50th question.³ One female subject in the PMT did not perform *Task 2* at all.⁴ Figure 3.3.1 draws the cumulative distribution function (CDF) of the number of correct answers in *Task 1* and *Task 2* respectively omitting the three outliers.

³Once they had completed the 50th question, there were no more new questions.

⁴Her role was Player 1 and she did not want to be chosen for delegation. We suspect that she misunderstood the instructions.

Figure 3.3.1: CDF of Number of Correctly Solved Questions by Gender



The CDF shows the proportion of subjects who can correctly solve a given number of questions or fewer at every level of performance. We test for a difference in distribution between male and female subjects for each task. The Kolmogorov-Smirnov test (exact) supports a significant gender difference in distribution of performances *in Task 2* ($p=0.024$) but not in *Task 1* ($p=0.065$).

3.3.1.2 Improvements in the Real Effort Task

We find that subjects on average perform two to three questions better in *Task 2* than in *Task 1* (one-sided $p<0.01$ with a pairwise t-test) despite the same piece rate payment.⁵ The improvement of the average performance does not differ significantly by gender ($p=0.4803$ with a two-sided rank-sum test and $p=0.5178$ with a two-sided t-test). The improvement may be caused either by learning or by the outcome of delegation. Before the repetition of the adding-up task, subjects were informed whether or not they were chosen for delegation. The chosen players were informed that a four Dollar bonus would be paid to them from the delegator and that the delegator’s payment from *Task 1* would depend on their performances

⁵We further remove two outliers whose improvement is beyond 20 questions. We suspect that they have employed a calculator without notice.

in *Task 2*. The non-chosen players were informed that they were not chosen by the delegator and that they would be paid the same piece rate for *Task 2*. Therefore, we can categorize the outcome of delegation into the following four cases: delegators who either delegated, i.e. 1) Delegators_D, or not, i.e. 2) Delegators_ND, and players who either were chosen, i.e. 3) Players_C, or not, i.e. 4) Players_NC. Table 3.3.2 gives a summary of performances using the pooled data conditional on the four outcome categories.

Table 3.3.2: Performance in the Adding-up Task by Roles and Delegation Outcomes

Outcome Categories	1) Delegators_ND	2) Delegators_D	3) Players_C	4) Players_NC
Performance in <i>T1</i>	9.92 (3.88)	14.58 (5.96)	15.89 (4.59)	12.57 (5.81)
Performance in <i>T2</i>	13.33 (5.08)	17.23 (6.74)	16.91 (5.06)	15.52 (6.68)
Improvement	3.42 (3.12)	2.65 (3.64)	1.03 (2.54)	2.96 (3.17)
Observation	36	77	35	189

Note: Average with standard deviation in parentheses.

[1] Improvement= Performance in *Task 2* - Performance in *Task 1*.

[2] We removed 5 outliers as mentioned before.

Delegation seems beneficial as the average performance of the chosen players in *Task 2* is seven questions better than the delegators' average performance in *Task 1*. The improvement in case 1), 2), and 4) have no significant difference with a pairwise rank sum test. Unexpectedly, the improvement of the chosen player is significantly ($p < 0.001$) lower compared to the case 1), 2) and 4). This finding is consistent with the idea that the better one already is the harder is to improve, since the chosen players on average have the highest performance in *Task 1*.

We further run an OLS regression of improvements on the four outcome categories controlling for individual performances in *Task 1*. We also include individual characteristic dummies such as gender, age, and high school math training that may correlate with individual performances. Table 3.3.3 reports the estimated coefficients and their standard errors. Compared to the improvements made by delegators who did not delegate, the chosen players

on average made significantly smaller improvements conditional on the same performance level in *Task 1*.

Table 3.3.3: OLS Regression of Improvements

	Coefficient	Standard Error
Performance in <i>Task 1</i>	-0.027	0.032
Player_C	-1.787***	0.655
Player_NC	0.278	0.436
Delegator_D	0.769	0.661
Gender (Female)	-0.216	0.353
Math training (Yes)	1.169***	0.379
Age (Below 30)	-0.130	0.519
Constant	2.484***	0.785

Note: ***,** denote significance at $p=0.01$ and 0.05 respectively.

[1] Sample size is 337 excluding five outliers.

[2] Adjusted R-squared = 0.0454.

If the bonus is viewed as a gift from the delegator to the chosen player, we would expect positive reciprocity in the form of increased effort. Charness (2004), an early gift exchange experimental study, has documented that agents provide more effort to higher wage offers from the employer. Recent field experiments suggest that gift exchange is sensitive to how the “gift” is presented. Workers reciprocate to gifts which show an employer’s effort and concern significantly stronger than a simple pay raise (Kube et al. 2012). An unexpected gift such as an increase of salary as a surprise rather than paying above-market wages leads to higher productivity (Gilchrist et al. 2016).

Possible explanations for the observed negative effect of improvement here could be either an income effect or an adverse reaction to additional pressure. Given the four Dollar payment, the chosen players may lose the drive to work harder. In this case, the bonus is interpreted as the payment deserved rather than a gift. Alternatively, the chosen players may feel more responsibility when they perform *Task 2* as the payment of the delegator depends on their performance. However, the sense of a greater responsibility can bring more

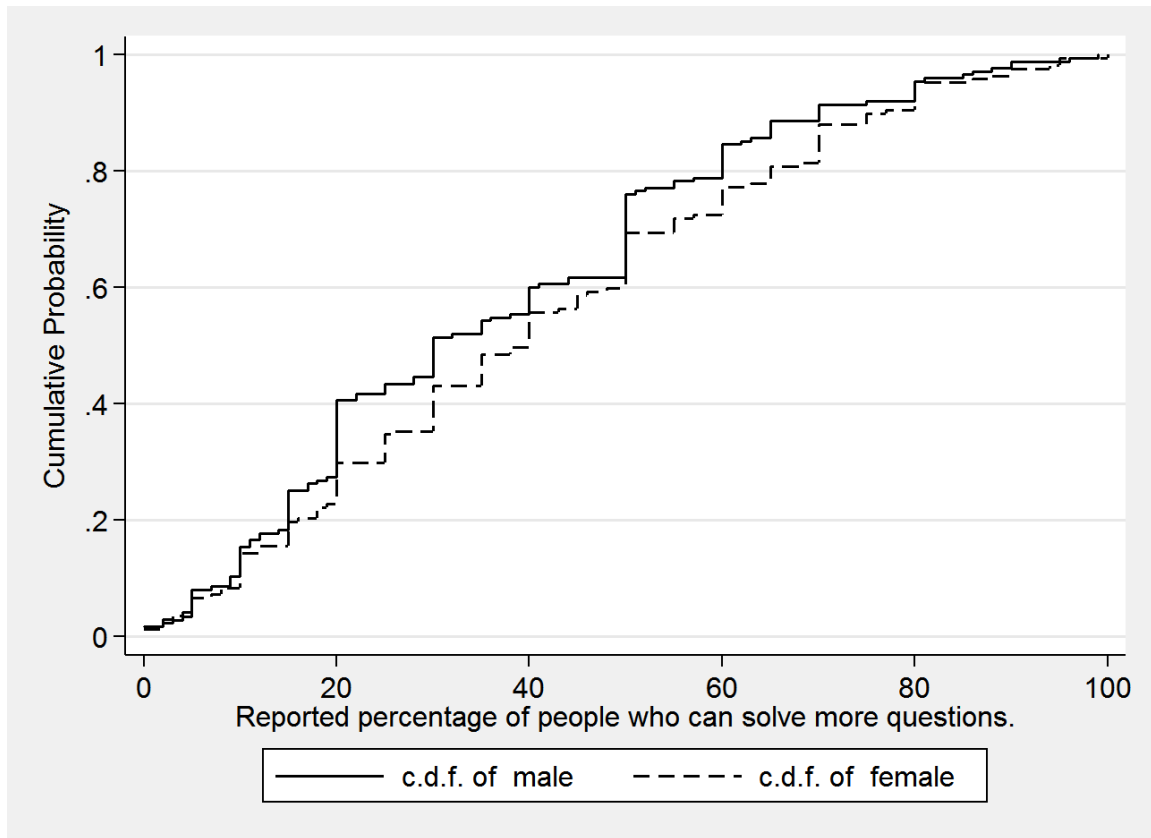
pressure which may adversely affect their performance in an real-effort task.⁶

3.3.1.3 Is There a Gender Gap in Confidence Level?

Subjects were reminded of their performance in the completed *Task 1* when they answered the two hypothetical questions about their relative performance and expected future performance. We further incentivised an objective guess of the relative performance by a five Dollar payment. Therefore, subjects' answer to the question of the proportion of people who can solve more questions reveals their relative confidence level. The larger the percentage reported, the less confident subjects are of their relative performance.

⁶We must be cautious here due to the regression toward the mean effect. Statistically speaking, some one who by chance has performed extremely well the first time is likely to perform to worse the next time. However, our regression shows that this effect is unlikely to be the driver, as past performance dose not have a significant impact on the improvement.

Figure 3.3.2: Cumulative Distribution of Self-reported Percentage Beliefs by Gender



Note: [1] Sample size is 342 (i.e. pooled data across three treatments).

As shown in Figure 3.3.2, the distribution of self-reported percentages of females first-order stochastically dominates that of males. For any fixed proportion, male subjects reported a lower percentage level than female subjects did. Women on average think 42.47% of the population outperform themselves, while men only think 36.58% of the population. However, both the distribution difference and the average difference is only weakly significant ($p=0.094$ with a one-sided Kolmogorov-Smirnov test and $p=0.051$ with a two-tailed t-test). The gender gap in relative confidence vanishes if we further control for performance using an OLS regression, i.e. see Table 3.3.4.

Table 3.3.4: An OLS Regression of the Elicited Beliefs

	Coefficient	Standard Error
Performance in <i>Task 1</i>	-1.059***	0.198
Gender (female)	3.438	2.553
Math training (Yes)	-3.224	2.752
Age (Young)	-8.435**	3.787

Note: ***, ** denote significance at $p=0.01$ and 0.05 respectively.

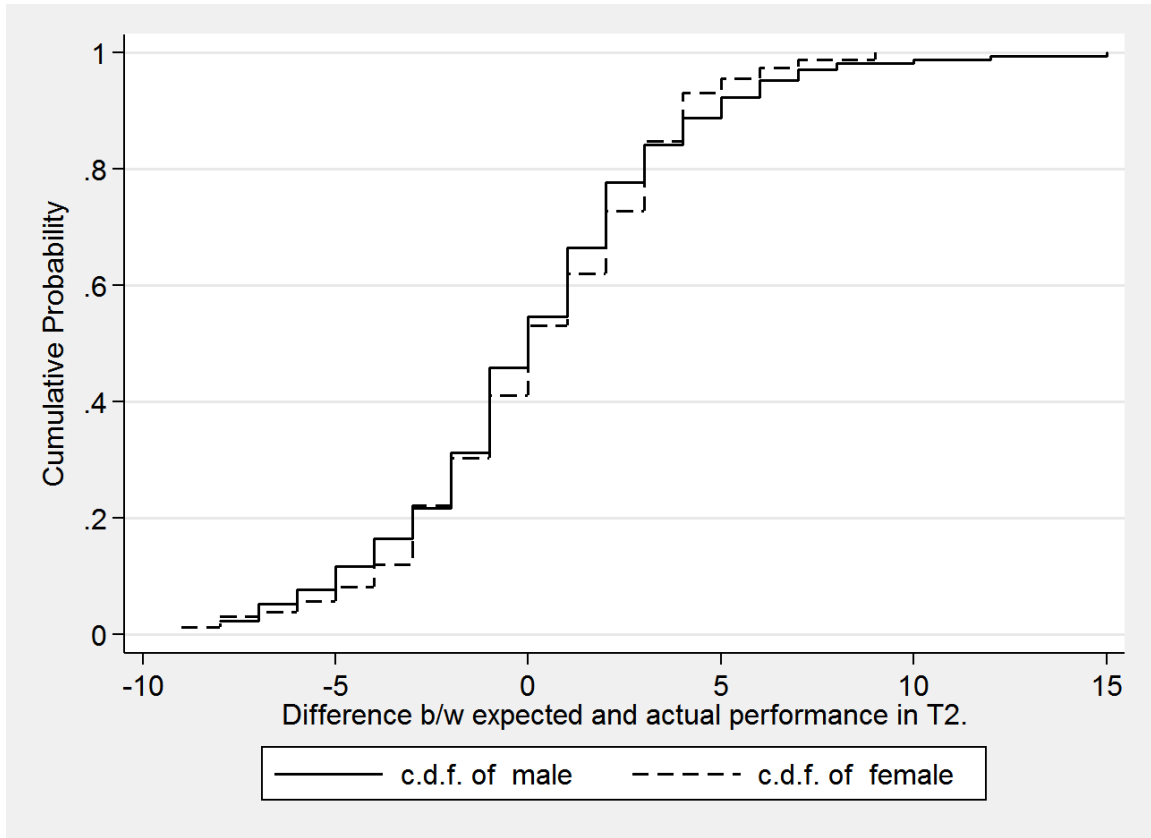
[1] Sample size is 342 (pooled data across treatments).

[2] Dependent variable is the guess rank among all subjects. We control for 1) performance in *Task 1*, 2) gender (1-female; 0-male), 3) math training (1-yes; 0-no) and 4) age (1-below 30 and 0-above 30).

Next, we construct an over-confidence measure by taking the difference between the self-reported expectation of future performance and the realized performance in *Task 2*. Figure 3.3.3 plots the cumulative distribution function (CDF) of over-confidence by gender. As one can see that the two distributions almost overlap. Both the two-sided K-S test ($p=0.432$) and the two-sided t-test ($p=0.153$) confirm that there is no gender difference in terms of over-confidence.⁷

⁷Sample size is 341 omitting one outlier who did not perform *Task 2*.

Figure 3.3.3: Cumulative Distribution of Over-confidence by Gender



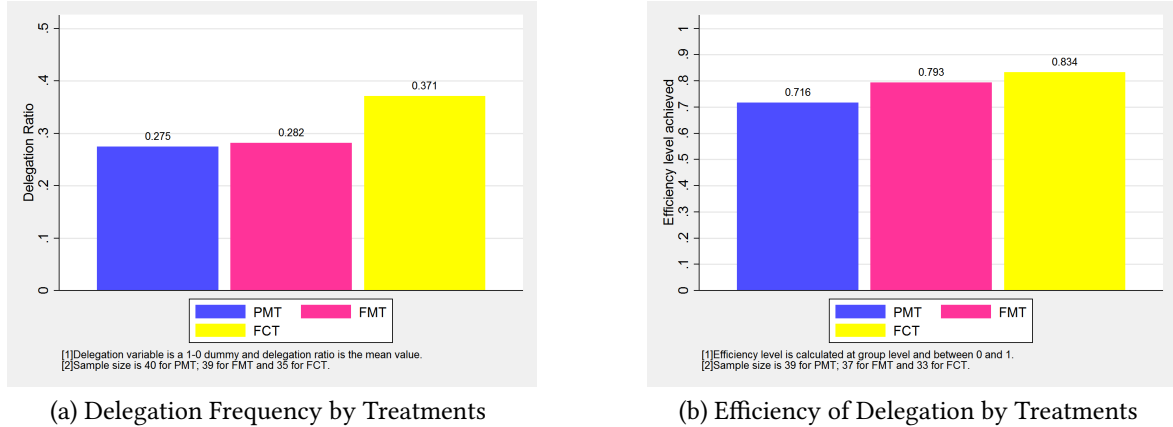
Note: We exclude 7 outliers (i.e. the difference is more than 20) to have a reasonable scale in horizontal line.

3.3.2 Delegation Frequency and Efficiency of Delegation

The observed delegation frequency in our experiments sits in between the two extremes. To one end we calibrated the four Dollar bonus to target approximately 50% of rational delegations when individual performances are public information. To the other end we would expect no delegation at all in theory according to the babbling equilibria. We calculated the actual delegation ratio by dividing the number of groups in which delegation occurs to the total number of groups for each treatment. As shown in Figure 3.3.5a, around 30% of delegators delegate within each treatment. A pairwise proportion test confirms no significant difference across treatments ($p > 0.3$, two-sided). Thus, different forms of a message have no

significant impact on the average frequency of delegation.

Figure 3.3.4: Comparisons across Treatments



Next, we check the treatment effect with respect to the efficiency of delegation. Recall that to delegate means to replace the delegator’s performance in *Task 1* with the chosen player’s performance in *Task 2* for calculating the delegator’s payout. In each group, delegation can improve social welfare only if there is some player whose performance in *Task 2* is better than the delegator’s in *Task 1*. If the delegator is able to allocate the task to the most capable person, then we say that delegation is fully efficient. We normalize the efficiency measure by dividing a delegator’s actual choice of performance to the best performance within a group (i.e. the best performance among the delegator’s in *Task 1* and the two players’ in *Task 2*). In this way, we construct an efficiency ratio to measure to what degree the actual delegation choice leads to the efficient outcomes. Figure 3.3.5b shows the average actual efficiency ratio across treatments.⁸

Interestingly, the efficiency ratio level in both the FCT and the FMT is higher than in the PMT. Compared to the baseline treatment, the difference is significant in the FCT ($p=0.019$, two-sided t-test and $p=0.025$, two-sided rank-sum test) and weakly significant in the FMT ($p=0.0737$, one-sided t-test and $p=0.0584$ one-sided rank-sum). Although it seems that the

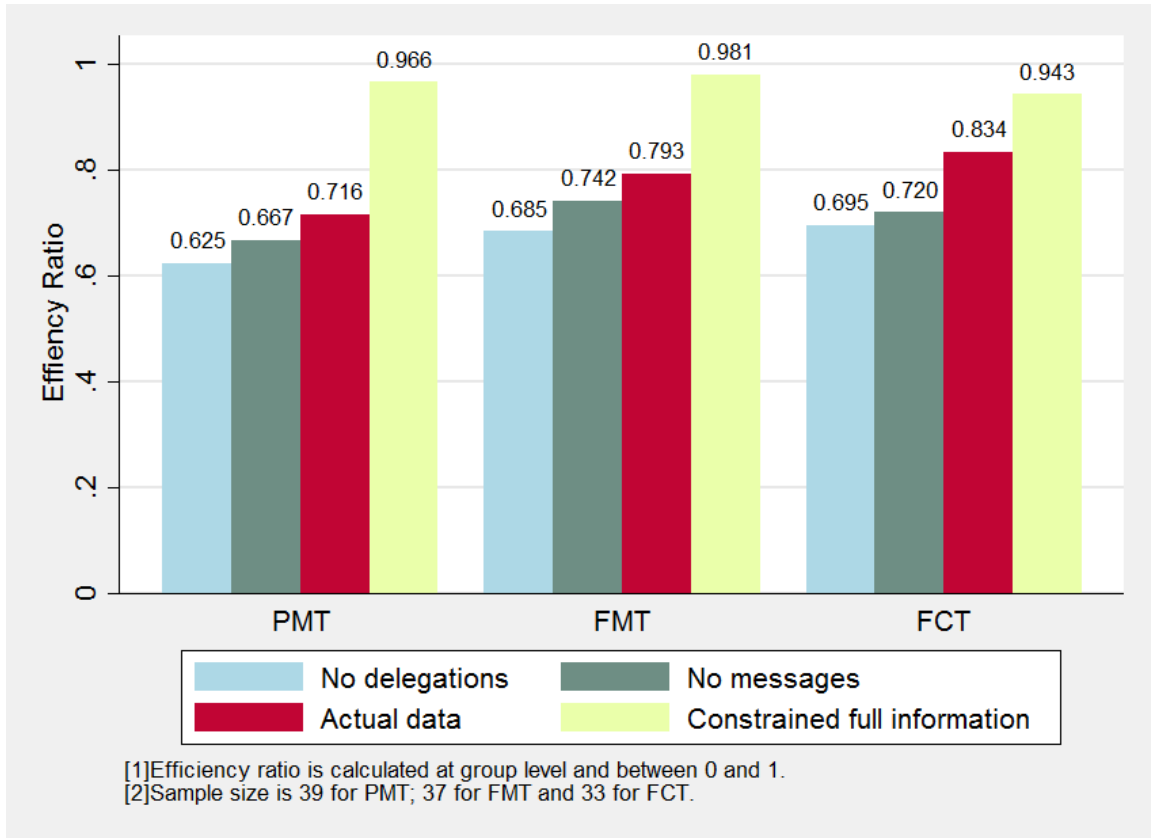
⁸We excluded five groups that contain outliers of extreme value in individual performances.

free form communication treatment leads to the most efficient allocations, there is no significant difference between the FMT and the FCT.

In what follows, we conduct three counterfactual exercises to compare efficiency within each treatment. The maximum efficiency is achieved when there is no hidden information on individual performances. Assume that information on individual performances are commonly known. Delegation leads to a Pareto improvement as long as the chosen player's performance is greater than the delegator's performance plus the bonus. Therefore, rational delegation takes place whenever there is a player in a group who can perform at least four questions (i.e. the size of the bonus) better than the delegator. We call the scenario the constrained full information and calculate the maximum efficiency ratio. Then we assume that individual performances are private information and shut down the communication channel to examine what efficiency level can be attained with and without a delegation option.

When the delegation option is not allowed, delegators have to use their own performance in *Task 1*. The no delegation scenario gives us the minimum efficiency level. In the no message scenario delegators can only resort to their initial beliefs about who would perform best to decide whether to delegate or not. More precisely, we need to know subjects' beliefs about the proportion of people who can solve at least four more questions. However, the elicited belief is about the percentage of people who can solve at least one more question. In this case we can calculate the upper bound for the efficiency level by assuming that rational delegation happens when delegators believe that more than half of the people can perform better than themselves. We further employ a random delegation strategy, i.e. with equal chance to use the performance of both players, to simplify the calculation. Figure 3.3.5 summarizes the average efficiency ratio calculated for 1) the actual experiment outcome, 2) the constrained full information scenario, 3) the no delegation scenario, and 4) the no message scenario within each treatment.

Figure 3.3.5: Average Efficiency Level in Four Scenarios



The ranking of efficiency level in different scenarios is the same across treatments. The actual outcomes in our experiments is highlighted in red. Compared to the situation where both cheap talk and delegation are banned, our actual data shows a higher efficiency level ($p < 0.01$, sign-rank, two-sided) in all treatments. Compared to the no message scenario, only the FCT obtains a significantly higher efficiency ($p < 0.01$, sign-rank, two-sided).⁹ Thus, we conjecture that information must be transmitted in the FCT as delegators were able to do better than a random delegation strategy. It is no surprise that compared to the constrained full information scenario, the actual efficiency level is significantly ($p < 0.01$, sign-rank, two-sided) lower due to private information. It seems that cheap talk in our experiments can alleviate the issue of asymmetric information to some degree but is not able to resolve the

⁹P-value is 0.2399 in the PMT and 0.1924 in the FMT

issue. In what follows we will have a closer look at individual behavior within each treatment to explain the treatment effect in efficiency.

3.3.3 Individual Behavior of Delegates

Players were allowed to express their interest for being chosen for delegation in a pre-message stage. They were asked, “Do you want to be considered as the person to whom the delegator will transfer some responsibility”. 21 (15 females and 6 males) out of 228 subjects declined the chance to take some responsibility for others in spite of the reward of the bonus. The majority of the unwilling performed below average (i.e. 14 correct answers). In some sessions we asked these subjects for the reason to get an idea about the motives for the unwillingness to take responsibility. Based on the answers we conjecture that for the under-performers a four Dollar bonus is not sufficient to counter the large discomfort from lying. The reason for the high-performers was that they did not want anybody else to benefit from their skill and effort.

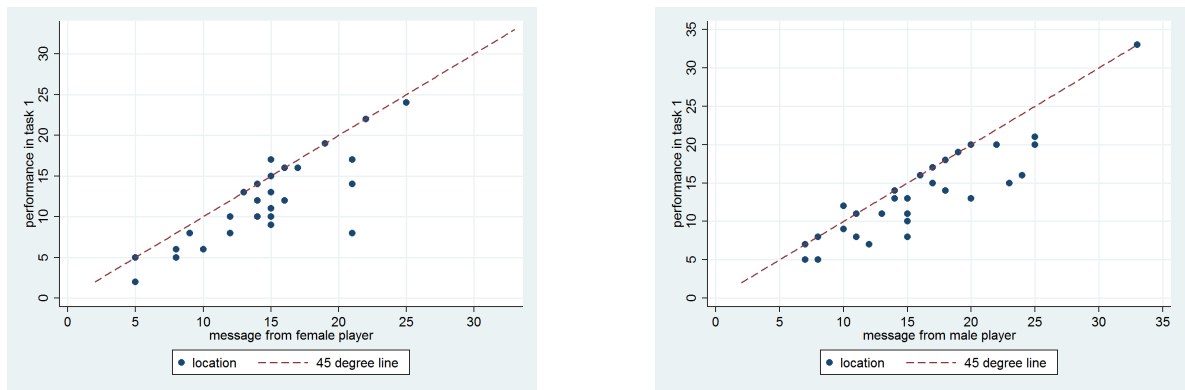
The remaining 207 players who were available for the delegation game entered the cheap talk stage. We kept the size of the bonus the same but changed message format across treatments. Players were reminded of how many correct questions they have solved and the fact that their actual performance is unknown to the delegator. If players are only motivated by the monetary reward, i.e. the bonus, we expect them to lie as much as possible to differentiate themselves from the other player in the group.

In the PMT, we constructed the message space such that players can only choose a number to represent how many questions they have correctly solved in *Task 1*. The simple message structure allows us to explore how many and to what degree players would exaggerate their past performance. Interestingly, we observe a large portion of truth-telling and a moderate degree of lies. The size of lies, on average, is smaller than the size of the bonus. We varied

the coarseness of the message space in order to investigate which kind of messages improve delegation outcomes. In the FMT, we added noise by partitioning the message space into intervals. In the FCT, we made the message space richer by allowing free-text communication.

3.3.3.1 How Players Message in the PMT?

In the PMT, approximately 40% of players truthfully reported their past performance and approximately 40% of players exaggerated their past performance by up to 4 questions. Figure 3.3.6 plots the players' messaged performance against their actual past performance by gender. A dashed 45 degree line indicates the truthful messages, and the area below the 45 degree line indicates lies. The further away a dot is from the 45 degree line the larger the lie.



(a) Graph of Female Players

(b) Graph of Male Players

Figure 3.3.6: Scatter-plot of Messages against Actual Performance in *Task 1*

The majority of dots locate either along or below the 45 degree line, which suggest either a truthful report or an upward lie.¹⁰ It seems that men lie slightly more aggressively than women. However, the small gender difference in the size of a lie is not significant ($p=0.339$ with a rank-sum test and $p=0.272$ with a one-sided t-test). About 54% of males exaggerated

¹⁰Strangely, there are two subjects who lie downward.

their past performance while the proportion of females was 64%. Again, the difference is not significant with a two-sided test of proportions.

An OLS regression reported in Table 3.3.5 confirms that there is no gender gap in self-reported individual performances. Regression (1) reveals the significantly positive relationship between the players' past performance and their messaged performance. Thus, we know that messages, on average, contain useful information about individual performances. In the PMT players can only signal their ability through one message about past performance, so they may exaggerate due to their positive expectation of future performance. As shown in regression (2), once the effect of expectation is partialled out, the marginal average impact of the past performance reduces to 0.453.

Table 3.3.5: An OLS Regression of Messaged Individual Performance

	(1)	(2)	(4)	(3)
Performance in <i>Task 1</i>	0.848*** (0.060)	0.453** (0.190)	0.863*** (0.062)	0.471** (0.191)
Expectation		0.380** (0.174)		0.375** (0.175)
Initial Belief			0.011 (0.013)	0.010 (0.013)
Gender (Female)	-0.280 (0.612)	-0.089 (0.602)	-0.336 (0.617)	-0.142 (0.608)
Age (Young)	2.004** (0.912)	2.239** (0.895)	2.134** (0.928)	2.355** (0.909)
Math Training (Yes)	-1.099 (0.650)	-0.996 (0.635)	-1.014 (0.660)	-0.919 (0.643)
R-squared	0.770	0.785	0.772	0.787

Note: Coefficient with standard error in parentheses.

***, **, * denote significance at $p=0.01$, 0.05 and 0.10 respectively.

[1] Sample size is 72 and constant term is included.

[2] Dependent variable is the messaged individual performances. We control for 1) gender (1-female; 0-male) 2) age (1-under 25 ; 0-above 25) and 3) math training (1-yes ;0-no). The variable expectation is the answer we elicit from subjects by asking how many questions they can correctly solve if the task were repeated while the initial belief is the percentage of people subjects believe who can solve more questions than themselves.

In order to check whether or not players message in a strategic way, we add the subjects' beliefs of their relative performance in regression (3) and (4). A strategic player would

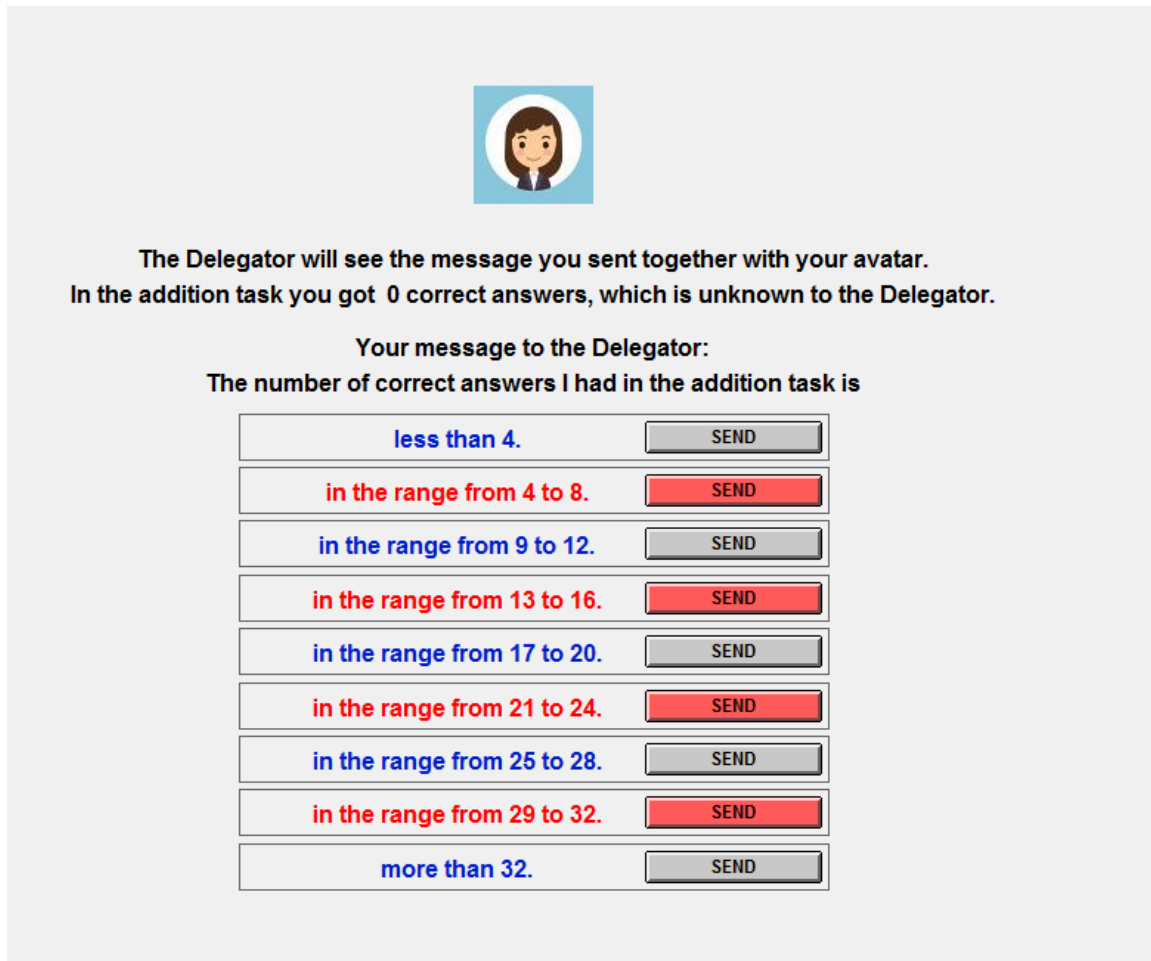
consider the likelihood that the other player performed better when deciding on a message. The coefficient of the initial belief in either regression model is close to zero and is not significant. Thus, we conjecture that useful information contained in messages is due to lie aversion (Sánchez-Pagés & Vorsatz 2007).

3.3.3.2 How Players Message in the FMT?

In the FMT, players chose an interval of four questions length to indicate where their past performance belongs. A typical screen shot is given in Figure 3.3.7.¹¹ We make the message space coarser with two aims. Firstly, we want to see whether players would lie harder to distinguish themselves from their competitor in the same group. Secondly, the coarser message space increases the chance of a tie in messaged performance. When the delegator receives two identical messages, avatars might become more salient for the delegation decision. In this way, we have a more favorable environment for observing a gender effect.

¹¹A careless mistake happened in the program where the left end point of the second band should be the right end point of the first band. However, the mistake only affects the decision of a subject whose role was delegator and performance was 1, 2, 3 or 4 which never happened in our data.

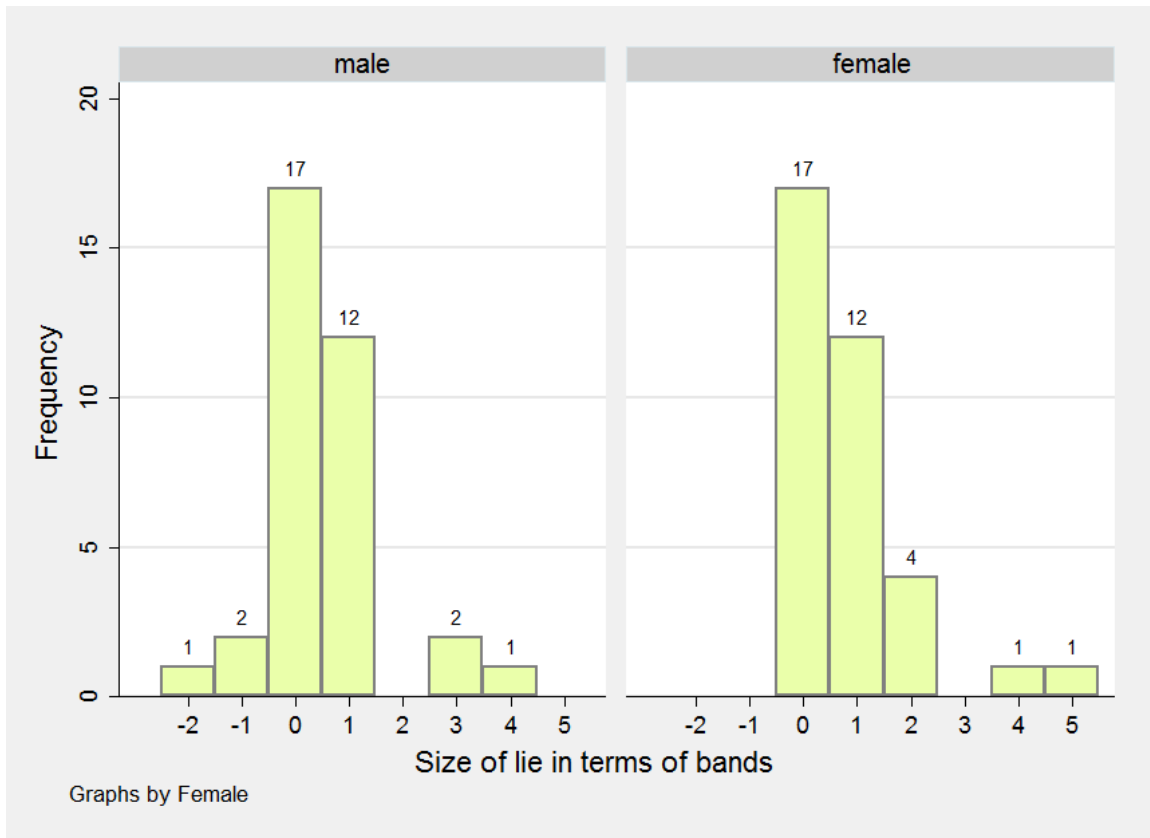
Figure 3.3.7: Message Screen in the FMT



Surprisingly, 49% of players truthfully report their past performance in the noisier environment. In the FMT, the size of a lie is measured in terms of intervals or bands. Figure 3.3.8 graphs the histogram of the size of lie by the two genders.¹² The majority of players exaggerated slightly by one or two bands and a few (3 male and 3 female) exaggerated by more than three bands. The shapes of the distribution for each gender look almost identical and the mean value is not significantly different ($p=0.588$ with a rank-sum test and $p=0.638$ with a two-sided t-test).

¹²Again, we had three subjects who messaged one or two bands worse than their past performance.

Figure 3.3.8: Discrete Histogram of Lie by Gender in the FMT



The interval setting further allows us to investigate the other-regarding motive beyond lie aversion. There are four positions in an interval, i.e. Left, Middle Left, Middle Right, and Right. Suppose that players have a stable preference over lying. Then, players would lie in the same fashion, i.e. same degree of a lie, regardless of their real position in an interval. However, if players care about welfare of the delegator in their group, their actual position in an interval affects how they would message about their past performance. The further away players' actual past performance locates to the right end, the larger the damage they could bring to the delegator if they exaggerate and the delegator believes them. Thus, players who have other-regarding preferences would refrain from lying to reduce the harm that would be done to the delegator. We conduct a probit regression of players' propensity for truthful reporting depending on their actual position in an interval controlling for their past

performance and individual characteristics. Results of the regression are given in Table 3.3.6.

Table 3.3.6: A Probit Regression of the Propensity for Truthful Reports

Dependent Variable	Coefficient	Conditional Marginal Effect
Performance in <i>Task 1</i>	-0.0136 (0.031)	-0.005 (0.012)
M_Right (Right)	0.601 (0.454)	0.229 (0.169)
M_Left (Right)	0.390 (0.482)	0.145 (0.180)
Left (Right)	1.012** (0.446)	0.387** (0.156)
Age (Young)	-1.182** (0.474)	-0.422*** (0.132)
Gender (Female)	0.041 (0.349)	0.016 (0.140)
(Pseudo) R-squared	0.1278	

Note: Coefficient with standard error in parentheses.

*** and ** denote significance at $p=0.01$ and 0.05 respectively.

[1] Sample size is 69 and constant term is included. We exclude one outlier whose position is beyond the four length of the best interval.

[2] The independent variable is the propensity of truthful which is a dummy (1-message the true performance and 0-lie about the true performance).

Agreeing with our hypothesis, sitting on the left of an interval has a significant effect, holding other variables constant. Compared to the reference level, i.e. the right end of an interval, being at the position to the left end increases the average propensity for truthful reporting by 38%. A pairwise check on the average marginal effects of different positions confirms the significant difference between the left and the right position ($p=0.013$).

Since we have found evidence for other-regarding preferences, we also expect that the further the position of a player is to the right end the more he or she will exaggerate. We conduct an OLS regression of the degree of exaggeration on the four possible actual positions of an interval controlling for actual individual performances and individual characteristics. As shown in Table 3.3.7, the coefficient of the right position is significantly positive. Compared to the left position, a player who locates on the right end, on average, exaggerates one band more.

Table 3.3.7: An OLS Regression of the Size of a Lie

Dependent	Coefficients	Standard Error
Performance in <i>Task 1</i>	-0.033	0.024
Right (Left)	1.089***	0.339
M_Right(Left)	0.150	0.353
M_Left (Left)	0.549	0.377
Age (Young)	0.707**	0.347
Gender (Female)	0.265	0.271
(Pseudo) R-squared	0.2586	

Note: *** and ** denote significance at $p=0.01$ and 0.05 respectively.

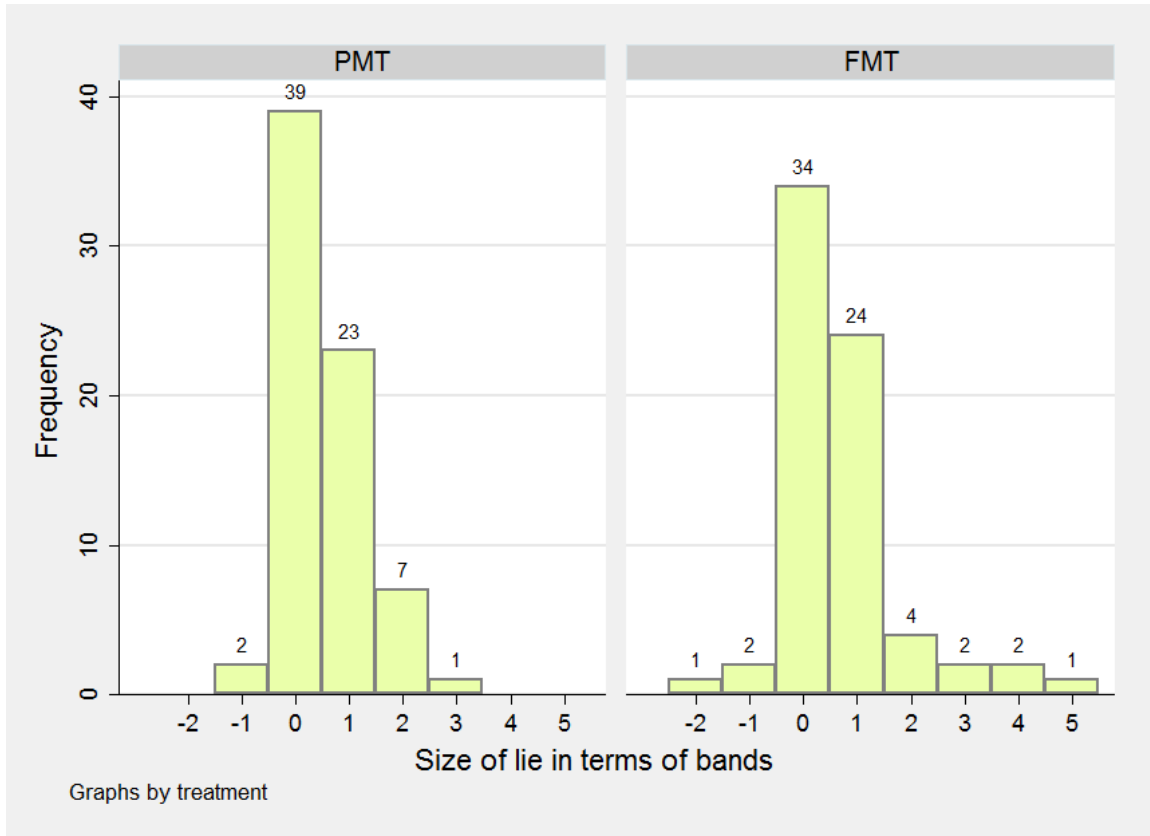
[1] Sample size is 69 and constant term is included. We exclude one outlier whose position is beyond the four length of the best interval.

[2] The Independent variable is the size of lie which is the difference of reported performance (i.e. the message) and the true performance in terms of bands.

3.3.3.3 Comparison of Messages between the Two Treatments

In this subsection, we want to address the question of whether subjects lie in the same fashion regardless of the coarseness of the message space? In the PMT, 40.28% of subjects truthfully stated their past performance, and in the FMT the percentage of truthful reports slightly increased to 48.57%. The small difference is not significant with a two-sided test of proportions ($p=0.160$). In order to compare the degree of exaggeration between the two treatments, we transfer messaged performance in the PMT into intervals in the same way we designed the interval setting in the FMT. Figure 3.3.9 depicts a discrete histogram comparison in the size of lies between the two treatments.

Figure 3.3.9: Histogram of the Size of Lies by Treatments



Although there are three very competitive players who exaggerated above four bands in the FMT, the two distributions look quite similar. The majority of players either reported their actual past performance or slightly exaggerated by one band above across both treatments. There is no difference across treatments ($p=0.647$ with a rank-sum test and $p=0.379$ with a two-sided t-test).

3.3.3.4 How Players Message in the FCT?

In the FCT, players were allowed to send a text message with a maximum of 500 characters to freely communicate anything they wanted to the delegator. We categorize the raw data into the following three categories. The most relevant information on a delegation

problem is about individual performances. Therefore, our first category is about whether individual past performance was mentioned or not. We name the first category variable “Past Performance”. Among 65 players who entered the message stage, 31 mentioned their past performance. Fourteen players truthfully stated their past performance, while seventeen exaggerated on average about four questions. We notice that statements such as “more effort”, or “better future performance”, or “better future outcome” frequently (38/65) appeared in messages. Thus, our second category captures whether a promise was made or not and call this variable “Promise”. Lastly, we observe that 19 out of 65 players mentioned their background such as accounting, computer programmers, math major and so on. Accordingly, we name the last “Background”, i.e. whether or not background information was mentioned. The three category variables summarize the typical pattern observed in the text data. Messages that do not fit into any categories are taken as babble.¹³

In order to study the amount of information revealed by the text messages, we further run probit models of the three category variables with players’ individual performances and their initial beliefs about the percentage of people who can solve more problems as independent variables. We also control for individual characteristics including gender (male or female), age (above 30 or under 30), math training, and ethnic backgrounds (asian/ white/ black/ others). Table 3.3.8 reports the main regression results of interest.

¹³Here are some examples: “Hi, I’d like to be delegator if you need help. thanks” or “excellent at it, believe it or not. 1+1=2”

Table 3.3.8: A Probit Regression of Text Messages

Dependent Independents	Past Performance (1)	Promise (2)	Background (3)
Performance in <i>Task 1</i>	0.013 (0.033) [0.004]	0.087* (0.045) [0.029**]	-0.084* (0.049) [-0.025*]
Initial Belief	-0.022** (0.009) [-0.008***]	0.001 (0.009) [0.000]	-0.015 (0.010) [-0.004]
Gender (Female)	0.234 (0.377) [0.082]	-0.001 (0.388) [-0.000]	0.677* (0.403) [0.204*]
Controls	Yes	Yes	Yes

Note: Coefficient with standard error in () and average marginal effect in [].

** and * denote significance at $p=0.05$ and 0.1 respectively.

[1] Sample size is 65 and constant term is included.

[2] Dependent variable is a dummy variable 1-mentioned; 0-not mentioned.

The significant estimated coefficients in each model suggest that information about either players' individual past performances or their relative confidence is revealed through these three categories. Past performances were mentioned by subjects who were more confident about their relative performance. Making a promise is positively correlated with individual performance, but mentioning of background is negatively correlated with performance. In summary, the information about individual performance is partially revealed in messages across the three treatments contained in the kind of messages sent by the players.

We have shown that in all treatments the messages contain some information. It remains to investigate how much the delegators actually pick up and utilize.

3.3.4 Behavior of Delegators

We first look at the behavior of delegators in the PMT and the FMT, since both treatments adopted a structured message space leading to a comparable delegation problem. In contrast,

delegators in the FCT received text messages to make a delegation decision, and therefore an analysis of the FCT is performed separately.

Unexpectedly, there is a similar average delegation frequency between the PMT and the FMT (i.e. 0.275 vs. 0.282). We would expect that the delegation problem in the FMT is harder than in the PMT, because the messages are less precise. Consider a delegator who believes that players would exaggerate their actual performance by two questions. In the PMT the delegator can make a precise rational decision based on such a belief, i.e. to delegate only if the messaged performance is at least 7 questions better than his own. However, in the FMT when the delegator receives the messaged performance which is two bands better than his, he is still unable to make a rational delegation decision. The difficulty arises in that a band includes four possible positions and the exact location of players' actual performance remains unknown. Therefore, we initially conjecture a lower delegation frequency in the FMT, as the delegation problem becomes harder. In order to have a deeper understanding, we first summarize the relevant factors for delegation in both treatments and then conduct a regression analysis at individual-data level.

3.3.4.1 Performance of Delegators

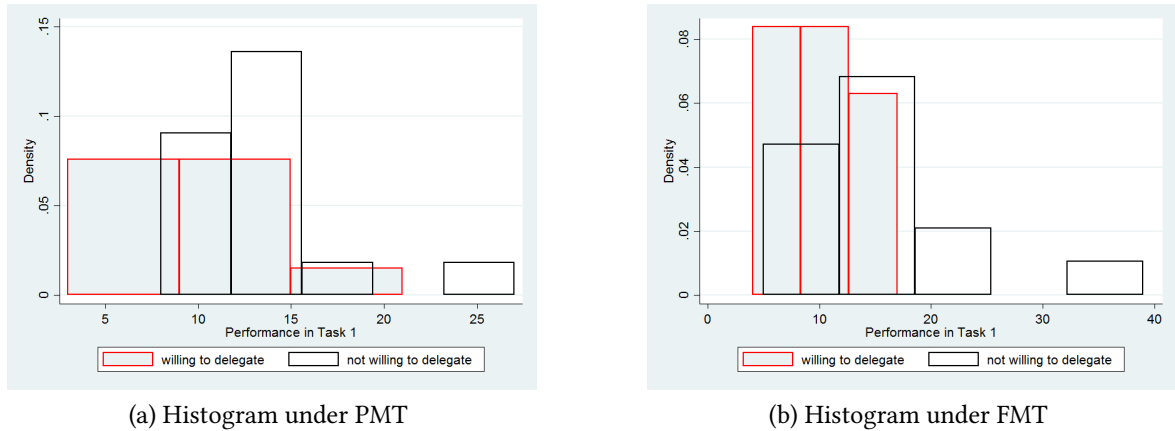
In our setting, only delegators who were not good at the addition task may benefit from delegation if a more capable candidate is chosen. Accordingly, we would expect a higher average past performance of the delegators who did not delegate than of those who delegated. As reported in Table 3.3.9 the delegators who did not delegate performed, on average, four questions better in the PMT and six questions better in the FMT than those who delegated.

Table 3.3.9: Performance in *Task 1* of Delegators Conditional on Their Decision

Treatment	Decision	Obs	Performance	One side t-test	Rank sum test
PMT	Delegate	11	9.64	p=0.0137	p=0.0092
	Not delegate	29	13.24		
FMT	Delegate	11	9.82	p=0.0114	p=0.0140
	Not delegate	28	15.86		

Below we provide a more intuitive comparison of the distributions of performance conditional on delegation decision. The overlapping area ranging from 5 to 20 correct answers suggests that delegators' own performance does not alone determine delegation decisions.

Figure 3.3.10: Performance of Delegators by Decision

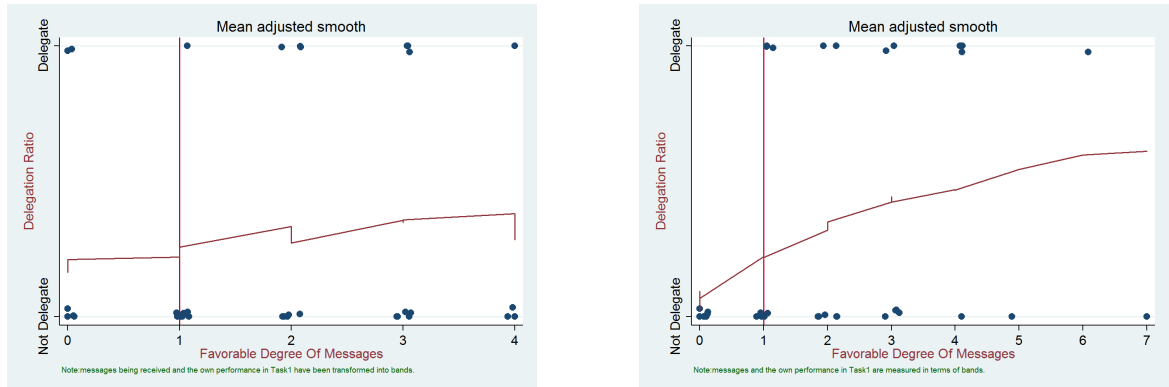


3.3.4.2 Comparison of Messages Received with the Own Performance

If delegators believe that players truthfully messaged, they should compare the messaged performance with their own performance in order to decide if to delegate. Therefore, we construct a variable that measures the maximum difference between the messaged performances and the delegator's own performance in a group. In order to have a consistent measure of messages between the two treatments, we transform the messaged numbers in the PMT into bands in the same fashion as in the FMT. We collapse the negative values of

the maximum difference variable into 0s. In this way, we treat all messages that indicate a worse performance than that of the delegators as equally irrelevant. Now we can conduct a local weighted regression to check if the relative attractiveness of the most favorable message affects frequency of delegation in the two treatments.

Figure 3.3.11: Lowess Graph by Treatments



(a) Lowess smoother in PMT

(b) Lowess smoother in FMT

Note: Sample is 40 in the PMT and 39 in the FMT. The dependent variable is a dummy (i.e. 1 for delegation and 0 for no delegation) and the independent variable is the maximum distance. The miss value from eight not available players is replaced by 0. The way we treat the missing value should not affect much since the maximum distance variable only pick the better message between alternatives.

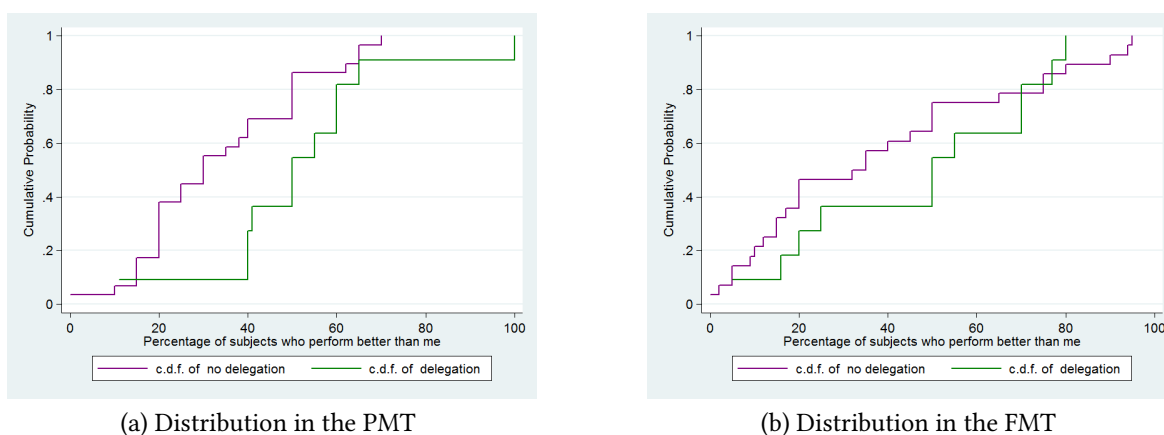
In Figure 3.3.11, the vertical line specified at one indicates the break-even point of delegation, i.e. the benefit equals to the bonus cost. The slope of the predicted delegation frequency line in Figure (b) is steeper than the slope in Figure (a). This suggests that delegators almost ignore the most favorable message in the PMT but not in the FMT.

3.3.4.3 Revisiting the Role of Initial beliefs

Since cheap talk is not verifiable, the delegators' initial belief about the distribution of individual performances is important to assess the truthfulness of messages. If delegators believe that more than fifty percent of people can outperform themselves, then they are more likely to believe a message that stated a better performance. Therefore, we graph

the distribution of elicited beliefs of delegators conditional on their delegation decision in Figure 3.3.12. In the PMT (3.3.13a) the distribution of initial belief of delegators who delegated is located to the right of those who did not delegate. The same pattern remains in the FMT(3.3.13b) for up to 80% of the population. In the FMT, there is approximately 20% of delegators who did not delegate but held the extreme belief that more than 80% of the people can perform better than themselves.

Figure 3.3.12: Cumulative Distribution of Initial Belief of Delegators



3.3.4.4 Probit Regressions

In what follows we investigate the determinants of delegation via probit regressions. For this we generate a variable favorable_high that captures how many bands the best message lies above the delegator's own performance. To make the regressions comparable, the messaged performances in the PMT are again transferred into bands as used in the FMT. As shown in Table 3.3.10 the estimated coefficient on the variable favorable_high is significant in the FMT, but not in the PMT. Surprisingly, we find that messages are informative in the FMT but not in the PMT. Delegators in the PMT discarded messages and resorted to their initial belief about their location in the distribution of performances to make a decision. The average marginal effect of the belief estimated in the PMT suggests that a 1% increase in relative confidence level increases the probability of delegation by 0.7%. In contrast, delegators make

use of received messages when delegating in the FMT. The average frequency of delegation increases by 7.2% when the best messaged performance in a group increases by one band. Once we control for the relative location of the best message in the FMT, the initial belief does not have any impact anymore.

Table 3.3.10: A Probit Model (1) of Delegation Decision

	PMT		FMT	
Number of Obs.	40		39	
	Coefficient	A.M.E.	Coefficient	A.M.E.
Initial belief	0.026** (0.012)	0.007** (0.003)	0.007 (0.008)	0.002 (0.002)
Favorable_high	0.016 (0.039)	0.004 (0.106)	0.251** (0.117)	0.072*** (0.027)
Gender (Female)	0.169 (0.524)	0.045 (0.138)	-0.079 (0.464)	-0.023 (0.133)
Math Training (Yes)	-0.627 (0.527)	-0.173 (0.142)	0.393 (0.580)	0.106 (0.144)
Age (Young)	-0.728 (0.891)	-0.218 (0.272)	0.211 (0.588)	0.058 (0.155)
Constant	-0.905 (1.014)		-1.806* (0.924)	
Pseudo R2	0.180		0.133	

Note: Coefficient with standard error in parentheses.

***, **, * denote significance at $p=0.01$, 0.05 and 0.10 respectively.

[1] Dependent variable is a dummy (1-delegation and 0-nodelegation).

A still open question is if the less favorable message has any information content. Therefore, we develop an alternative approach to examine the process of the delegation decision. We label the messaged performance that is strictly better than the delegator's own performance as a "good message". We conjecture that delegators would be more likely to delegate when they receive two good messages, less likely to delegate when receives one good message, and the least likely to delegate when receives no good message at all. Consequently, we can create a set of category variables to combine information contained in the two messages into three cases: 1) two good messages, 2) only one good message, and 3) no good messages. In the case where delegators only receive one message, since the other player does not want

to be considered, we either will have one or no good message.

Table 3.3.11: A Probit model (2) of Delegation Decision

	PMT	FMT
Number of Obs.	40	28
Initial belief	0.026** (0.012)	-0.001 (0.010)
No Good Messages	omitted	empty
One Good Messages	-0.254 (0.806)	-1.351** (0.596)
Two Good Messages	0.423 (0.679)	omitted
Pseudo R2	0.209	0.194

Note: Coefficient with standard error in parentheses.

***, **, * denote significance at $p=0.01$, 0.05 and 0.10 respectively.

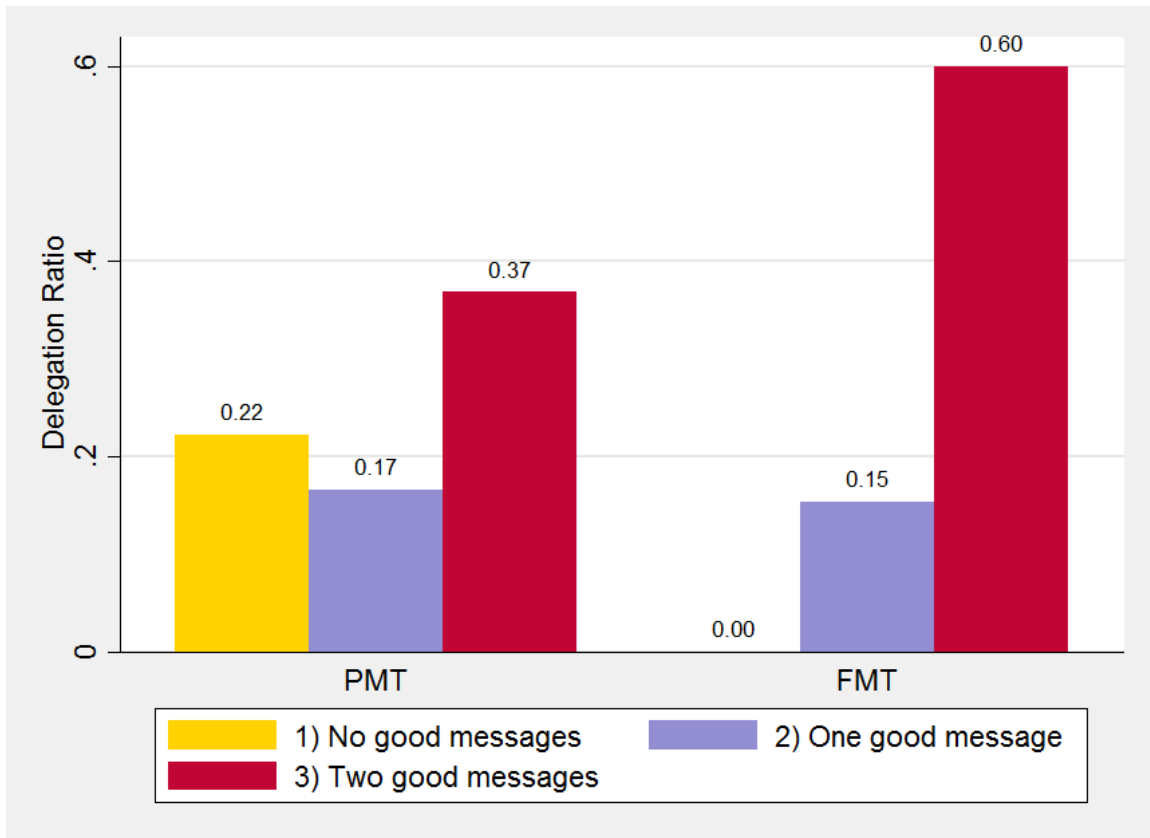
[1] Dependent variable is a dummy (1-delegation and 0-nodelegation). [2] We control for the characteristics of the delegator such as gender, math training and age group.

Results of the resulting probit regression are reported in Table 3.3.11 and confirm that information transmission takes place in the FMT, but not in the PMT. We observe that in the PMT it does not make a difference for delegation decision how many good messages a subject receives, while the initial belief has a significant impact. In the FMT, all eleven subjects that did not receive a good message refrained from delegating. Therefore, the category “no good message” perfectly predicted failure and the observations were dropped. There is a further significant difference between receiving two or one positive message, which provides further evidence that the messages do not only contain information but that it is also transmitted to the delegator.

In Figure 3.3.13, we present the average delegation frequency conditional on how many good messages delegators received. The differences of the average delegation frequencies are more pronounced in the FMT than in the PMT. A pairwise proportion test confirms the difference is significant in the FMT but not in the PMT.¹⁴

¹⁴In the PMT, the one-sided p-values are 0.6258, 0.1140, and 0.2196 in each case, and in the FMT the one-sided p-values are 0.087, 0.0080, and 0.0007.

Figure 3.3.13: Delegation Frequency by Message Categories

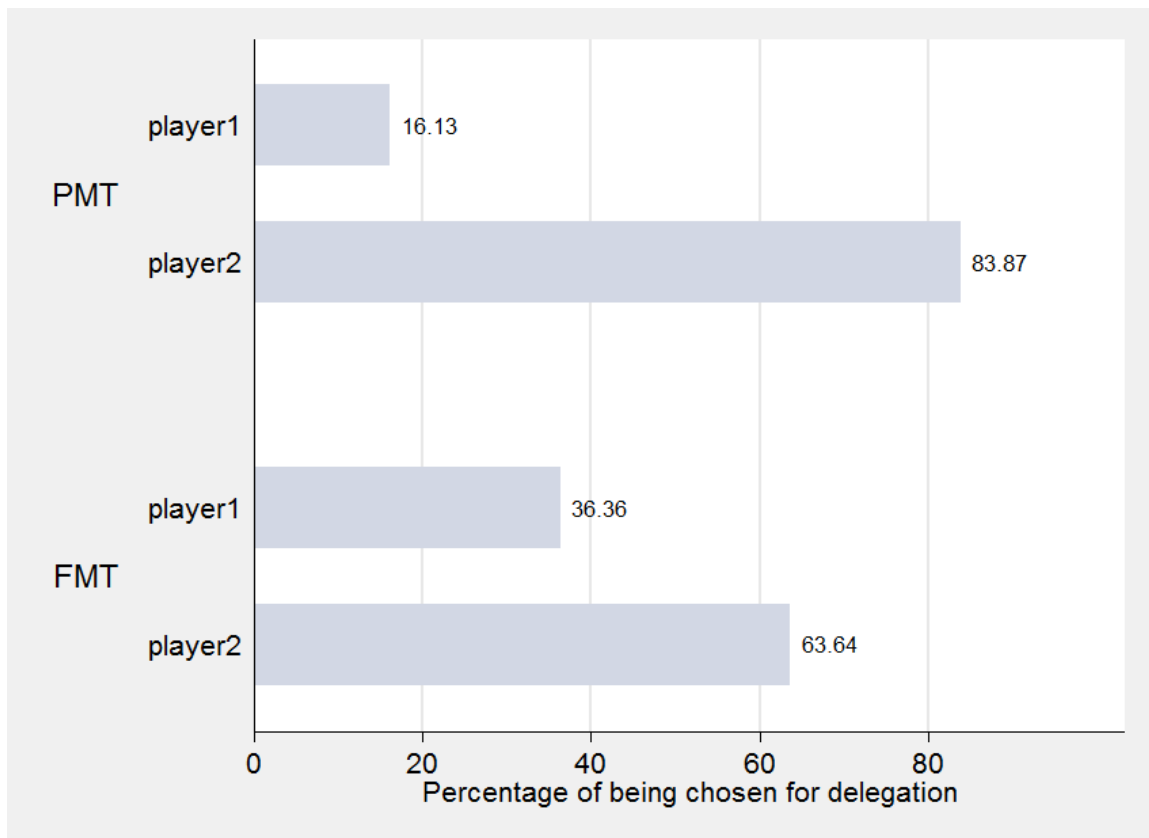


Next, we have a closer look at to whom delegators delegated. Since only 11 delegators delegated in each treatment, regression analysis is not a sensible option. Instead we just check whether delegation always went to the player who sent the better message. We find only one exception in each treatment. A male delegator who solved three correct answers in the PMT delegated to the male player who messaged 20 but not to the female player who messaged 22. A male delegator whose actual performance belonged to band 3 in the FMT delegated to the male player who messaged band 6 but not to the female player who messaged band 7.

3.3.4.5 Is There a Gender Gap in Delegation in the PMT and the FMT?

An earlier study (Reuben et al. 2012) shows that male players are chosen more frequently than female players for a group task due to the stereotype that males are better at solving math problems than females. Figure 3.3.14 depicts the average frequency of being chosen for delegation by gender across the two treatments. In the PMT, males are chosen four times more often than females and in the FMT males are chosen at a bit more than half the frequency than females. The frequency of being chosen for male is significantly different from 0.5 in the PMT ($p=0.0327$) but not in the FMT ($p=0.2744$) according to a binomial probability test. The gender gap of being chosen for delegation seems to disappear when information is transmitted. Unfortunately, we do not have enough cases of delegation to sensibly investigate where the gender gap in the PMT comes from.

Figure 3.3.14: Frequency of Being Chosen for Delegation



3.3.4.6 How Delegates Delegate in the CFT?

In the CFT, we have created three dummy variables (i.e. Past Performance; Promise; Background) in order to characterize the content of the text messages and found that the choice of content categories when messaging contains information about the past performance of the sender. Here we investigate if the delegators are able to extract this information.

In what follows we run a probit model where delegation is the dependent variable. The independent variables are the delegator's performance in *Task 1*, her initial belief about her relative performance. And dummy variables for the content of the free text messages. These indicate if a) at least one player mentioned a performance that would lead to a profit if true, b) at least one player made a promise and c) at least one player mentioned his or her background. Table 3.3.12 displays the results.

Table 3.3.12: A Probit Regression of Delegation Decision in the FCT

Variables	(1)	(2)	(3)	(4)	(5)
Performance in <i>Task 1</i>	-0.267* (0.142)	-0.322** (0.129)	-0.298*** (0.111)	-0.333* (0.191)	-0.333** (0.132)
Initial Belief	0.026 (0.018)	0.014 (0.014)	0.014 (0.013)	0.026 (0.018)	0.015 (0.014)
Better Performance (at least one better)	2.045*** (0.707)			2.100*** (0.744)	
Promise (at least one making)		0.905 (0.695)		0.990 (0.990)	0.928 (0.715)
Background (at least one mention)			0.288 (0.537)		0.298 (0.555)

Note: Coefficient with standard error in parentheses.

***, **, * denote significance at p=0.01, 0.05 and 0.10 respectively.

[1] Constant term is included.

[2] Dependent variable is a dummy variable 1-delegation; 0-no delegation.

Column (1) suggests that delegators increase the frequency of delegation significantly if they receive at least one message which mentioned a past performance that was at least four

questions better than her own. As shown in columns (2) and (3), neither making promises alone, or mentioning the background alone, has a significant impact on delegation decisions. We further control for making a promise in column (4) as a robustness check for the positive effect on the variable Better Performance. The coefficient remains highly significant. Also, the average marginal effect of the variable Better Performance remains significant at 42% compared to 45% in column (1). Due to the severe co-linearity issue the regression can not differentiate the effect of the variable Better Performance from the variable Background.¹⁵ However, as shown in column (5) “stating promises” and “backgrounds” without “mentioning the at least 4 questions better performance” do not have a significant effect on delegation decisions.

3.4 Conclusion

In this paper, we investigate the efficiency of delegation when cheap talk is allowed in the context of a real-effort task. If individual performances in the task are public information, allocating the task to the candidate who performs best maximizes efficiency. However, individual ability is in reality more often than not private information. One can credibly signal individual ability, for instance, by taking costly education (Spence 1973). A simpler solution is to directly communicate. In reality applications and interviews to a certain extent serve this purpose. Theorists are skeptical about the benefits from cheap talk as they expect that rational players always lie if the incentive to do so exists. On the other hand, experimental research (Fischbacher & Föllmi-Heusi 2013) shows that people in general are averse to lying. Thus, we ask if cheap talk can facilitate information transmission with human players despite that in standard equilibrium information revelation is impossible. We create a novel three-person cheap-talk delegation game in the laboratory to investigate how differ-

¹⁵Delegation always happened when at least one better performance and background are mentioned in all 4 groups. Delegation never happened when none of better performance and background are mentioned in all 9 groups.

ent forms of messages affect information transmission and consequently the efficiency of delegation. We implement three treatments. In the baseline treatment, i.e. PMT, players can choose any number to represent their past performance. Then we vary the cardinality of the message space. In the FMT, we make the message space coarser by partitioning the space into intervals. In the CFT, we make the message space much richer by allowing free text communication.

We find that on average the majority of subjects either truthfully report their true past performance or moderately exaggerate. The interval setting in the FMT allows us to further identify other-regarding preferences beyond lie aversion. Therefore, information about individual ability is partially revealed in messages across all treatments due to either lie aversion or other-regarding preferences. However, information is transmitted in both the FMT and the CFT but not in the PMT. Unexpectedly, delegators in the PMT disregarded messages and relied on their initial beliefs about their relative individual performance when making delegation decisions. Although the average delegation frequency is similar across all treatments, efficiency obtained in both the FMT and the CFT is significantly higher than in the PMT.

We also add the avatar setting to expand the social dimension in the cheap talk delegation game, which allows us to explore the gender gap and non-monetary motives. We find that male avatars are significantly chosen more frequently than female avatars in the PMT. However, we are not able to differentiate whether the observed gender gap is caused by the demand or the supply side.

There are two potential extensions in the current work. Real life communication takes place in the form of conversations in which people talk sequentially. It would be interesting to see if a higher efficiency is possible with two-way communication rather than a one-way communication. Alternatively, one could provide information of delegators' performances to the available players before they have to send a message. Given that our subjects exhibited

either lie aversion or other-regarding motives, we wonder whether there is a way to amplify the positive effects of such pro-social motivation. If information on the employers' ability is commonly known, then players would have a clear idea of how much harm they can incur by lying.

Chapter 4

Cooperation and Distributional Conflict with a Costly Contract Option

4.1 Introduction

One important reason for drawing up a contract is to regulate future actions. A premarital agreement is one such example. Marriage is like a joint venture in which the cooperation of both spouses generates joint wealth for the family. In the event of the breakup of a marriage, division of family's property can be a long and emotional process. In order to avoid protracted and bitter fights, a contract can be entered into prior to marriage to specify the split of the joint property in case of a break-up. Often it is difficult, expensive or even impossible to include all possible contingencies in a contract. Thus, contracts usually use general rules which sometimes require interpretation. Contract and mechanism design theory has shown that in theory appropriately designed contracts can improve welfare. In this paper, we study the impact of a contract option to resolve distributional conflicts that could be caused by uncertainty over property rights over proceeds from cooperation in advance.

Our study relates to recent papers that explore other mechanisms of conflict resolution, for instance, bargaining (Herbst et al. 2016), a side payment (Kimbrough & Sheremeta 2013) or non-binding random devices (Kimbrough et al. 2015). The mentioned studies consider a conflict situation between two players fighting over a prize of given value. This paper, however, considers a situation where the value of a prize is endogenously determined by a voluntary contribution game played between the two players. In this richer environment, we can study how anticipated distributional conflict and contract options to preempt its influence on the degree of beneficial cooperation.

An exemplary situation captured by our setup is a joint venture. Two firms each possess one of two complementary technologies. Both technologies alone are useless but together they are productive. If the two firms decide to produce together, proceeds become their joint assets. A distributional conflict (i.e. how to divide proceeds) arises when the property right over those joint assets are not well defined.

This paper extends the framework of Bayer (2016*b*), who combines a voluntary contribution game with a second-stage share contest over the proceeds from contributions. In this setting the second stage game creates a hold-up problem. Rational players who foresee the severity of future conflicts should refrain from cooperating in the first place. Bayer finds that subjects still cooperate to a certain extent in the social dilemma situation, but that welfare losses from competition in distributional contest destroy welfare gains from voluntary cooperation. Consequently, we ask in this paper if providing players with an ex-ante contract option that if taken determines a sharing rule improves social welfare.

We extend Bayer's model by providing an enforceable and costly contract option in advance, and then experimentally study the degree of social welfare improvement. A contract in our context specifies a pre-determined sharing rule for the proceeds generated in the cooperation game. If such a contract is agreed upon, then there is no uncertainty over the outcome of a future distributional conflict. In the given environment it is easy to show that an opti-

mal contract that theoretically achieves the first best exists if the sharing rule can condition on the contributions in the voluntary contribution stage. Splitting the proceeds according to the proportions of contributions (henceforth PSC for Proportional Split Contract) aligns individual incentive with the aim to maximize total welfare. Both the social dilemma and the hold-up problem are solved at the same time. Such a contract requires that contributions are verifiable in court. This is a very strong assumption that rarely holds in real life.

When circumstances become too complicated, parties usually turn to some commonly accepted social norm for a simple solution. For the case that contributions are not verifiable, an ex-ante rule can only condition on the total proceeds but not on how much each partner has contributed to it. A natural rule would be an equal split (henceforth ESC for Equal Split Contract). If adopted, then the equal split sharing rule turns the game into a standard public goods game where agent have an incentive to free ride by investing nothing. An ESC does not resolve the social dilemma of cooperation. In terms of total surplus the PSC is the first best implementation, while the ESC should achieve the same level of welfare as achieved without a contract option. In what follows we will give subjects the choice to agree or not on a contract before they play the joint-venture game. Which contract they can choose and what the transaction cost is differs across treatments. If they sign a contract then instead of a costly Tullock share contest determining the split of the proceeds, the sharing rule specified in the contract will be enacted.

We use the *unconstrained_low* treatment in Bayer's paper as the base and control treatment. There all players have to enter a contest game and exert costly effort to determine their share of the proceeds. In this two-stage cooperation and contest game, the social dilemma of a cooperation game and the hold-up problem due to future costly conflict coexist. On the equilibrium path, the subgame-perfect Nash solution predicts no investments in the cooperation game and no efforts in the contest game. When a ex-ante contract option is present, standard theory predicts different equilibrium behaviors depending on the type

contract offered. Since the PSC incentivizes full contributions in the cooperation game, rational players should choose to implement the contract as long as the cost is not too high and then contribute their full endowment. In contrast, the ESC does not remove the social dilemma in the cooperation game and accordingly rational players should invest nothing and reject the contract as long as it is costly. The theoretical predictions for the ESC and the control treatment with respect to contributions and payoffs are identical.

In *unconstrained_low* treatment, subject's actual behavior deviates from the prediction. They are able to overcome the social dilemma to a certain degree and play, on average, close to equilibrium in the continuation game. Optimal contest efforts destroy more welfare than the contribution create. Subjects in the end are worse off than if they did not contribute at all. Contributions and therefore losses decline over time. We are now interested in seeing if providing costly contract options can improve welfare. In the treatment *ESC_high*, we offer subjects to enter an expensive ESC. The cost is calibrated such that it is equal to the average amount of resources wasted in the distributive contest in the base treatment. Here we only expect a welfare improvement if signing a contract induces higher contributions, which is not predicted by standard theory though. In contrast, theory would predict a sizable welfare gain in our *PSC_high* treatment, where an expensive PSC is offered. As writing an ESC is less complicated since it does not have to condition on contributions we add a treatment with a cheap ESC (i.e. *ESC_low*).

Roughly consistent with theory, the majority of subjects (about 72%) in the *PSC_high* treatment are able to reach a mutual agreement to implement the contract and achieve maximum contributions. In light of this, it is no surprise that the highest social welfare is obtained in this treatment. More interestingly, even in the *ESC_high* treatment a significantly higher social welfare is obtained than in the control treatment. We find that agreeing on an equal split contract per se has a significantly positive impact on individual contributions. In the *ESC_high* treatment, approximately 33% of subjects agree to implement a contract, while

in the ESC_low treatment the percentage increases up to approximately 54%. When an equal split contract is carried out, the two-stage game is reduced to a standard public-goods game. In line with the survey of (Chaudhuri 2011, Ledyard 1995) on the dynamics of contributions in voluntary contribution games, we find that subjects, on average, contribute positively (rather than the Nash equilibrium prediction of zero contribution) and consecutively reduce their contributions with repetition. A surprising difference to the dynamics in a standard public goods games is that subjects that agree to a contract do not further reduce in contributions in later stages.

When a contract fails to be implemented, since one or both parties do not want to sign the contract, then subjects enter a contest stage to compete for their share of the proceeds. As in Bayer's experiments we find behavior in the contest if entered that is on average close to equilibrium. We further identify some asymmetric fighting behavior in the ESC_low treatment where subjects fight harder if they contribute more to the group project.

To our best knowledge, we are the first to experimentally investigate the role of a contract to resolve future distributional conflicts. Various theories of the firm (Gibbons 2005) has been developed to address the issue of rent-seeking behaviors, i.e. haggling over appropriable quasi-rents. Our setting provide a simple two-stage stage game with an ex ante contract option to test how different transaction costs and contract options affect social efficiency. A contract in our context defines the property right of the joint asset in advance to avoid future costly rent-seeking. In this sense, our design focuses on ex ante incentive alignment (Grossman & Hart 1986, Hart & Moore 1990) rather than ex post decision governance (Klein et al. 1978, Williamson 1971, 1979).

There are two other strands of experimental research studying contract in a different setting. Most earlier experimental investigate the behavioral validity of the incentive of contract in a principal-agent setting and focus on the discussion of various social preferences (e.g.: Anderhub et al. 2002, Fehr et al. 2007, Frey & Jegen 2001). A smaller line of research

studies relational contracts as the reference points for later negotiation in a seller and buyer situation with a hold-up problem (Hoppe & Schmitz 2011, Fehr et al. 2015).

4.2 Experimental Design:

We extend the design of Bayer (2016a) by providing a contract option in advance. We first present the extended model and show how the equilibrium behavior changes when a contract option is introduced. Then we show how we determine the value of parameters for the different treatments. We conclude this section by summarizing how we conducted the experiments.

4.2.1 Basic Model and Predictions

We modify the two-stage cooperation and contest game introduced by Bayer (2016b) by adding a costly contract option. An ex-ante contract specifies how to split the proceeds from investments in the cooperation game. The introduction of a contract transforms the two-stage game to a one-stage cooperation game if the contract is signed by both players.

The sequence of play in our experiment is summarized in Figure 4.2.1. In the first stage the two players in a group are offered a contract, depending on the treatment either an ESC or a PSC. They simultaneously make a decision of whether or not to accept the proposed contract by choosing either “Yes” or “No”. Once both players make their decision, they are informed of the other group member’s choice and whether the contract is implemented or not. If they both choose “Yes”, then the contract will be implemented with a cost of T for each. The game continues following the green arrow. If either of them chooses “No”, then the contract will not be carried out and no cost will be incurred. The game proceeds along the red arrow. In the second stage, subjects play the voluntary contribution game

independently. Only the players who are on the path indicated by the red arrows enter a third stage to exert costly effort to compete for their share. In the contest stage, subjects know the value of the project they are competing for. A profit calculator is also available at the bottom screen where players can enter hypothetical effort choices for themselves and their opponents and see the payoff consequences. Once choices are made, players are shown a summary of their own and their partner's investment, value of the project, their own and partner's effort if no contracts were signed, their share of the proceeds, and the final profit.

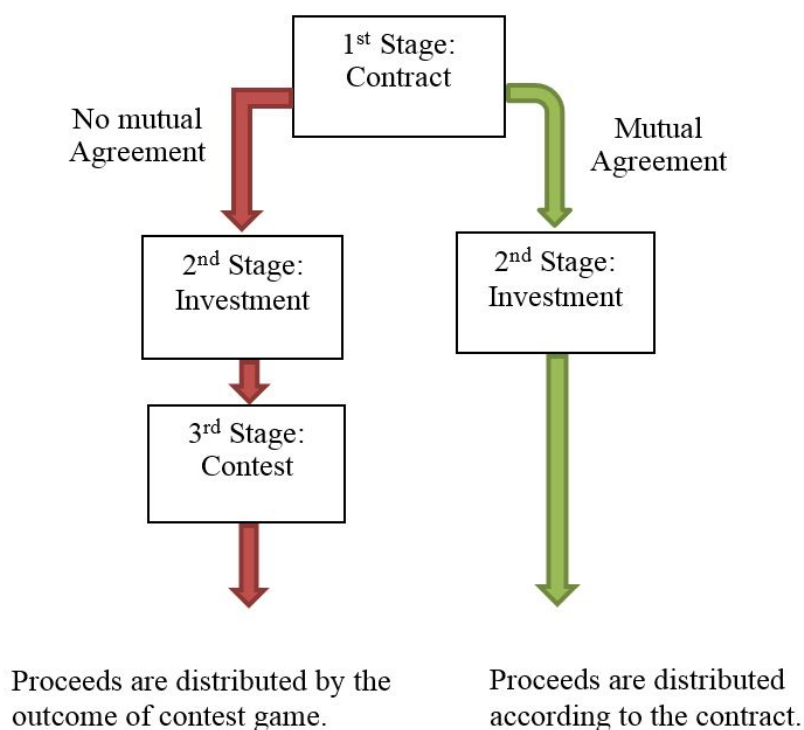


Figure 4.2.1: Game structure

The cooperation game at the heart of the model is a linear two-player voluntary contribution mechanism. Two players i and j are given an initial endowment C to voluntarily invest in a group project.

The value of the group project V is linear in each player's investment, denoted by c_i and c_j , with a constant investment technology ϕ :¹

¹In this paper we use investment and contribution as interchangeable.

$$V := \phi(c_i + c_j).$$

The investment technology ϕ measures the marginal social return to one unit of investment. It is socially efficient to invest to the full endowment, i.e. $c_i = c_j = C$, as long as $\phi > 1$. A social dilemma arises when it is not rational for an individual to fully invest. The individual investment strategy depends on how the group value is divided between the two players.

We consider two variations of a sharing rule. The first one implements an equal split (i.e. ESC) and the second one gives a subject a share equal to the proportion of her contributions (i.e. PSC). Let us denote the share of the value V accruing to i as $\rho_i(c_i, c_j)$. We write the two sharing rules compactly as follows:

$$\rho_i(c_i, c_j) := \begin{cases} \frac{1}{2} & \text{if ESC is adopted} \\ \frac{c_i}{c_i + c_j} & \text{if PSC is adopted} \end{cases}.$$

If both players agree to implement a contract, each is charged a fixed cost T . The payoff of player i , who has signed a contract, is specified as:

$$U_i := C + \rho_i V - c_i - T = C + \rho_i \phi(c_i + c_j) - c_i - T.$$

Now let us solve the maximization problem for player i case by case.

Suppose an equal split contract is implemented, i.e. $\rho_i = \frac{1}{2}$.

$$U_i = C + \frac{1}{2}\phi(c_i + c_j) - c_i - T = (C - T) + \left(\frac{1}{2}\phi - 1\right)c_i + \frac{1}{2}\phi c_j.$$

Since half of the investment of player i also goes to player j , we require $\phi > 2$ to induce full investment. The marginal benefit of investing, i.e. $\partial U_i / \partial c_i = 1/2\phi - 1$, is negative if $\phi < 2$. When $\phi < 2$, the net return to one unit investment of player i is negative. A rational player would minimize the loss by investing nothing. Therefore, when $1 < \phi < 2$, a social dilemma arises. Zero contribution is a dominant strategy for a rational player, but full investment maximizes the social welfare.

Suppose a conditional proportional split contract is implemented, i.e. $\rho_i = c_i / (c_i + c_j)$.

$$U_i = C + \frac{c_i}{c_i + c_j} \phi (c_i + c_j) - c_i - T = (C - T) + (\phi - 1) c_i.$$

Now player i gets the whole ϕ back for each unit of his or her investment. The marginal benefit of investing, i.e. $\partial U_i / \partial c_i = \phi - 1$, is positive if $\phi > 1$. Therefore, it is beneficial to fully invest individually and socially if $\phi > 1$. Hence, the social dilemma in the cooperation game disappears.

If player i and j fail to reach a mutual agreement to carry out a contract, then they enter a contest game to compete for their share of the group proceeds V . In the contest game the two players make irreversible costly efforts e_i and e_j simultaneously knowing the outcome in the cooperation game. As in Bayer (2016a) we assume a standard Tullock contest function and the share is increasing proportionally with the share of the total effort, i.e. player i receives a share $\rho_i = e_i / (e_i + e_j)$. If both players exert zero effort, then they equally split the proceeds. The payoff of player i is specified as follows:

$$U_i := C + \rho_i (e_i, e_j) V - c_i - e_i.$$

If the realized group value is zero, then there is nothing to fight over. Hence, the optimal effort is zero. If the realized group value is greater than zero, then the optimal effort is determined by the first-order-condition:

$$\frac{\partial}{\partial e_i} U(V, e_i, e_j) = \frac{e_i}{(e_i + e_j)^2} V - 1 = 0.$$

Since player i and player j face the same problem, we can exploit symmetry by setting $e_i^* = e_j^*$ and solve for the optimal effort:

$$e_i^*(V) = e_j^*(V) = V/4.$$

Now we can solve the two-stage game with an ex-ante contract option backwards. If $1 < \phi < 2$ and no contract is assigned, the zero contribution and zero effort are expected on the equilibrium path as shown in Bayer (2016a). When $1 < \phi < 2$, we expect zero contribution if an ESC is assigned and full contribution if a PSC is signed. Therefore, as long as the transaction cost T is positive no rational player should sign the ESC. In contrast, as long as the transaction cost is not too high, i.e. $T < \phi C$, rational players should sign the PSC. We summarize the equilibrium predictions as follows:

Suppose an ESC is offered. For $1 < \phi < 2$ and $0 < T < \phi C$, no contract is signed, $c_i^* = c_j^* = 0$, $e_i^* = e_j^* = 0$ and predicted welfare equals $2C$.

Suppose a PSC is offered. For $1 < \phi < 2$ and $0 < T < \phi C$, the contract is signed, $c_i^* = c_j^* = C$, and predicted welfare equals $2(\phi C - T)$.

4.2.2 Parameters of Treatments

In order to make our results comparable to Bayer's *unconstrained_low* treatment we chose $\phi = 1.6$ and $C = 20$ points.² When no contract option is available before the two-stage cooperation and contest game, rational players would contribute zero in the first stage and invest zero effort in the second stage on the equilibrium path. The payoff on the equilibrium path is equal to the endowment C . Suppose an equal split contract is available and implemented. The two-stage game is reduced to a standard public goods game where zero contribution is still a dominant strategy. The payoff from playing the dominant strategy

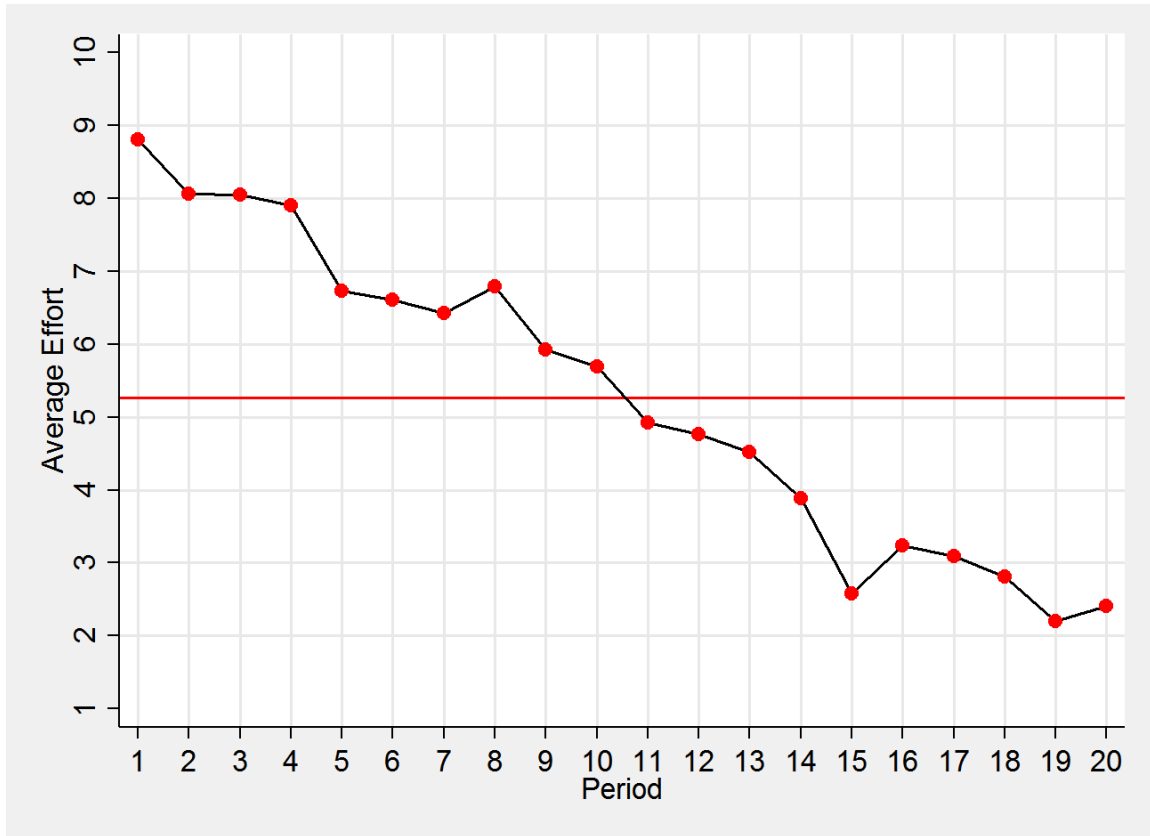
² Here $\phi = 1.6$ corresponds a VCM with the marginal per capita return of 0.8 where subjects contribute substantially despite of the social dilemma.

with an equal split implemented is $C - T$. As long as the cost of contract is positive, i.e. $T > 0$, rational players would prefer to participate in the two-stage game instead of choosing the ESC. Now suppose a conditional proportional split contract is carried out. Then the social dilemma in the public goods game is internalized by the full contribution incentive. In this case the payoff becomes $1.6C - T$. As long as the cost of the contract is less than $0.6C$, i.e. 12 points, it is profitable to take the PSC and fully invest in the cooperation game.

In absence of a contract option, there is no budget constraint to restrict how much resources can be spent in the contest stage. By committing to a contract the cost of resolving the future distributional conflict is bounded by the implementation cost. To calibrate the value of the implementation cost, we use the data from Bayer's *unconstrained_low* treatment. Figure 4.2.2 depicts dynamics of the average effort incurred in each period in the distributional contest. The red line indicates the average taken over all periods.

As a benchmark treatment, we set the cost of an equal split contract to be the average effort incurred in the *unconstrained_low* treatment, i.e. 5 points. If subjects all choose the contract and contribute in the same fashion as in the *unconstrained_low* treatment, we would expect no treatment effect in terms of the average profit. As shown in Figure 4.2.2 the average effort starts at around 9 points and drops over time to about 2 points. The decline of the average effort follows the decreasing proceeds over time. Subjects on average play closely to the subgame-perfect continuation effort level, which means that they exerted on average an effort equal to a quarter of the proceed from investment. An ESC with 5 points cost is beneficial in early periods, especially for subjects who contribute large portions of their endowment but will become unprofitable if contributions falls and efforts are close to equilibrium.

We denote this ESC treatment with $T = 5$ by ESC_high. We vary our treatments in two dimensions. Along one dimension we keep the equal split sharing rule the same but decrease the contracting cost to 1. This is the treatment ESC_low. By reducing the cost of contract-



Note: Pooled data are used to calculate the average over each period, i.e.108 subjects per period.

Figure 4.2.2: Average effort in unconstrained low treatment

ing, we expect an increase in the average profit if the contracting frequency increases and subjects keep making high contributions if contracted. Along the other dimension, holding the cost constant at 5 we replace the sharing rule by the conditional proportional split which theoretically overcomes both the social dilemma in the cooperation game and the hold-up problem created by future distributive conflict. By implementing the PSC, we expect subjects to fully contribute and achieve the highest average profit. We refer to this treatment as PSC_high.

4.2.3 Experiment Procedure

For each contract treatment we ran three sessions with an even number of subjects. Subjects were recruited through ORSEE (Greiner 2015). In each session subjects played a three-stage game using software z-tree (Fischbacher 2007) for 24 rounds of repetition. We conducted 12 sessions for the three treatments in a shuffled order at Adelaide University Laboratory for Experimental Economics. We formed a matching group among every four subjects if possible and employed a stranger design within the matching group.³ After each period new groups of two were randomly determined within the subgroup. Subjects were informed about the random stranger matching protocol but not the subgroup design. The partition of subjects into subgroups provided us with a reasonable number of independent observations.

In addition, we use Bayer's *unconstrained_low treatment* as our control treatment where all subjects have to enter a contest stage to determine their share of the group proceeds. In this treatment five sessions were conducted with 20 rounds of repetition. The same matching protocol was adopted for the last three sessions and a stranger matching without subgroup was applied to the first two sessions. Table 4.2.1 summarizes session information across different treatments.

Table 4.2.1: Session information

Treatments	Control (Old)	Control (New)	ESC_high	ESC_low	PSC_high
Sessions	2	3	3	3	3
Subjects	108		62	60	66
Repeated Rounds	20		24		
Conducted Time	Nov, 2012 & 2013	Oct , 2014	Sep & Oct , 2016		
Exchange Rate	70 Points = \$1	40 Points = \$1	30 Points = \$ 1 (AUD)		
Matching Pool	all subjects	six subjects	four subjects if possible		

Overall we had 188 subjects participate in our experiment. The majority of subjects were

³If the number of subjects is not a multiple of 4, then we have one subgroup of 2 subjects only.

university students from different disciplines such as economics, computer science, accounting and so on. We adopted imaginary money, i.e. points, rather than Dollars to make it more intuitive for subjects to understand their earnings in each round. Subjects were informed the fixed exchange rate at which points were converted to Australian Dollars by the end of the experiments. As shown in Table 4.2.1, we adjusted the exchange rate in accordance to nominal inflation. On average we paid subjects 22 Dollars for one and a half hours which includes the reading of instructions and payment.

4.3 Results

Social welfare in our setting depends on: 1) contracting decisions, 2) investments in the voluntary contribution stage and 3) efforts incurred in the distributional contest stage if no contracts are signed. We start by identifying the treatment effect with respect to social welfare. Then we look into behaviors in each stage at both the aggregate and individual levels to identify the driving factors behind the observed treatment effect.

4.3.1 Social Welfare

The focus of our study is on how much welfare improvement can be made with a contract option in advance to resolve future distributional conflicts. We compare the data at an aggregate level where each matching group qualifies as an independent observation.⁴ In all treatments losses in social welfare arise from the lack of contributions. In addition, welfare losses arise from costly efforts incurred in the distributional contest game in the control treatment. Across our three contract treatments, additional welfare losses arise from either the implementation cost if a contract is carried out, or costly contest efforts if a contract fails. Without loss of generality, we use the average profit subjects earned to represent the

⁴In the control treatment the whole session is one independent observation for the first two sessions.

average social welfare in a treatment. Figure 4.3.1 plots the dynamics of the average profit for different treatments over time to give an intuitive look of the treatment effect.⁵

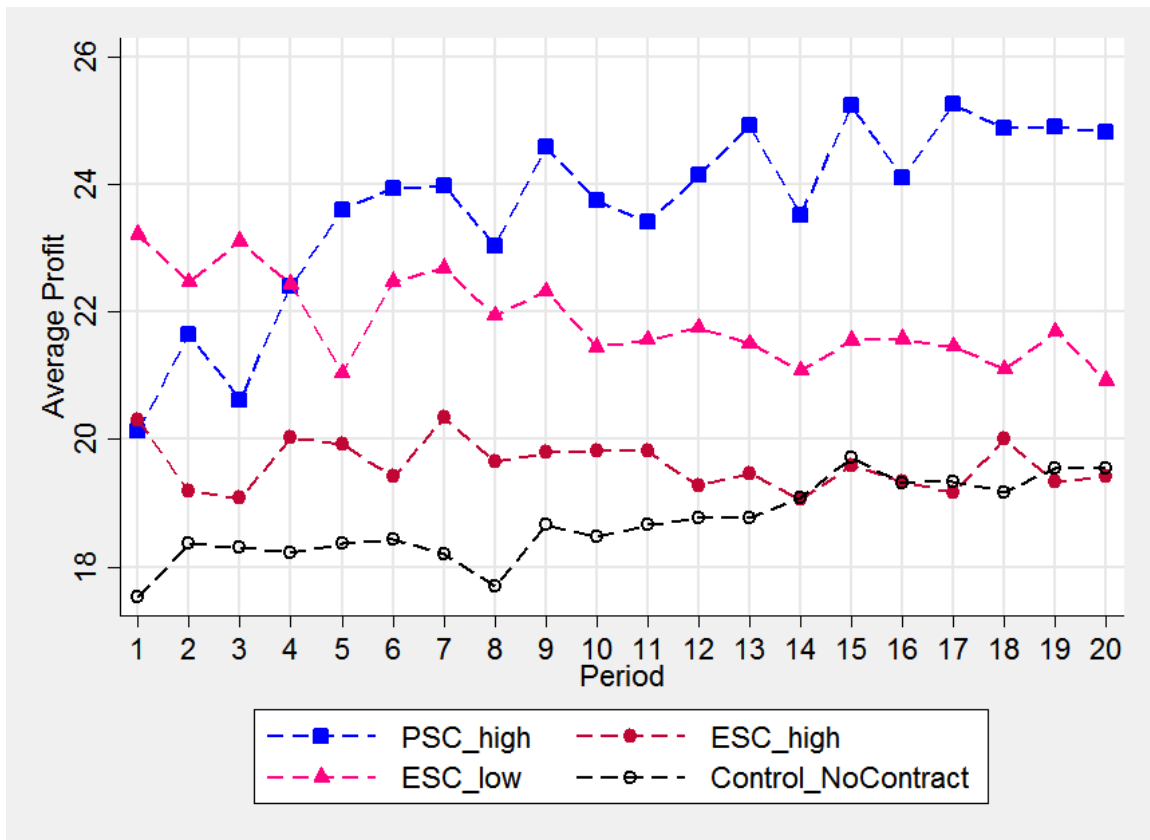


Figure 4.3.1: Average profit by period and treatment

Standard theory predicts that rational players take a PSC with full investments and no players take an ESC as long as the contract is costly. On the equilibrium path, subjects would make a profit of 27 points in the PSC_high treatment and 20 points in the remaining treatments, by just keeping their endowment. With 20 rounds of repetition the trend of the dynamics across treatments follows the equilibrium prediction. What seems interesting is that the two ESC treatments obtain a higher level of individual average profit compared to the control treatment.

Next we consider a hypothetical third party who enforces a contract and receives the cor-

⁵We only use our data up to 20 periods to match the data from the control in Ralph’s paper.

responding implementation fee. In this sense the cost of contract is not a welfare loss but a transfer to this third party. We adjust the average surplus by adding the contract cost back to the average profit if a contract is carried out. In this way we calculate social welfare for the three contract treatments as if there were no contract cost. Table 4.3.1 summarizes the two measures of social welfare, i.e. the average profit and the average surplus over matching groups, for different treatments.

Table 4.3.1: Average profit and average contribution

Treatment	Control	ESC_high	ESC_low	PSC_high
Average Profit	18.78	19.58	21.94	23.65
Average Surplus	18.78	21.40	22.52	27.21
Sample Size	13	16	16	17

Not surprisingly, the highest social welfare is achieved in the PSC_high where both the social dilemma and the hold-up problem are solved. However, the realized social welfare is lower than the predicted equilibrium outcome (i.e. 27 in terms of the average profit and 32 in terms of the average surplus). We perform a two-sided Wilcoxon rank-sum test (i.e. MWW test) on both the profit and the surplus averaged over all periods and subjects in a matching group. Remarkably, even the small difference of 0.8 points in the average profit is significant ($p=0.0353$) between the control treatment and the ESC_high treatment. A pair-wise comparison across the contract treatments further confirms the significant ($p<0.01$) difference in average profit. When we add the implementation cost of a contract back to the average profit, the ranking of the average surplus remains unaltered. However, the difference in average surplus between the ESC_high treatment and the ESC_low treatment is no longer significant ($p= 0.3461$, MWW test, two-sided).

4.3.2 Investment Stage: Contributions

Contributions have a positive effect on social welfare if contracts are signed. By committing to a contract in advance the hold-up problem is removed from all three contract treatments. In the PSC_high treatment the social dilemma is also removed by the conditional proportional sharing rule. Firstly, we compare the aggregate contributions across treatments by averaging the data within each matching group. As expected, subjects on average contribute significantly more (15.09) in the PSC_high treatment. The mean value of individual contributions in the ESC_high treatment and the ESC_low treatment (6.34 and 6.69 respectively) is similar to that in the control treatment (6.57). A two sided MWW test fails to reject the null hypothesis of no difference compared to the control treatment ($p=0.6610$ in the ESC_high and $p=0.7924$ in the ESC_low).

Next we compare the aggregate distribution of contributions across treatments by box plots drawn in Figure 4.3.2. Compared to the control treatment the distribution is skewed above the median in both ESC treatments. The increased skewness in the distribution indicates that an equal split contract may have a positive effect on contributions among those who stay on a contract. We will investigate this possibility next.

4.3.3 How does an equal split contract affect contributions?

Although an equal split contract does not solve the social dilemma in the cooperation game, committing to the contract at least removes the hold-up problem and future distributional conflicts. Figure 4.3.3 depicts the average contributions conditional on the outcome of contracting over time, i.e. either stay on a contract or enter into a fight. Interestingly, the average contributions of the subjects who stay on a contract is not only higher than those who enter a fight, but also stabilizes over time. Committing to a contract seems to prevent the decay pattern of contributions documented in public goods game experiments (Chaud-

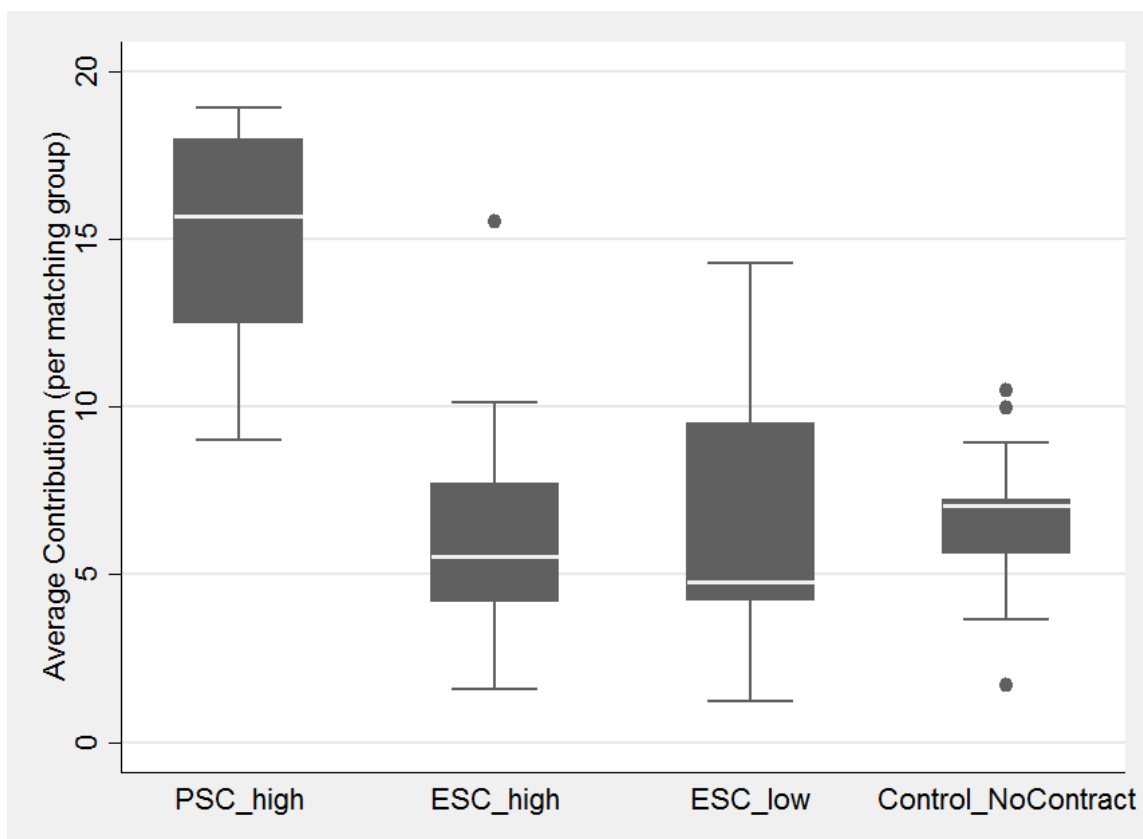


Figure 4.3.2: Average contribution by treatments

huri 2011). This is interesting, as an equal split contract transforms the remainder of the game to exactly the public goods game studied.

Next we plot the cumulative distribution of individual contributions conditional on the outcome of contracting for the two ESC treatments and the control treatment. As shown in Figure 4.3.4 there is a clear selection effect in the ESC_high treatment where more low contributors enter a fight and more high contributors stay on a contract. Note that the blue line of the control treatment partially overlaps with the dashed green line in the ESC_low treatment. This overlap suggests that low contributors also select themselves into the ESC if the cost is low. Does an ESC separate the higher contributors from low ones via self-selection, or does committing to a contract facilitate individual contributions? To address this question, we perform a regression analysis on individual level data.

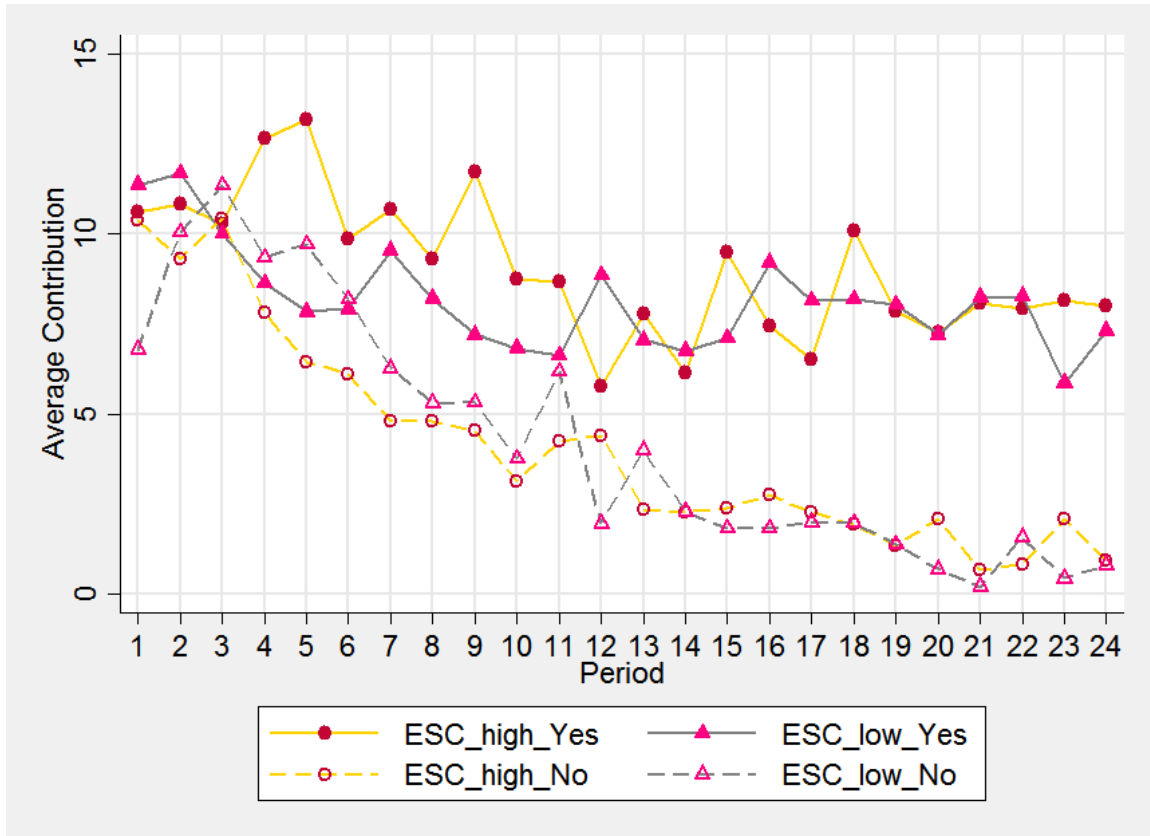


Figure 4.3.3: Average contribution conditional on the outcome of an equal split contract

To isolate the self-selection effect, we need an instrument to identify the pure contract effect on individual contributions. Before the outcome of contracting is revealed, subjects independently express their interests for a contract, i.e. opt in or not. A new partner is re-matched randomly in each round. We exploit the variation in the partner's intention for a contract as an instrument for the outcome of having a contract. We interact the choice dummy of individual i , i.e. $OptIn_i$, with the choice dummy of his or her partner j , i.e. $OptIn_j$. A contract is carried out only when both players in a match mutually opt in. When i opts in for a contract, implementation of the contract is determined by the choice of i 's partner j . If there is only a selection effect, then we would expect subjects who opt in to contribute more regardless of the outcome of contracting.

Let us first check our data structure before specifying the regression model. We allowed

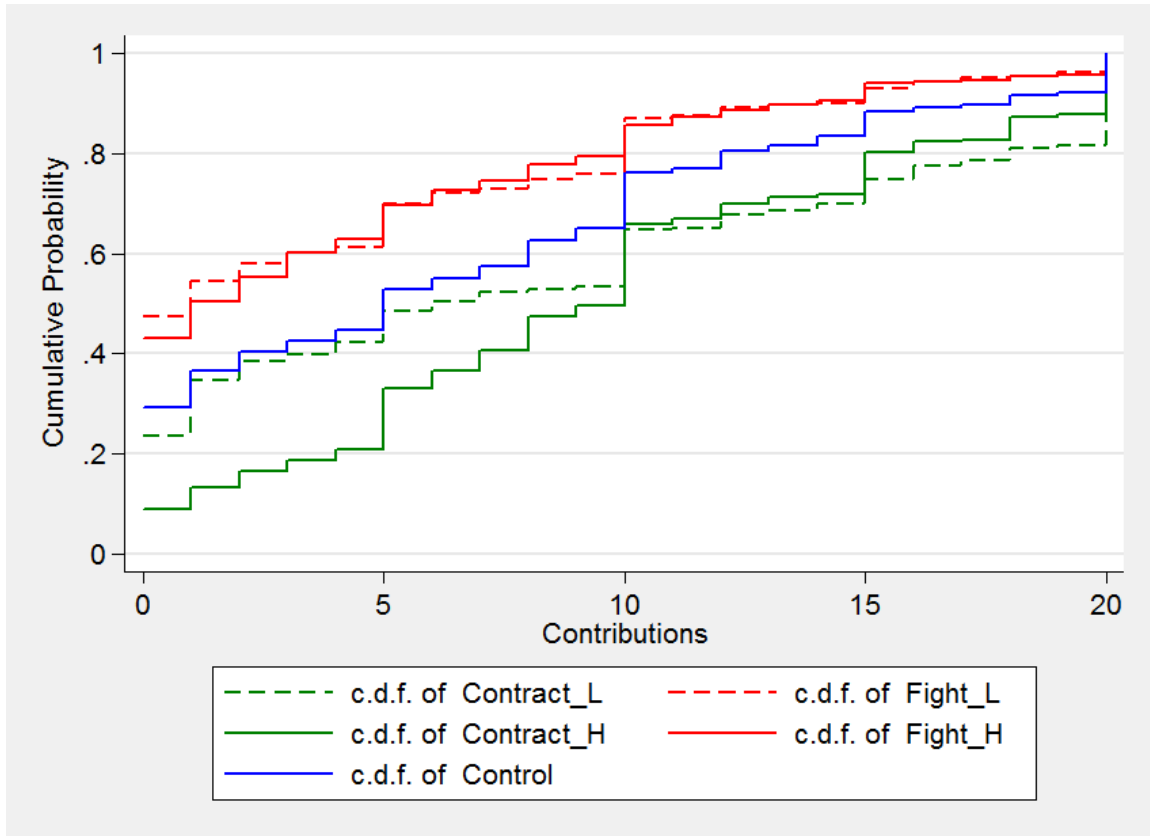


Figure 4.3.4: CDF of individual contributions conditional on the outcome of contracting 24 periods of repetition with feedback of all relevant information of both group members, including individual decision to accept a contract; contribution level; share of the proceeds and effort level if a contest stage follows, in each round and employed a subgroup random matching among every four players if possible. Thus individual data may be correlated along the time dimension and within a subgroup if we pool the data. In order to take account of the correlation, we run an OLS regression using a clustering option at the individual level and the matching group level respectively. This allows us to have a robust estimation of a significant effect.

The dependent variable is the individual contribution, i.e. $Contribution_i$. To avoid the dynamic panel bias, we refrain from using any lagged terms of the dependent variable as a regressor. We use the lagged term of partner's contribution (i.e. $L1Contribution_j$) to

control for past experience. We further include a set of dummy variables for each period (i.e. Period_t , $t=1,\dots,24$) to control for learning effect over time. Individual characteristic controls such as gender, age and math training are also included. We run the regression separately for the ESP_high treatment and the ESP_low treatment to see if the cost of an ESC has any effect on contributions.⁶ Table 4.3.2 reports the estimated coefficients for each clustering option across the two treatments. The negative coefficient for period dummies suggests that individual contributions erode over time. Controlling for this dynamic effect, we find a positive correlation between individual contribution and the past contribution of a randomly matched partner. The effect size in the ESC_high treatment is not significantly different from the the ESC_low treatment, i.e. 0.278 and 0.374 respectively.⁷

Our major interest is in the interaction terms. The difference between the fourth and the third categories captures the pure effect of a contract on the average individual contribution. Given that a subject opts for a contract, the contract actually being implemented increases contributions on average ($p < 0.01$ using F -statistic) by about 3 points in the ESC_high treatment. A similar positive effect is also observed in the ESC_low treatment. However, the effect in the ESC_low treatment becomes insignificant ($p = 0.1205$ and $F(1, 5) = 3.49$) with a clustering option at the subgroup level. A look at individual data reveals that the positive effect of a contract in the ESC_low treatment only happens within some matching groups. This explains the higher standard error of the estimated coefficients.

The difference between the second category and the base group captures the signal effect of demanding a contract. We have a significantly positive effect in the ESC_high treatment ($P < 0.05$ using a Wald test for the equality of two coefficients) but not in the ESC_low treatment. It seems that opting in for a contract with a high cost serves as a signal of good will to avoid future fights and thus increase contributions in the ESC_high treatment.

⁶We have also conducted a multilevel mixed-effects logistic regression by imposing more structure on error terms and have confirmed the robustness of our results.

⁷We use seemingly unrelated estimation in Stata to perform the chow test. The p-value is 0.1473 if clustering is on the individual level and 0.4537 if clustering is on the matching group level.

Table 4.3.2: Pooled OLS regressions for individual contributions

Clustering Treatment	at individual level		at subgroup level	
	ESC_high	ESC_low	ESC_high	ESC_low
OptIn_i × OptIn_j [base group: 1) i No, j No]				
2) i No, j Yes	1.309*** (0.491)	0.553 (0.528)	1.309** (0.469)	0.553 (0.519)
3) i Yes, j No	1.376** (0.570)	-0.817 (0.657)	1.376*** (0.261)	-0.817 (0.727)
4) i Yes, j Yes	4.687*** (0.570)	2.103** (0.826)	4.687*** (0.530)	2.103 (1.249)
L1Contribution_j	0.278*** (0.0499)	0.374*** (0.0456)	0.278*** (0.0747)	0.374*** (0.0758)
Controls [gender, age, math] are not significant				
Period dummies are negative significant for the 6th period and onwards.				
Constant	4.104*** (1.450)	4.903*** (1.457)	4.104** (1.601)	4.903*** (0.805)
Adjusted clusters	62	60	7	6
R-squared	0.375	0.306	0.375	0.306

Notes: Robust standard errors in parentheses ; *** p<0.01, ** p<0.05, * p<0.1 .

[1] Sample size is 1,410 in ESC_high and 1,380 in ESC_low. Constant term is included.

[2] Characteristics controls are all 0-1 dummies where age has a threshold value at 25 years old.

[3] Period dummies include 23 periods from the 2nd to the 24th.

4.3.4 Contest Stage: Efforts

If subjects fail to reach an agreement on a contract in advance, they enter into a contest to compete for their share of the proceeds. We know that the average effort of subjects is remarkably close to the subgame-perfect Nash continuation efforts in the two-stage game without a contract option (Bayer 2016b). Does the cost of a contract serve as a reference point to restrict the maximum efforts one would like to incur in the contest stage in our context? If it is the case, we would expect subjects on average to exert higher efforts in the ESC_high treatment than in the ESC_low treatment. If not, we would observe similar close to equilibrium play between the two ESC treatments.

We plot the individual data of the actual efforts against the predicted Nash efforts (i.e. a

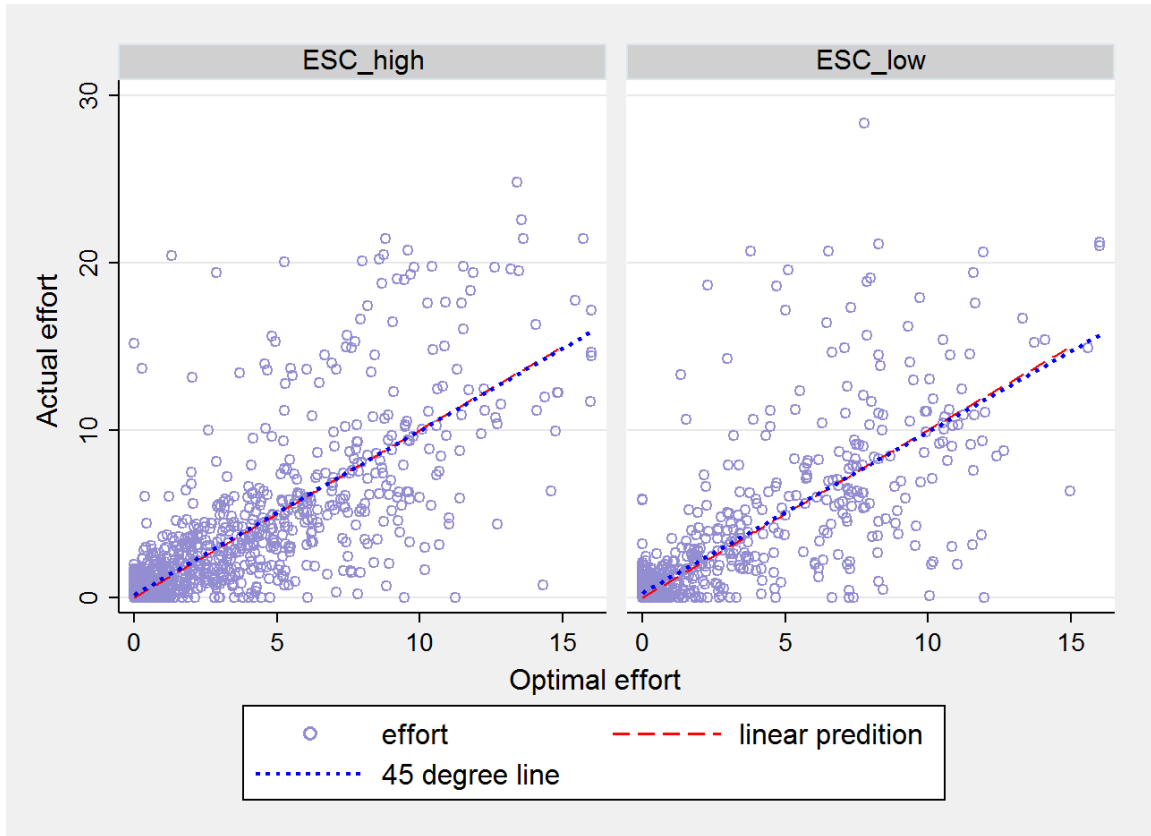


Figure 4.3.5: Actual efforts vs. Nash efforts in the ESC treatments

quarter of the realized proceeds) in Figure 4.3.5. The majority of the data scatters along the 45 degree line and the linear predicted line is very close to the 45 degree line. We replicate the robust pattern of equilibrium play on average in the contest game with a sharing Tullock rule (Dechenaux et al. 2015). It seems that the existence of a costly contract option in advance does not affect the average contest behavior in the later stage.

However, we wonder if subjects who self-select into a fight would fight harder for their fair share? Suppose a high contributor i meets a free-riding partner j in the cooperation game. Equal shares as in equilibrium would be extremely unfair for player i given his or her high contribution. Thus we expect player i to fight harder than her or his partner j in the contest game for a fair share. To test our conjecture, we run an OLS regression for individual efforts on individual contributions and their partner's contribution. Results of

the regression across treatments are reported in Table 4.3.3.

Table 4.3.3: An Pooled OLS Regression for Individual Efforts

Dependent	Treatments		
	ESC_low	ESC_high	Control
Own Contribution	0.555*** (0.061)	0.412*** (0.037)	0.416*** (0.019)
Partner's Contribution	0.243*** (0.043)	0.390*** (0.029)	0.327*** (0.015)
Observations	652	984	2,160
Adjusted clusters	60	62	621
R-squared	0.735	0.731	0.713

Notes: Robust standard errors in parentheses ; *** p<0.01, ** p<0.05, * p<0.1 ..

[1] In the ESC treatments we include 24 periods data for subjects who enter a contest game.

[2] In the control treatment every one enters a contest game with 20 rounds of repetition.

Interestingly, the coefficient of the Own Contribution variable is double that of the coefficient for the Partner's Contribution variable in the ESC_low treatment. A Wald test confirms the significant difference between the two coefficients. The difference in the slope suggests that subjects fight harder if they contribute more in the ESC_low treatment (p=0.0017). The difference in the slope becomes weaker in the control treatment (p=0.0241) and vanishes in the ESC_high treatment (p=0.6798). It seems that the free-riding issue is more severe in the ESC_low treatment compared to the ESC_high treatment.

In summary, we know that an ESC has welfare enhancing effects and the wasteful efforts incurred in the contest game are close to the equilibrium amounts on average. We conclude that the higher welfare achieved in the two ESC treatments stems from high contributors who can avoid a fight by committing to a contract. We now switch our attention to contracting behavior.

4.3.5 Frequency of Contracting

In the PSC_high treatment 72% of groups use a contract to resolve future distributional conflicts while in both the ESC_high treatment and the ESC_low treatment only about 33% and 55% of groups implement a contract respectively. A pairwise rank sum test across treatments suggests the difference in average frequency of contracting is significant ($p < 0.05$, two-sided). The percentages are not surprising because the PSC in theory solves both the social dilemma and the hold-up problem, but the ESC leaves the social dilemma unsolved. Figure 4.3.6 depicts the dynamics of contracting frequencies over time and across the three treatments.

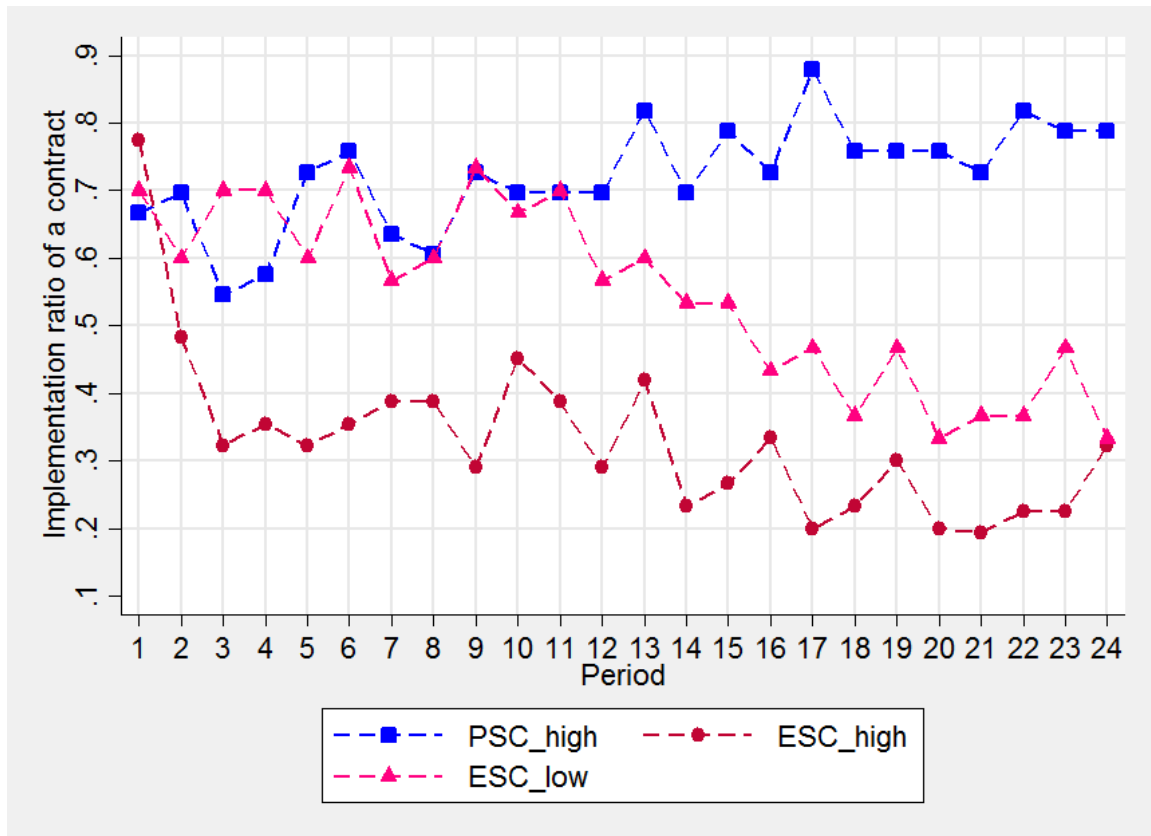


Figure 4.3.6: Frequency of contracting over periods by treatments

The majority of subjects start by trying to implement a contract during the initial periods. There is a further upwards trend in the PSC as people learn that contracts work well. For

the ESC the trend is downwards, while the decay is faster if the contract costs are high. It is interesting to point out that a low cost ESC helps to delay the decline but does not entirely prevent it. We are interested in what drives people to switch out of contracts as the experiment progresses.

4.3.6 What drives people into a distributive fight?

Since an ESC does not solve the social dilemma in the cooperation game, subjects always have an incentive to play the dominant strategy to free ride. Suppose your partner invests nothing while you fully invest. By implementing the equal split sharing contract, you would get only 16 points back for the 20 points invested and be worse off than you initially were endowed. A distributional contest at least gives you a chance to fight an unfair allocation on the other hand. So we suspect the difference in group contributions is the driving factor of the decline of contracting in the ESC_low treatment.

We take the difference between the individual contribution and their partner's contribution to measure the absolute uneven degree at a group level. In order to adjust for the fact that contributions declines over time, we divide the absolute degree by the sum of the two group members' contribution. This gives us a measure of the relative uneven degree in group contributions. If only one player makes a contribution, then the relative uneven degree attains the maximum value 1. If both players contribute equally, then the relative uneven degree reaches the minimum value 0. Figure 4.3.7 graphs the cumulative distribution of the relative uneven degree in contributions conditional on the outcome of contracting (i.e. on a contract or on a fight) for the two treatments.

As shown in Figure ?? the solid green distribution lies beneath (i.e. first-order stochastically dominates) the dashed red distribution everywhere. This demonstrates a strong uneven contribution pattern in the ESC_low treatment. At every percentile the degree of uneven-

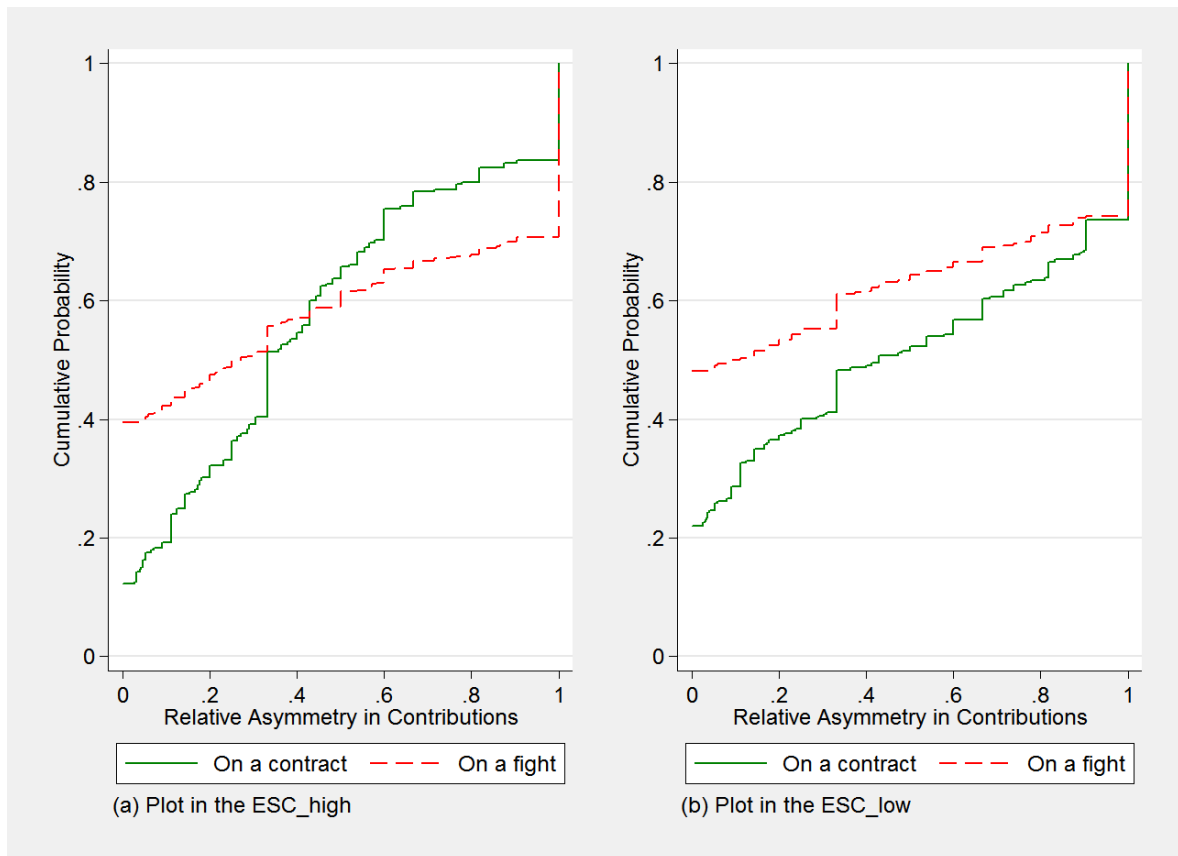


Figure 4.3.7: CDF of relative uneven degree in contributions by outcome of contracting. The plot shows that the degree of relative unevenness in group contributions is more severe when a contract takes place than when a contest follows. In contrast no such patterns are observed in the ESC_high treatment.

Next we run a probit regression on individual contracting decisions (i.e. 1 go for a fight and 0 demand a contract) with a clustering option at the individual level. We use past contributions of the randomly matched partners as a proxy to indicate how severe the free-riding issue is. Table 4.3.4 reports the estimated coefficients and average marginal effects. The strongly significant negative sign of the coefficients of the lagged terms confirms our conjecture. The less one's partner has contributed in the past, the more likely an individual chooses to enter a fight for a fair share. The negative effect is persistent over time. On average subjects increase the chance to enter a fight by 1% if their previous partner has decreased contributions by 1 point holding other factors constant.

Table 4.3.4: A random-effects Probit model of entering a contest

	Coefficients	Average Marginal Effects
PartnerContribution_L1	-0.047*** (0.008)	-0.013*** (0.002)
PartnerContribution_L2	-0.023*** (0.008)	-0.006*** (0.002)
PartnerContribution_L3	-0.023*** (0.007)	-0.006*** (0.002)
PartnerContribution_L4	-0.036*** (0.008)	-0.010*** (0.002)
Math_training (Yes)	-0.594** (0.238)	-0.178** (0.075)
Age (Above 25)	0.222 (0.183)	0.061 (0.049)
Gender (Male)	-0.172 (0.174)	-0.049 (0.050)
Constant	0.460 (0.302)	
Pseudo R2	0.173	

Notes: Robust standard errors in parentheses ; *** p<0.01, ** p<0.05, * p<0.1 .
 [1] Number of observations is 1,200 and the number of subject clusters is 60.

4.4 Conclusion

In this paper we study the role of a contract option to resolve future distributional conflicts. We extend Bayer's two-stage cooperation and contest game by proving a costly contract option. In Bayer's model a second stage contest game is employed to decide how to split the proceeds of the first stage cooperation game. However, welfare gains from positive contributions are wiped out by the costly contest investments needed to fight over the share of the proceeds. In order to remedy the waste of resources in distributional fights we offer a contract option to commit to a sharing rule in advance. In our first treatment, we offer an expensive contract that splits the proceeds of cooperation equally. This ESC does not remove the social dilemma in the cooperation game. Interestingly, we find that contracts have a

positive effect on individual contributions. However, the average frequency of contracting is only about 33%.

In a further treatment, we change the nature of the contract by specifying a conditional proportional split sharing rule. The PSC aligns the incentive of individuals to fully invest in the cooperation game and thus removes the social dilemma as well as the hold-up problem. Over time, subjects learn this and converge to equilibrium which results in the highest social welfare. Along the second dimension, we decrease the implementation cost of an ESC. A lower cost increases the frequency of contracting to about 55%. However, lower cost contracts also attract more low contributors to sign contracts. Remarkably, we find that the ESC can stabilize the average contributions over time among subjects who keep committing to a contract. The significant welfare enhancing effect of the ESC comes from the subjects who have a preference for contribution and can avoid future distributional fights by opting for a contract.

In general, a contract option helps to resolve future distributional conflicts. However, the magnitude of the welfare enhancing effect depends on the nature and the cost of a contract. Contracts that condition on individual contribution levels might not be feasible or more expensive than simple share contracts. One possible future extension of our work is to study in the situation where subjects can choose between the PSC with a high cost or the ESC with a low cost. We observe heterogeneous behavior in the cooperation game when the social dilemma exists. A further efficiency improvement is still possible if the higher contributors implement the low cost ESC and the low contributors commit to the high cost PSC. It would be interesting to see if subjects are able to choose the right mechanism that favors their type. This idea bears similarities to the experimental study on endogenous choice of various voting mechanisms (Engelmann & Gruner 2013).

Bibliography

- Anderhub, V., Gächter, S. & Königstein, M. (2002), 'Efficient contracting and fair play in a simple principal-agent experiment', *Experimental Economics* 5(1), 5–27.
- Ariely, D. & Simonson, I. (2003), 'Buying, bidding, playing, or competing? value assessment and decision dynamics in online auctions', *Journal of Consumer Psychology* 13(1-2), 113–123.
- Aronson, J., Quinn, D. M. & Spencer, S. J. (1998), 'Stereotype threat and the academic underperformance of minorities and women.', (*Mimeo*).
- Bargh, J. A. & Pietromonaco, P. (1982), 'Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation.', *Journal of Personality and Social Psychology* 43(3), 437.
- Bartling, B., Fehr, E. & Herz, H. (2014), 'The intrinsic value of decision rights', *Econometrica* pp. 2005–2039.
- Bartling, B. & Fischbacher, U. (2011), 'Shifting the blame: On delegation and responsibility', *Review of Economic Studies* p. rdr023.
- Bayer, R.-C. (2016a), 'Cooperation and distributive conflict', *Games and Economic Behavior* 97, 88–109.
- Bayer, R.-C. (2016b), 'Cooperation in partnerships: The role of breakups and reputation', *Journal of Institutional and Theoretical Economics* 172(4), 615–638.
- Becker, G. M., DeGroot, M. H. & Marschak, J. (1964), 'Measuring utility by a single-response sequential method', *Behavioral Science* 9(3), 226–232.
- Berg, J., Dickhaut, J. & McCabe, K. (1995), 'Trust, reciprocity, and social history', *Games and Economic Behavior* 10(1), 122–142.
- Bolton, G. E. & Ockenfels, A. (2000), 'Erc: A theory of equity, reciprocity, and competition', *American Economic Review* pp. 166–193.
- Bosman, R., Sonnemans, J., Zeelenberg, M. et al. (2001), 'Emotions, rejections, and cooling off in the ultimatum game', (*Mimeo*).

- Cai, H. & Wang, J. T.-Y. (2006), 'Overcommunication in strategic information transmission games', *Games and Economic Behavior* **56**(1), 7–36.
- Cardella, E. & Chiu, R. (2012), 'Stackelberg in the lab: The effect of group decision making and cooling-off periods', *Journal of Economic Psychology* **33**(6), 1070–1083.
- Charness, G. (2004), 'Attribution and reciprocity in an experimental labor market', *Journal of Labor Economics* **22**(3), 665–688.
- Charness, G., Cobo-Reyes, R., Jiménez, N., Lacomba, J. A. & Lagos, F. (2012), 'The hidden advantage of delegation: Pareto improvements in a gift exchange game', *American Economic Review* **102**(5), 2358–2379.
- Charness, G., Cobo-Reyes, R., Lacomba, J. A., Lagos, F. & Pérez, J. M. (2016), 'Social comparisons in wage delegation: Experimental evidence', *Experimental Economics* **19**(2), 433–459.
- Chaudhuri, A. (2011), 'Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature', *Experimental Economics* **14**(1), 47–83.
- Cooper, D. J. & Fang, H. (2008), 'Understanding overbidding in second price auctions: An experimental study', *Economic Journal* **118**(532), 1572–1595.
- Crawford, V. P. & Sobel, J. (1982), 'Strategic information transmission', *Econometrica* pp. 1431–1451.
- Croson, R. & Gneezy, U. (2009), 'Gender differences in preferences', *Journal of Economic Literature* **47**(2), 448–474.
- Dechenaux, E., Kovenock, D. & Sheremeta, R. M. (2015), 'A survey of experimental research on contests, all-pay auctions and tournaments', *Experimental Economics* **18**(4), 609–669.
- Dessein, W. (2002), 'Authority and communication in organizations', *Review of Economic Studies* **69**(4), 811–838.
- Dickhaut, J. W., McCabe, K. A. & Mukherji, A. (1995), 'An experimental study of strategic information transmission', *Economic Theory* **6**(3), 389–403.
- Dufwenberg, M. & Kirchsteiger, G. (2004), 'A theory of sequential reciprocity', *Games and Economic Behavior* **47**(2), 268–298.
- Ehrhart, K.-M., Ott, M. & Abele, S. (2015), 'Auction fever: Rising revenue in second-price auction formats', *Games and Economic Behavior* **92**, 206–227.
- Engelmann, D. & Gruner, H. P. (2013), 'Tailored bayesian mechanisms: Experimental evidence from two-stage voting games', *CEPR Discussion Paper, (No.9544)*. .
- Falk, A. & Fischbacher, U. (2006), 'A theory of reciprocity', *Games and Economic Behavior* **54**(2), 293–315.

- Fehr, E., Hart, O. & Zehnder, C. (2015), 'How do informal agreements and revision shape contractual reference points?', *Journal of the European Economic Association* **13**(1), 1–28.
- Fehr, E., Herz, H. & Wilkening, T. (2013), 'The lure of authority: Motivation and incentive effects of power', *American Economic Review* **103**(4), 1325–1359.
- Fehr, E., Klein, A. & Schmidt, K. M. (2007), 'Fairness and contract design', *Econometrica* **75**(1), 121–154.
- Fershtman, C. & Gneezy, U. (2001), 'Strategic delegation: An experiment', *RAND Journal of Economics* pp. 352–368.
- Fischbacher, U. (2007), 'z-tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics* **10**(2), 171–178.
- Fischbacher, U. & Föllmi-Heusi, F. (2013), 'Lies in disguise-an experimental study on cheating', *Journal of the European Economic Association* **11**(3), 525–547.
- Frey, B. S. & Jegen, R. (2001), 'Motivation crowding theory', *Journal of Economic Surveys* **15**(5), 589–611.
- Gibbons, R. (2005), 'Four formal (izable) theories of the firm?', *Journal of Economic Behavior & Organization* **58**(2), 200–245.
- Gilchrist, D. S., Luca, M. & Malhotra, D. (2016), 'When $3+1 > 4$: Gift structure and reciprocity in the field', *Management Science* .
- Gneezy, U. (2005), 'Deception: The role of consequences', *American Economic Review* **95**(1), 384–394.
- Gneezy, U., Rockenbach, B. & Serra-Garcia, M. (2013), 'Measuring lying aversion', *Journal of Economic Behavior & Organization* **93**, 293–300.
- Greiner, B. (2015), 'Subject pool recruitment procedures: organizing experiments with orsee', *Journal of the Economic Science Association* **1**(1), 114–125.
- Grimm, V. & Mengel, F. (2011), 'Let me sleep on it: Delay reduces rejection rates in ultimatum games', *Economics Letters* **111**(2), 113–115.
- Grossman, S. J. & Hart, O. D. (1986), 'The costs and benefits of ownership: A theory of vertical and lateral integration', *Journal of Political Economy* **94**(4), 691–719.
- Hamman, J. R., Loewenstein, G. & Weber, R. A. (2010), 'Self-interest through delegation: An additional rationale for the principal-agent relationship', *American Economic Review* **100**(4), 1826–1846.
- Hart, O. & Moore, J. (1990), 'Property rights and the nature of the firm', *Journal of Political Economy* **98**(6), 1119–1158.

- Herbst, L., Konrad, K. A. & Morath, F. (2016), 'Balance of power and the propensity of conflict', *Games and Economic Behavior* .
- Heyman, J. E., Orhun, Y. & Ariely, D. (2004), 'Auction fever: The effect of opponents and quasi-endowment on product valuations', *Journal of Interactive Marketing* **18**(4), 7–21.
- Holmström, B. R. (1977), *On incentives and control in organizations*, Doctoral Dissertation, Stanford University.
- Hoppe, E. I. & Schmitz, P. W. (2011), 'Can contracts solve the hold-up problem? experimental evidence', *Games and Economic Behavior* **73**(1), 186–199.
- Hurkens, S. & Kartik, N. (2009), 'Would i lie to you? on social preferences and lying aversion', *Experimental Economics* **12**(2), 180–192.
- Izmalkov, S. (2001), 'English auctions with reentry', *Working Paper, Pennsylvania State University* .
- Jones, M. T. (2011), 'Bidding fever in ebay auctions of amazon. com gift certificates', *Economics Letters* **113**(1), 5–7.
- Kagel, J. H. & Levin, D. (1993), 'Independent private value auctions: Bidder behaviour in first-, second- and third-price auctions with varying numbers of bidders', *Economic Journal* **103**(419), 868–879.
- Kagel, J. H. & Roth, A. E. (2016), *Handbook of Experimental Economics, Volume 2: The Handbook of Experimental Economics*, Princeton university press.
- Kahneman, D., Knetsch, J. L. & Thaler, R. H. (1990), 'Experimental tests of the endowment effect and the coase theorem', *Journal of political Economy* pp. 1325–1348.
- Kahneman, D. & Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* pp. 263–291.
- Kaiser, J. (2007), 'An exact and a monte carlo proposal to the fisher–pitman permutation tests for paired replicates and for independent samples', *Stata Journal* **7**(3), 402–412.
- Kamecke, U. (1998), 'Dominance or maximin: How to solve an english auction', *International Journal of Game Theory* **27**(3), 407–426.
- Katz, M. L. (1991), 'Game-playing agents: Unobservable contracts as precommitments', *The RAND Journal of Economics* pp. 307–328.
- Kimbrough, E. O., Rubin, J., Sheremeta, R. M. & Shields, T. W. (2015), 'Commitment problems in conflict resolution', *Journal of Economic Behavior & Organization* **112**, 33–45.
- Kimbrough, E. O. & Sheremeta, R. M. (2013), 'Side-payments and the costs of conflict', *International Journal of Industrial Organization* **31**(3), 278–286.

- Klein, B., Crawford, R. G. & Alchian, A. A. (1978), 'Vertical integration, appropriable rents, and the competitive contracting process', *Journal of Law and Economics* **21**(2), 297–326.
- Knetsch, J. L. (1989), 'The endowment effect and evidence of nonreversible indifference curves', *American Economic Review* pp. 1277–1284.
- Kockesen, L. & Ok, E. A. (2004), 'Strategic delegation by unobservable incentive contracts', *Review of Economic Studies* **71**(2), 397–424.
- Kube, S., Maréchal, M. A. & Puppea, C. (2012), 'The currency of reciprocity: Gift exchange in the workplace', *American Economic Review* **102**(4), 1644–1662.
- Lai, E. K. & Lim, W. (2012), 'Authority and communication in the laboratory', *Games and Economic Behavior* **74**(2), 541–560.
- Lange, A. & Ratan, A. (2010), 'Multi-dimensional reference-dependent preferences in sealed-bid auctions—how (most) laboratory experiments differ from the field', *Games and Economic Behavior* **68**(2), 634–645.
- Ledyard, J. (1995), *Public goods: a survey of experimental research*. In: Kagel, J.H., Roth, A.E. (Eds) *The Handbook of Experimental Economics*, Princeton University Press Princeton, New Jersey.
- List, J. A. (2003), 'Does market experience eliminate market anomalies?', *Quarterly Journal of Economics* **118**(1), 41–71.
- López-Pérez, R. & Spiegelman, E. (2013), 'Why do people tell the truth? experimental evidence for pure lie aversion', *Experimental Economics* **16**(3), 233–247.
- Lundquist, T., Ellingsen, T., Gribbe, E. & Johannesson, M. (2009), 'The aversion to lying', *Journal of Economic Behavior & Organization* **70**(1), 81–92.
- Malmendier, U. & Lee, Y. H. (2011), 'The bidder's curse', *American Economic Review* pp. 749–787.
- McAfee, R. P. & McMillan, J. (1987), 'Auctions and bidding', *Journal of Economic Literature* pp. 699–738.
- Milgrom, P. R. & Weber, R. J. (1982), 'A theory of auctions and competitive bidding', *Econometrica* pp. 1089–1122.
- Niederle, M. & Vesterlund, L. (2007), 'Do women shy away from competition? do men compete too much?', *Quarterly Journal of Economics* **122**(3).
- Owens, D., Grossman, Z. & Fackler, R. (2014), 'The control premium: A preference for payoff autonomy', *American Economic Journal: Microeconomics* **6**(4), 138–161.
- Pierce, J. L., Kostova, T. & Dirks, K. T. (2003), 'The state of psychological ownership: Integrating and extending a century of research.', *Review of General Psychology* **7**(1), 84.

- Plott, C. R. & Zeiler, K. (2005), 'The willingness to pay–willingness to accept gap, the', *The American Economic Review* **95**(3), 530–545.
- Reuben, E., Rey-Biel, P., Sapienza, P. & Zingales, L. (2012), 'The emergence of male leadership in competitive environments', *Journal of Economic Behavior & Organization* **83**(1), 111–117.
- Reuben, E., Sapienza, P. & Zingales, L. (2014), 'How stereotypes impair women's careers in science', *Proceedings of the National Academy of Sciences* **111**(12), 4403–4408.
- Sánchez-Pagés, S. & Vorsatz, M. (2007), 'An experimental study of truth-telling in a sender–receiver game', *Games and Economic Behavior* **61**(1), 86–112.
- Siegel, S. (1956), 'Nonparametric statistics for the behavioral sciences.'
- Spence, M. (1973), 'Job market signaling', *Quarterly Journal of Economics* **87**(3), 355–374.
- Thaler, R. (1980), 'Toward a positive theory of consumer choice', *Journal of Economic Behavior & Organization* **1**(1), 39–60.
- Thaler, R. H. & Johnson, E. J. (1990), 'Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice', *Management Science* **36**(6), 643–660.
- Tversky, A. & Kahneman, D. (1991), 'Loss aversion in riskless choice: A reference-dependent model', *Quarterly Journal of Economics* **106**(4), 1039–1061.
- Vickrey, W. (1961), 'Counterspeculation, auctions, and competitive sealed tenders', *The Journal of Finance* **16**(1), 8–37.
- Williamson, O. E. (1971), 'The vertical integration of production: market failure considerations', *American Economic Review* **61**(2), 112–123.
- Williamson, O. E. (1979), 'Transaction cost economics: The governance of contractual relations', *Journal of Law and Economics* **22**(2), 223–261.

Appendix Samples of Instructions

Instructions

Introduction

Welcome to the second experiment! This time you will participate in an auction, in which you will act as a bidder for the purchase of a box of chocolates.

We will invite you to have a taste of the chocolate before the auction. A box contains 150 grams of chocolates (12 Balls).

You will be paired with two other participants in a group. The three of you then will bid in an auction (like the ones used to auction off houses) for a box of *Lindt* chocolates. The winner will receive the box of chocolates but will pay the winning bid in Australian Dollars. (We will deduct the winner's bid from the earning in the previous experiment.) Unsuccessful bidders do not receive the chocolates but will not have to pay anything.

Auction rules

There is a clock on the upper left of your computer screen indicating the remaining time of the auction. The initial time has been set to 30 seconds. The auction ends when the clock runs out. Any new bid restarts the clock. (In a sense the clock is the auctioneer. When the clock runs out

the hammer falls.) The bidder holding the highest bid, when the clock runs out is the winner.

You can submit a bid at any time, as long as the clock has not run out yet. The only restriction is that your bid has to be higher than the currently highest bid. You will be shown all previous bids ordered from the highest to the lowest.

Remember that the bid you enter is the amount of money you have to pay in exchange for the box of chocolates if you win the auction. You can enter any amount. Note that the bids are in Australian Dollars. So if you enter, e.g. 2.3, then this means that you bid 2 Dollars and 30 cents. The minimal increment of bids is 0.1 Dollar (i.e. 10 Cents).

Auction outcomes

The bidder with the highest bid in the end of the auction receives the box of chocolates and pays the winning bid in real money. The payment for the winning bid will be deducted from the amount of money being earned in the previous experiment. The bidders who do not win the auction do not need to pay and get nothing in this experiment.

Here is the screenshot of our auction experiment:

The remaining time of the auction: sec

- The clock is restarted once a new bid has been entered.
- The auction ends when the clock runs out!

•A list of all bids submitted so far:

Bidder ID	Bids
-----------	------

Please enter your bid (in AUD) here:

Are there any questions?

Welcome

We are pleased to welcome you to the experiment today. During the experiment you will earn money depending on your actions. Communicating with other participants during the experiment is not allowed and will lead to exclusion from the experiment without payment.

The experiments today

In what follows we will ask you to complete a few tasks on the screen. The tasks are very simple and the instructions are given on the screen. You can make use of the provided pen and scratch paper during the tasks. Please raise your arm if you have a question and we will come to you and answer your question.

Don't turn over as yet!!!

A screen shot of the adding-up task is shown as below:

#	#	#	#	#	The Sum
					<input type="text"/>
Submit this sum as your final result					<input type="submit" value="Submit"/>

Your correct answers so far:

Your wrong answers so far:

Remaining Time [sec]:

Instructions (freeNoinfo)

So far you have completed one adding-up task for 7 minutes where you have earned one Dollar per correct answer. Your **performance** in the adding-up task refers to the number of correct answers given during a task.

Overview

In this part of the experiment, you will be randomly grouped with two other participants before you play a **delegation game**. In the delegation game one person (the **Delegator**) decides if he¹ wants the performance of somebody else to count instead of his own performance. So a delegator can make a nice profit if he chooses a person who performs much better than himself. In the meantime the other group members try to convince the Delegator to choose them, since being chosen results in a handy bonus.

Delegation game

As part of the delegation game the adding-up task you have already performed will be repeated. We call the adding-up task you have done already **Task 1** and the repetition you will do **Task 2**. Everybody will be repeating the adding up task and will be paid again one Dollar per correct answer.

However, one participant of the group will be randomly chosen to become the **Delegator**. The **Delegator** can decide if he wants to transfer the responsibility for his payoff from **Task 1** to any of the two other players (this is called to delegate). In each group there are one Delegator and two non-delegator players.

If a **Delegator** decides to delegate, then his performance from the completed **Task 1** will be replaced by the performance of the chosen person in the upcoming adding-up **Task 2**. So the Delegator's earnings from **Task 1** will be replaced by one Dollar for each correct answer the chosen person achieves in **Task 2**. A **Delegator** will pay a bonus of four dollars to the person he delegates to. The chosen player will receive this bonus on top of his normal earnings from performing Task 2.

A **Delegator** who decides not to delegate will be paid according to his own performance in the already completed **Task 1**.

Before the Delegator has to decide if and to whom he is delegating, non-delegators will have to indicate if they are available to be chosen. Then the non-delegator can send a short paragraph up to 400 characters message to the Delegator. When the Delegator decides he will see the messages sent and the avatars (the little pictures you chose) but will have **no information** about the actual past performance of the non-delegator players.

¹ We adopt a male perspective instead of saying she or he to avoid the cumbersome expression.

After the delegation decision has been taken, it will be shown to all group members. Then all participants will do **Task 2** and again add up numbers for 7 minutes, for which the payment is one Dollar per correct answer.

Earnings summary

The computer will randomly decide if you will be paid your earnings from Task 1 **or** Task 2.

Earnings for Task 1

Below you can find the payoffs for Task 1 separately for the case where the Delegator chose to delegate and if not.

The three shaded cells show where payoffs differ depending on whether the **Delegator** delegates or not.

Earnings after delegation:

	<i>Delegator who delegated</i>	<i>Chosen player</i>	<i>Not chosen player</i>
<i>Earnings for Task 1</i>	1 Dollar * performance of the chosen player in Task 2	1 Dollar * own performance in Task 1	1 Dollar * own performance in Task 1
<i>Bonus</i>	- 4 Dollars bonus to the chosen player	+ 4 Dollars bonus for being chosen	--

Earnings without delegation:

	<i>Delegator who did not delegate</i>	<i>Other player</i>	<i>Other player</i>
<i>Earnings for task 1</i>	1 Dollar * own performance in Task 1	1 Dollar * own performance in Task 1	1 Dollar * own performance in Task 1
<i>Bonus</i>	--	--	--

Earnings for Task 2

Delegation does not influence the payoff from **Task 2**. Everybody will be paid 1 Dollar times the performance in **Task 2**.

Please turn over if you would like to see a summary of the sequence of decisions in the delegation following delegation game.

Sequence of decisions summary

Stage 1

Everyone is informed about their role either as a Delegator or a non-delegator player.

Stage 2

Non-delegator players decide whether they want to be considered by the Delegator as a person he can delegate to.

Stage 3

Non-delegators send messages, to the Delegator of their group.

Stage 4

Delegators receive messages but don't know if they are truthful. Then they see the avatars of the interested non-delegator players and make a delegation choice.

Stage 5

Everyone is informed the result of the delegation.

Stage 6

Everyone repeats the adding-up task.

Stage 7

Earnings are displayed.

Are there any questions before we begin?

Instructions

Welcome to the experiment! Before we start, please read the instructions carefully.

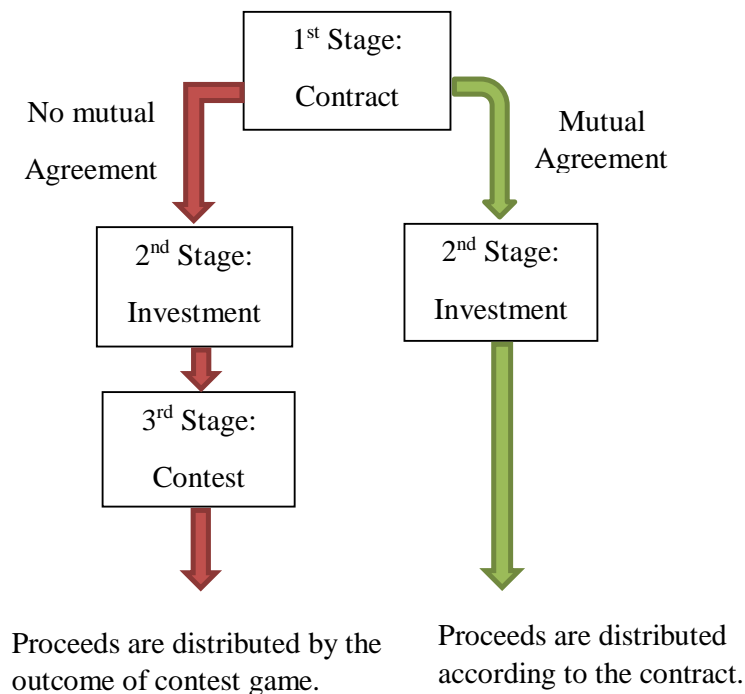
During the experiment, your earnings will be calculated in points rather than Dollars. Points are converted to Dollars at the following exchange rate at the end of the session to determine your payment:

20 Points = AUD 1.00

You will be paid in cash immediately after the experiment. You are not allowed to communicate with other participants during the experiment. If you have any questions, please raise your hand and we will attend to you individually. Failure to comply with the outlined rules will result in exclusion from the experiment and you will forfeit your payment.

Summary:

In this experiment you are asked to decide how much you and another group member want to invest in a group project. Your investments will jointly generate some **proceeds (value of the group project)**. There are two ways of splitting the value of the project. One is to mutually agree on a contract which specifies a binding sharing rule in advance. There is a fixed cost for implementing the contract. If you do not mutually agree on the contract, then you will be competing for a share of the value in a contest by exerting costly effort after the investment.



In the first stage a contract specifying a “50-50” equal sharing rule is offered to both players. Each player needs to decide whether to accept or decline.

In the second stage both players decide how much to invest in the group project.

If no agreement has been reached in the first stage, both players will enter a third stage to complete over the share of the proceeds. Each player needs to choose how much effort to invest in an attempt to acquire a share of the group project. The share each player receives in this case is determined by both players' efforts.

You will participate 24 of these investment and distribution games. In each of these you will be paired with a newly randomly chosen other subject. Whether you will enter the contest stage depends on the outcome of the contract stage. Only a mutual agreement on the contract can spare you both from the contest stage. The sequence of the game goes as follows:

The experiment in detail:

In what follows, we will explain the stages and how your payoff is determined in detail.

1st Stage: Contracting

You and your group member are offered to sign a contract that will implement an equal split of the proceeds of the group project.

The contract will only be implemented if both of you click accept. In that case both of you will pay a fixed fee of five points and there will be no third stage (contest stage). Instead of a contest with costly efforts determining the split of the project you will get half of the proceeds from the project.

If no mutual agreement has been made, the contract will not be carried out and no fee will be incurred. However, both of you will enter a third stage where your and your group member's costly effort determines the split of the value.

2nd Stage: Investment:

On the following page is a screen shot to familiarize you with how the investment stage will appear on your screen:



Your task is to divide your endowment (**20 points**) between what you keep for yourself and what you invest in the group project. The other group member has to do the same by choosing his or her investment at the same time as you.

The value of the group project depends on your investment and the investment of the other group member.

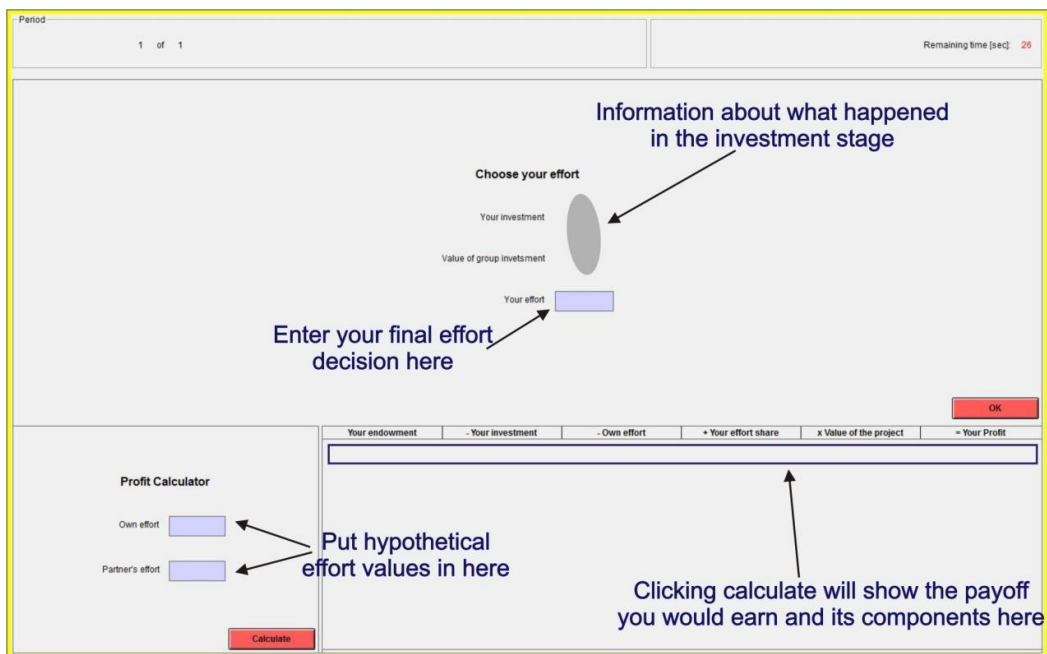
Once your investment has been made, you will be notified how much you and the other group member have individually invested in the project. The sum of your investments will be multiplied by **1.6** and that will be the total value of the project for that round. This means that in every round:

$$\text{Value of the project} = 1.6 \times (\text{your investment} + \text{other group member's investment})$$

3rd Stage: Distributional Contest (only if no contract was signed):

If the contract has not been agreed upon, then both of you enter the contest stage.

The following is a screen shot to familiarize you with what the distribution or effort stage will look like on your screen:



In this stage of the experiment, your task is to determine an amount of effort that you would like to invest in order to acquire a share of the group project. Your group member has to do the same.

The more effort you put in for a given level of the other group member's effort, the larger will be your share of the project, however, the higher will be your effort cost. On the other hand, the smaller your investment of effort is for the given effort of your group member, the smaller will be your share of the project, however, the effort cost you incur will also be low. The same is true for the other group member.

As a guide, on the back of these instructions is a table attached which represents values of percentage share of the project that you can expect to get for any given values of your own and your group member's effort.

In addition, you will be provided with a profit calculator on your screen (as visible in the screen shot above) which you can use to calculate what your expected profit will be for any combination of your own and the other group member's effort input.

Please note that the profit calculator is there only for your help. It does not affect your final profit in any way. You can play around with it using different values of effort for yourself and the other group member. You can then make your decision about what would be the **optimal** level of effort for you to put in.

Your payoff:

The total income you earn will be the sum of two parts:

1. Points that you keep (endowment - investment)
2. Your income from the group investment project.

The profit you earn will be the total income minus the cost incurred either by implementing the contract (the fixed fee) or by competing in the contest (the effort cost):

- ❖ If the contract has been carried out, then your total payoff at the end of each round is calculated as:

$$\text{Profit} = (\text{endowment} - \text{investment}) + \text{half share of the proceeds} - \text{5 points.}$$

- ❖ If the contract has not been implemented, then your total payoff at the end of each round is calculated as:

$$\text{Profit} = (\text{endowment} - \text{investment}) + \text{your share of proceeds} - \text{your effort.}$$

Your share of the group project depends on your and your group member's effort.

The other group member's profit is calculated in the same way.

After the 24th round your total profit will be recorded and you will be paid in cash.

The following table represents percentage share of the group project that you can expect to get for a given combination of your effort and your partner's effort. For example, if your effort input is 1 and your partner invests effort equal to 4, then you will get 20% of the group investment project. Note that your effort share (%) is calculated as follows:

		Other group member's effort																				
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Your effort	0	50.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	1	100.0%	50.0%	33.3%	25.0%	20.0%	16.7%	14.3%	12.5%	11.1%	10.0%	9.1%	8.3%	7.7%	7.1%	6.7%	6.3%	5.9%	5.6%	5.3%	5.0%	4.8%
	2	100.0%	66.7%	50.0%	40.0%	33.3%	28.6%	25.0%	22.2%	20.0%	18.2%	16.7%	15.4%	14.3%	13.3%	12.5%	11.8%	11.1%	10.5%	10.0%	9.5%	9.1%
	3	100.0%	75.0%	60.0%	50.0%	42.9%	37.5%	33.3%	30.0%	27.3%	25.0%	23.1%	21.4%	20.0%	18.8%	17.6%	16.7%	15.8%	15.0%	14.3%	13.6%	13.0%
	4	100.0%	80.0%	66.7%	57.1%	50.0%	44.4%	40.0%	36.4%	33.3%	30.8%	28.6%	26.7%	25.0%	23.5%	22.2%	21.1%	20.0%	19.0%	18.2%	17.4%	16.7%
	5	100.0%	83.3%	71.4%	62.5%	55.6%	50.0%	45.5%	41.7%	38.5%	35.7%	33.3%	31.3%	29.4%	27.8%	26.3%	25.0%	23.8%	22.7%	21.7%	20.8%	20.0%
	6	100.0%	85.7%	75.0%	66.7%	60.0%	54.5%	50.0%	46.2%	42.9%	40.0%	37.5%	35.3%	33.3%	31.6%	30.0%	28.6%	27.3%	26.1%	25.0%	24.0%	23.1%
	7	100.0%	87.5%	77.8%	70.0%	63.6%	53.5%	53.8%	50.0%	46.7%	43.8%	41.2%	38.9%	36.8%	35.0%	33.3%	31.8%	30.4%	29.2%	28.0%	26.9%	25.9%
	8	100.0%	88.9%	80.0%	72.7%	66.7%	61.5%	57.1%	53.3%	50.0%	47.1%	44.4%	42.1%	40.0%	38.1%	36.4%	34.8%	33.3%	32.0%	30.8%	29.6%	28.6%
	9	100.0%	90.0%	81.8%	75.0%	69.2%	64.3%	60.0%	56.3%	52.9%	50.0%	47.4%	45.0%	42.9%	40.9%	39.1%	37.5%	36.0%	34.6%	33.3%	32.1%	31.0%
	10	100.0%	90.9%	83.3%	76.9%	71.4%	66.7%	62.5%	58.8%	55.6%	52.6%	50.0%	47.6%	45.5%	43.5%	41.7%	40.0%	38.5%	37.0%	35.7%	34.5%	33.3%
	11	100.0%	91.7%	84.6%	78.6%	73.3%	68.8%	64.7%	61.1%	57.9%	55.0%	52.4%	50.0%	47.8%	45.8%	44.0%	42.3%	40.7%	39.3%	37.9%	36.7%	35.5%
	12	100.0%	92.3%	85.7%	80.0%	75.0%	70.6%	66.7%	63.2%	60.0%	57.1%	54.5%	52.2%	50.0%	48.0%	46.2%	44.4%	42.9%	41.4%	40.0%	38.7%	37.5%
	13	100.0%	92.9%	86.7%	81.3%	76.5%	72.2%	68.4%	65.0%	61.9%	59.1%	56.5%	54.2%	52.0%	50.0%	48.1%	46.4%	44.8%	43.3%	41.9%	40.6%	39.4%
	14	100.0%	93.3%	87.5%	82.4%	77.8%	73.7%	70.0%	66.7%	63.6%	60.9%	58.3%	56.0%	53.8%	51.9%	50.0%	48.3%	46.7%	45.2%	43.8%	42.4%	41.2%
	15	100.0%	93.8%	88.2%	83.3%	78.9%	75.0%	71.4%	68.2%	65.2%	62.5%	60.0%	57.7%	55.6%	53.6%	51.7%	50.0%	48.4%	46.9%	45.5%	44.1%	42.9%
	16	100.0%	94.1%	88.9%	84.2%	80.0%	76.2%	72.7%	69.6%	66.7%	64.0%	61.5%	59.3%	57.1%	55.2%	53.3%	51.6%	50.0%	48.5%	47.1%	45.7%	44.4%
	17	100.0%	94.4%	89.5%	85.0%	81.0%	77.3%	73.9%	70.8%	68.0%	65.4%	63.0%	60.7%	58.6%	56.7%	54.8%	53.1%	51.5%	50.0%	48.6%	47.2%	45.9%
	18	100.0%	94.7%	90.0%	85.7%	81.8%	78.3%	75.0%	72.0%	69.0%	66.7%	64.3%	62.1%	60.0%	58.1%	56.3%	54.5%	52.9%	51.4%	50.0%	48.6%	47.4%
	19	100.0%	95%	90.5%	86.4%	82.6%	79.2%	76.0%	73.1%	70.4%	67.9%	65.5%	63.3%	61.3%	59.4%	57.6%	55.9%	54.3%	52.8%	51.4%	50.0%	48.7%
	20	100.0%	95.2%	90.9%	87.0%	83.3%	80.0%	76.9%	74.1%	71.4%	69.0%	66.7%	64.5%	62.5%	60.6%	58.8%	57.1%	55.6%	54.1%	52.6%	51.3%	50.0%