Universität Konstanz

Four Essays on Cognitive Processes of Social Decision Making

Doctoral thesis for obtaining the

academic degree Doctor of economics

submitted by

Chen Fadong

at the University of Konstanz

Faculty of Politics, Law and Economics

Department of Economics

Date of the oral examination: 24.05.2017

1. Reviewer: Prof. Dr. Urs Fischbacher

2. Reviewer: Prof. Dr. Ian Krajbich

Acknowledgement

First and foremost, I am deeply grateful to my dedicated supervisor Prof. Dr. Urs Fischbacher for supervising this thesis, for being an inspiring co-author, for providing an environment where ideas can grow, for being an excellent guide to the world of academia, and for his continuous help and support throughout the past few years. Words cannot express the gratitude I feel for the time he spent discussing my research and providing detailed comments, supporting my exploration of different ideas, filling me with faith as well as suggestions when I ran into challenges and obstacles.

I am grateful to Prof. Dr. Ian Krajbich and Prof. Dr. Sebastian Fehrler for serving as committee members of this dissertation, for their insightful comments, challenging and inspiring questions, and reliable support. Furthermore, I am thankful to Prof. Dr. Ian Krajbich and his team for making my research visit at the Ohio State University possible and great.

I benefited from many discussions with and comments from the lovely TWI group. There was always something new that I could learn from them, including the importance of staying focused on the big picture research question while also keeping track of every piece of institutional detail, how to keep presentation concise and to the point, and much, much more.

Furthermore, I would like to thank my family for all their love and encouragement. I am thankful for my parents and my brother for raising me and caring for me to be who I am, for encouraging me to do what I like, and for feeling proud of what I have become but never giving me any stress about what to accomplish.

The final words belong to the persons to whom this dissertation is dedicated to, my wife Ling and my son Xiao-Ou. Successfully completing a Ph.D. program requires many sacrifices. No one sacrificed more for me to be able to achieve this goal than my wife, Ling. Although my name is on the title, the cost was shared by the both of us, and I will never forget that. For this, for your support, and for being partners in this great adventure, I am forever grateful. I would also like to thank my son, Xiao-Ou, who brought joy into my life at the end of the study.

Table of Contents

List of Tables

List of Figures

Summary

One of the basic assumptions of standard economic theory is that individuals are purely self-interested. However, overwhelming evidence from a variety of disciplines has indicated that a substantial number of people exhibit social preferences. That is, rather than caring exclusively about their own material outcomes, they also care positively or negatively for the material outcomes of others (see, e.g., Berg, Dickhaut, and McCabe 1995, Charness, Frechette, and Kagel 2004, Charness and Rabin 2002, Fehr, Kirchsteiger, and Riedl 1993, Güth, Schmittberger, and Schwarze 1982). Social preferences have a significant impact on economic consequences and are very important in understanding fundamental economic questions (Fehr and Fischbacher 2002). Thus, understanding the nature of social preferences and how they interact with material incentives is highly relevant to improving economists' ability to predict behavior, analyze welfare, and make policy recommendations.

There is a huge body of theoretical research that investigates the structure of social preferences on the behavioral level. These theoretical models assume that, additional to the selfish motive, people have other motives that capture their social preferences, for example, altruism, inequality aversion, envy, and reciprocity (Bolton and Ockenfels 2000, Dufwenberg and Kirchsteiger 2004, Falk, Fehr, and Fischbacher 2005, Falk and Fischbacher 2006, Fehr and Schmidt 1999, Levine 1998, Rabin 1993). These models are "black box" models and do not correspond to internal processes. To better understand social preferences, recently economists have become interested in the cognitive processes which govern these preferences. The analysis of response times (RTs) is a very simple method to acquire insights into these processes. However, results of studies using RT to investigate the cognitive process of social decision making are still disputed and inconclusive.

One debate is whether people's automatic response is to be prosocial or selfish. This line of studies investigates the cognitive process of social decision making using dual-process theories (Alós-Ferrer and Strack 2014, Brocas and Carrillo 2014, Evans 2008, Frederick 2005, Kahneman 2011). Some studies find that people have a spontaneous tendency to be prosocial (Cappelen et al. 2015, Cappelletti, Güth, and Ploner 2011, Lotito, Migheli, and Ortona 2013, Nielsen, Tyran, and Wengström 2014, Rand, Greene, and Nowak 2012, Schulz et al. 2014), while other studies find that people's intuition favors selfishness (Duffy and Smith 2014, Fiedler et al. 2013, Hauge et al. 2016, Lohse, Goeschl, and Diederich forthcoming, Piovesan and Wengström 2009, Tinghog et al. 2013). Another

line of research has argued that the cognitive process of social decision making is in line with sequential sampling models (SSMs), which assume that information is sampled continuously until sufficient evidence is accumulated for one of the available options (Ratcliff 1978, Ratcliff and McKoon 2008). In particular, these papers have argued that variations in RT are due to the utility differences between choice options, not to parallel competing processes (Hutcherson, Bushong, and Rangel 2015, Krajbich, Bartling, et al. 2015, Krajbich, Hare, et al. 2015, Krajbich, Oud, and Fehr 2014). This causes another debate that whether social decisions are the result of competing processes or the result of a single deliberative sequential sampling process.

This thesis conducts experimental, theoretical, and methodological studies to investigate the cognitive process of social decision making. Chapters 1 and 2 study the cognitive process of social preferences, and chapter 3 studies the cognitive process of social norm. These studies seek to give clear answers about the cognitive process of social decision making, and propose a process-based model to describe and explain social behavior. Chapter 4 shows that, additional to RT, click position can also be used as an indicator of people's preferences. Chapters 1 and 4 are joint work with Urs Fischbacher, and chapter 2 is a joint work with Ian Krajbich.

Specifically, the first chapter of this thesis (*Cognitive Processes of Distributional Preferences: A Response Time Study*) investigates the cognitive process of distributional preferences. A huge body of evidence indicates that people are willing to sacrifice their own material resources to benefit or hurt others, and it is a well-established fact that people are heterogeneous in the relevant motives and in the strength of preferences (Andreoni and Miller 2002, Engelmann and Strobel 2004, Erlei 2008, Fisman, Kariv, and Markovits 2007, Kerschbamer 2015). For instance, some people care more about efficiency, while others care more about fairness (Murphy, Ackermann, and Handgraaf 2011). Since studies concerning the cognitive process of social decision making are still disputed and inconclusive, we conjecture that individual heterogeneity plays an important role in explaining the cognitive process of social decision making.

In this study, we explicitly take individual heterogeneity into account, not only with respect to what kind of social motives people care about, but also with respect to the strength of people's selfish motives. Our experiment includes two types of binary three-person dictator games: the third-party (TP) dictator game and the second-party (SP) dictator game. In TP games, the dictator's payoffs are the same in the two allocations, while in SP games, the dictator's payoffs differ between the two allocations. Decisions in TP games allow us to identify subjects' social motives. And then we can study how people

with different social motives react to various decision situations in SP games using RT analysis. In addition, we can study which kind of motive, the selfish or the social motives, can be considered as more intuitive. Using finite mixture model, we classify subjects into three norm types which differ with respect to the relevant social motives based on the decisions for TP games. Based on this identification, a within-subject analysis on the SP games shows that the RT increases with the number of conflicts between individually relevant motives and decreases with the utility difference between choice options. These results are consistent with the predictions of sequential sampling models. A between-subject analysis reveals the heterogeneity with respect to whether the selfish or the social motive is more intuitive. It turns out that the selfish motive is more intuitive for subjects who are more selfish, which can also be corroborated by the eye-tracking data. We also show that the heterogeneity in social motives and selfishness is reflected in process differences. Our findings demonstrate that the cognitive processes of social decision making comply with both, sequential sampling models and dual-process theories. And our findings also show that it is important to take the heterogeneity of preferences into account when investigating the cognitive processes of social decision making.

The second chapter (*Sequential Sampling with Intuition*) studies the cognitive process of social decision making using a series of mini-dictator games and time manipulation. In this chapter, we seek to understand why time pressure favors prosocial behavior in some cases and selfishness in others, and propose a model to describe the cognitive process underlying social decision making by integrating the notion of intuition into a sequential sampling framework. To do so, we use a series of mini-dictator games in which there is always a conflict between selfishness and prosociality. Subjects make binary decisions for these games under time-free, time-pressure and time-delay conditions in the experiment. The experiment allows us to identify whether subjects have an early bias to be prosocial or selfish by comparing how their preferences change across different time conditions. The experiment also allows us to test whether the cognitive process of social decision making complies with SSMs, even under time pressure and time delay.

We find that subjects are heterogeneous in whether their intuition favors prosociality or selfishness. In particular, prosocial subjects become more prosocial under time pressure and less prosocial under time delay, while selfish subjects do the opposite. This indicates that the intuition favors the selfish option for selfish subjects, and the intuition favors the prosocial option for prosocial subjects. The experiment also provides strong evidence that social decision making is in line with SSMs. That is, RTs decrease

with the strength of preference in all three conditions. These results help to resolve the current debate and reconcile the conflicting results concerning the cognitive process of social decision making. Going a step further, we integrate the intuition of selfishness or prosociality into the starting point of a SSM process as a bias and propose a SSM with intuition (SSMi) to model the cognitive process. We validate the model with out-of-sample predictions. The out-of-sample prediction results show that the SSMi has higher predictive power than the standard unbiased SSM and SSM with biased drift rates or both biased starting points and drift rates. Moreover, the SSMi outperforms standard logistic choice models, highlighting the importance of modeling the choice process.

The third chapter of this thesis (*Cognitive Processes of Prosocial and Antisocial Punishment*) investigates the cognitive process of prosocial and antisocial punishment. To better understand punishment behavior and how it affects cooperation, research across a wide variety of fields has studied the motivations underlying prosocial and anti-social punishment (Bone and Raihani 2015, Fehr and Gächter 2002, Fowler, Johnson, and Smirnov 2005, Pfattheicher and Schindler 2015, Sylwester, Herrmann, and Bryson 2013). However, the underlying process which governs punishment behavior remains largely uninvestigated. This chapter is to fill this gap, and examine whether the cognitive process is in line with sequential sampling models and whether people have a spontaneous tendency to punish others or not. More specifically, I study this topic using behavioral and RT data from a repeated public goods game. In the contribution stage of the public goods game, subjects decide how many points they would like to contribute to the public goods. In the punishment stage, subjects decide whether to punish each of their group members or not based on the information of the own contribution, the other's contribution, the cost and the impact of punishment. To create different situations for subjects, the cost and the impact of punishment are randomly selected from four preset levels and provided to subjects together with other information in each decision situation.

The results show that, similar to perceptual or value-based decision making, RT decreases with the strength of preference on punishment or non-punishment, which indicates that both the cognitive process of prosocial and antisocial punishment complies with sequential sampling models. This result can be corroborated by the RT difference between punishment and non-punishment. Subjects who are more likely to punish others are quicker in making punishment decisions, and subjects who are less likely to punish others are slower in making punishment decisions. In addition, the sequential sampling process starts near the threshold of non-punishment when cooperators punishing defectors, and the process starts near the threshold of punishment when defectors

punishing cooperators. That is, prosocial punishment is more deliberative and antisocial punishment is more intuitive than non-punishment. Furthermore, the sequential sampling process of punishment can be corroborated by the RT difference between punishment and non-punishment. And prosocial (antisocial) punishment is more deliberative (intuitive) can be corroborated by out-of-sample predictions. The drift diffusion model which takes the intuition into account improves the ability to make out-of-sample predictions compared to the standard drift diffusion model.

The fourth chapter of this thesis (*Response Time and Click Position: Cheap Indicators of Preferences, Journal of Economic Science Association, 2016*) is a methodological study. There is an increasing interest in using process data to get a better understanding of economic decision making. Different methods such as RTs (Rubinstein 2007), pupil dilations (Wang, Spezio, and Camerer 2010), eye movements (Reutskaja et al. 2011) and even neural activity (Smith et al. 2014) have been introduced and have shown significant correspondences between this kind of data and people's decisions. In this chapter, we study how RT and click position relate to economic decisions in a social value orientation task (Murphy, Ackermann, and Handgraaf 2011). The SVO task consists of several decision situations, each consisting of a menu of allocations between the deciding subject and an anonymous partner. Thus, the SVO measures how much people care about their own as compared to the other player's welfare. The SVO task represents a prototypical situation in which people have to assess and trade off different attributes of an option—in the case of the SVO their own vs the other's payoff.

We find that more individualistic subjects have shorter RTs than prosocial subjects in the SVO task. Individualistic subjects click more often on their own payoffs than on the others' payoffs, and they click more often on their own payoffs than prosocial subjects. Moreover, the RT information and the click position information are complementary in explaining subjects' preferences. These results show that RTs and click positions can be used as indicators of people's preferences. In addition, the general idea can be applied to any type of choice between multi-attribute options such as consumer decisions, where options differ in price and brand attributes. Also in this kinds of decisions, RT data and click positions provide additional information about people's preferences and can be used to improve the predictive power.

Zusammenfassung

Eine der wesentlichen Annahmen ökonomischer Standardtheorie besteht darin, dass Menschen egoistisch sind. Eine große Fülle wissenschaftlicher Evidenz zeigt allerdings, dass viele Menschen sozial handeln und sogenannte soziale Präferenzen aufweisen. Dies heißt, dass viele Menschen nicht allein egoistischen Motiven folgen, sondern auch die negativen oder positiven Konsequenzen für andere in ihre Entscheidungen mit einbeziehen (see, e.g., Berg, Dickhaut, and McCabe 1995, Charness, Frechette, and Kagel 2004, Charness and Rabin 2002, Fehr, Kirchsteiger, and Riedl 1993, Güth, Schmittberger, and Schwarze 1982). Soziale Präferenzen haben einen grundlegenden Einfluss auf ökonomische Ergebnisse und sind damit wichtig für das Verständnis fundamentaler ökonomischer Fragestellungen (Fehr and Fischbacher 2002). Somit ist es von großer Bedeutung zu verstehen, wie genau soziale Präferenzen und materielle Anreize interagieren. Dies ermöglicht es, ökonomisch relevantes Verhalten sowie Auswirkungen auf das Gemeinwohl besser vorherzusagen und daraus politische Empfehlungen abzuleiten.

Es existiert eine umfangreiche theoretische Forschungsliteratur, die die Struktur sozialer Präferenzen auf der Verhaltensebene untersucht. Diesen theoretischen Modellen liegt die Annahme zugrunde, dass Individuen, zusätzlich zu ihren egoistischen Motiven, auch andere, soziale Motive verfolgen. Hierzu gehören beispielsweise Altruismus, Ungleichheitsaversion, Neid oder Reziprozität (Bolton and Ockenfels 2000, Dufwenberg and Kirchsteiger 2004, Falk, Fehr, and Fischbacher 2005, Falk and Fischbacher 2006, Fehr and Schmidt 1999, Levine 1998, Rabin 1993). Diese Modelle sind jedoch *Black-Box*-Modelle und haben keine Entsprechung in den kognitiven Prozessen, die tatsächlich zu den beobachteten Entscheidungen führen. Um soziale Präferenzen besser zu verstehen, haben nun aber auch Ökonomen damit begonnen, die kognitiven Prozesse, die den weithin beobachteten Präferenzen zugrunde liegen, näher zu ergründen. Dabei ist die Analyse von Reaktionszeiten eine einfache Methode, erste Erkenntnisse über diese kognitiven Prozesse zu erhalten. Dennoch sind viele Resultate, die auf der Analyse von Reaktionszeiten beruhen, umstritten und als vorläufig zu betrachten.

So besteht beispielsweise Uneinigkeit darüber, ob Menschen intuitiv eher pro-sozial oder orientiert am materiellen Eigennutzen handeln. In dieser Hinsicht untersuchen Studien die kognitiven Prozesse in sozialen Entscheidungssituationen auf Basis sogenannter *Dual-Process*-Theorien (Alós-Ferrer and Strack 2014, Brocas and Carrillo 2014, Evans 2008, Frederick 2005, Kahneman 2011). Einige Studien haben dabei

gezeigt, dass intuitive Handlungen eher pro-sozialer Natur sind (Cappelen et al. 2015, Cappelletti, Güth, and Ploner 2011, Lotito, Migheli, and Ortona 2013, Nielsen, Tyran, and Wengström 2014, Rand, Greene, and Nowak 2012, Schulz et al. 2014), während andere postulieren, dass selbige eher egoistisch sind (Duffy and Smith 2014, Fiedler et al. 2013, Hauge et al. 2016, Lohse, Goeschl, and Diederich forthcoming, Piovesan and Wengström 2009, Tinghog et al. 2013). Wiederum andere haben argumentiert, dass kognitive Prozesse, die das Verhalten in Entscheidungssituationen bestimmen, in Übereinstimmung mit *Sequential-Sampling*-Modellen (SSM) stehen. Der SSM-Ansatz nimmt an, dass Informationen kontinuierlich gesammelt werden bis genügend Evidenz für eine der verfügbaren Handlungsalternativen vorhanden ist (Ratcliff 1978, Ratcliff and McKoon 2008). Auf Basis des SSM-Ansatzes haben verschiedene Autoren argumentiert, dass Unterschiede in den Reaktionszeiten von der Größe der Nutzenunterschiede zwischen den zur Wahl stehenden Alternativen abhängt, und nicht von parallelen, in Wettbewerb stehenden Prozessen (Hutcherson, Bushong, and Rangel 2015, Krajbich, Bartling, et al. 2015, Krajbich, Hare, et al. 2015, Krajbich, Oud, and Fehr 2014). Dies hat eine Debatte darüber initiiert, ob soziale Entscheidungen das Resultat eines Wettbewerbs zwischen konkurrierenden Prozessen sind oder durch einen einzigen sequentiellen *Sampling*-Prozess beschrieben werden können.

Diese Dissertation untersucht die kognitiven Prozesse in sozialen Entscheidungssituationen auf experimenteller, theoretischer und methodischer Ebene. Kapitel 1 und 2 befassen sich mit den kognitiven Prozessen hinter sozialen Entscheidungen, Kapitel 3 beleuchtet die kognitiven Prozesse hinter sozialen Normen. Alle drei Kapitel verfolgen das Ziel, die kognitiven Prozesse hinter sozialen Entscheidungen genauer herauszuarbeiten und prozessbasierte Modelle zur Beschreibung sowie Erklärung von sozialem Verhalten zu finden. Kapitel 4 zeigt auf, dass die Analyse der Klickposition eine weitere Möglichkeit neben der Betrachtung der Reaktionszeiten darstellt, um individuelle Präferenzen zu untersuchen. Kapitel 1 und 4 sind in Zusammenarbeit mit Urs Fischbacher entstanden, Kapitel 2 in Zusammenarbeit mit Ian Krajbich.

Das erste Kapitel dieser Arbeit *Cognitive Processes of Distributional Preferences: A Response Time Study* untersucht die kognitiven Prozesse hinter Verteilungspräferenzen. Es gibt starke Evidenz dafür, dass Menschen bereit sind, eigene materielle Ressourcen zu opfern, um andere besser zu stellen oder ihnen zu schaden. Außerdem ist es eine bereits gut untersuchte Tatsache, dass Menschen unterschiedlichen Motiven folgen und diese für sie in unterschiedlicher Stärke relevant sind (Andreoni and Miller 2002, Engelmann and Strobel 2004, Erlei 2008, Fisman, Kariv, and Markovits 2007, Kerschbamer 2015).

Beispielsweise gibt es Personen, denen Effizienz wichtiger ist, wohingegen andere Fairness als wichtiger erachten (Murphy, Ackermann, and Handgraaf 2011). Da Studien, die sich mit den kognitiven Prozessen hinter sozialen Entscheidungen befassen noch immer umstritten und widersprüchlich sind, vermuten wir, dass inter-individuelle Heterogenität eine wichtige Rolle in der Erklärung von kognitiven Prozessen in sozialen Entscheidungssituationen spielt.

In unserer Studie haben wir Heterogenität daher explizit berücksichtigt - nicht nur in Hinblick darauf, welche sozialen Motive Menschen wichtig sind, sondern auch in Bezug auf die Stärke des egoistischen Motivs. Unser Experiment besteht aus zwei Varianten eines binären Drei-Personen-Diktatorspiels: dem Drittpartei-Diktatorspiel (TP-Diktatorspiel) und dem Zweitpartei-Diktatorspiel (SP-Diktatorspiel). In TP-Spielen ist die Auszahlung des Diktators in beiden verfügbaren Allokationen gleich, wohingegen sich in SP Spielen die Auszahlung des Diktators zwischen den beiden Allokationen unterscheidet. Entscheidungen in TP-Spielen ermöglichen es uns, die sozialen Motive der Experimentteilnehmer zu identifizieren. Wir verwenden die TP-Spiele, um die Relevanz der unterschiedlichen sozialen Motive einzuschätzen. Darauf basierend untersuchen wir dann, auch mit Reaktionszeit-analysen, wie sich die Teilnehmer in den SP-Spielen verhalten. Dies erlaubt es festzustellen, ob soziale Motive oder die Eigennutzorientierung intuitiver sind. Mit Hilfe eines *Finite-Mixture*-Modells können wir die Teilnehmer in drei verschiedene Normkategorien klassifizieren, welche sich in Bezug auf die relevanten sozialen Motive in den TP-Spielen unterscheiden. Basierend auf dieser Klassifikation zeigt sich, dass die Reaktionszeiten in den SP-Spielen mit der Anzahl der Konflikte zwischen den verschiedenen Motiven zunehmen und mit dem Nutzenunterschied zwischen den Alternativen abnehmen. Die Ergebnisse sind konsistent mit den Vorhersagen von *Sequential-Sampling*-Modellen. Ferner zeigt sich Heterogenität darin, ob Eigennutzorientierung oder Prosozialität intuitiver ist. Die Eigennutzorientierung zeigt sich intuitiver für egoistischere Experimentteilnehmer, was auch durch Eye-tracking-Daten bestätigt wird. Es stellt sich auch heraus, dass die Heterogenität in den sozialen und egoistischen Motiven mit Unterschieden in den zugrunde liegenden kognitiven Prozessen korrespondiert. Unsere Ergebnisse zeigen, dass die kognitiven Prozesse in sozialen Entscheidungssituationen sowohl mit *Sequential-Sampling*-Modellen wie auch mit *Dual-Process*-Theorien in Einklang zu bringen sind. Zusätzlich weisen unsere Ergebnisse darauf hin, dass es für die Untersuchung kognitiver Prozesse von entscheidender Bedeutung ist, die Heterogenität von Präferenzen zu berücksichtigen.

Das zweite Kapitel *Sequential Sampling with Intuition* untersucht die kognitiven Prozesse hinter sozialen Entscheidungen anhand einer Variation der Entscheidungszeit in einer Reihe von Mini-Diktatorspielen. Das Ziel dieses Kapitel ist es, zu verstehen, warum Zeitdruck in manchen Fällen zu pro-sozialerem Verhalten und in anderen Fällen zu egoistischerem Verhalten führt. Zudem schlagen wir ein Modell vor, das die kognitiven Prozesse in dieser Art von Entscheidungssituation genauer beschreibt. Dabei wird der Begriff der intuitiven Entscheidung in einen *Sequential-Sampling*-Rahmen eingefügt. Das zugrundeliegende Experiment beruht auf einer Serie von Mini-Diktatorspielen, die so konstruiert sind, dass jeweils ein Konflikt zwischen materiellem Eigennutzen und prosozialen Motiven entsteht. Die Teilnehmer trafen in diesen Spielen binäre Entscheidungen in drei Experimentalbedingungen: (1) ohne Vorgabe einer Entscheidungszeit, (2) unter Zeitdruck und (3) mit zeitlicher Verzögerung. Durch den Vergleich des Verhaltens in den drei Bedingungen erlaubt uns das Experiment zu identifizieren, ob die Teilnehmer intuitiv pro-sozial oder egoistisch entscheiden. Das Experiment ermöglicht es uns ferner zu testen, ob die kognitiven Prozesse in diesen Entscheidungssituationen auch unter Zeitdruck und Zeitverzögerung kompatibel mit dem SSM-Ansatz sind.

In den Ergebnissen stellt sich heraus, dass die Teilnehmer heterogen in Bezug darauf sind, ob sie intuitiv pro-sozial oder egoistisch handeln. Insbesondere werden pro-soziale Personen unter Zeitdruck pro-sozialer und bei zeitlicher Verzögerung weniger pro-sozial. Für egoistische Teilnehmer beobachten wir den gegenteiligen Effekt. Dies deutet darauf hin, dass Menschen mit egoistischeren Präferenzen intuitiv egoistisch handeln und dass pro-sozialer eingestellte Personen intuitiv pro-sozial entscheiden. Das Experiment liefert zudem starke Hinweise dafür, dass Entscheidungen in sozialen Situationen ein kognitiver Prozess zugrunde liegt, der in Übereinstimmung mit dem SSM-Ansatz steht. Dies zeigt sich beispielsweise darin, dass die Reaktionszeiten in allen drei Experimentalbedingungen bei stärkeren Präferenzunterschieden kürzer sind. So liefern unsere Ergebnisse einen Beitrag, die Diskussion über die widersprüchlichen Ergebnisse bezüglich kognitiver Prozesse in sozialen Entscheidungssituationen aufzulösen. In einem weiteren Schritt haben wir ein spezielles SSM entwickelt, das die jeweilige Intuition direkt berücksichtigt (SSMi). In diesem Modell haben wir die individuelle Intuition in Richtung materiellem Eigennutzen oder Pro-Sozialität über unterschiedliche Startpunkte in ein SSM integriert, um den kognitiven Prozess genauer zu modellieren. Das resultierende Modell haben wir dann mithilfe von Kreuzvalidierungsverfahren getestet. Diese Tests zeigen, dass das SSMi prädiktiv besser abschneidet als ein Standard-SSM, als ein SSM mit unterschiedlichen *Drift*-Raten und ein SSM, das Startpunkt und *Drift*-Raten variieren lässt.

Zudem übertrifft das SSMi-Modell in der prädiktiven Güte auch ein Standard-Logit-Modell. Dies unterstreicht die Bedeutung einer genauen Modellierung der zugrunde liegenden Entscheidungsprozesse unter Berücksichtigung intuitiv egoistischer oder pro-sozialer Tendenzen.

Das dritte Kapitel der Dissertation *Cognitive Processes of Prosocial and Antisocial Punishment* untersucht die kognitiven Prozesse hinter pro-sozialer und anti-sozialer Bestrafung. Eine Vielzahl von Untersuchungen unterschiedlicher Disziplinen hat sich bereits mit den grundlegenden Motiven hinter pro-sozialer und anti-sozialer Bestrafung befasst, um Bestrafungsverhalten und dessen Wirkung auf Kooperation besser zu verstehen (Bone und Raihani 2015, Fehr und Gächter 2002, Fowler, Johnson, und Smirnov 2005, Pfattheicher und Schindler 2015, Sylwester, Herrmann, und Bryson 2013). Allerdings blieben die kognitiven Prozesse, welche dem jeweiligen Bestrafungsverhalten zugrunde liegen, bislang größtenteils unerforscht. Die Untersuchung, die ich in diesem Kapitel genauer ausführe, beabsichtigt, diese Lücke zu schließen und näher zu beleuchten, ob die zugrunde liegenden kognitiven Prozesse im Einklang mit der SSM-Idee stehen und ob Menschen eine intuitive Tendenz haben, andere zu bestrafen oder nicht. Konkret untersuche ich diese Fragestellung durch Analyse von Verhaltensdaten und Reaktionszeiten in einem wiederholten Öffentlichen-Güter-Spiel. In der Beitragsphase des Öffentlichen-Güter-Spiels entschieden die Teilnehmer, wie viele Punkte sie zu einem öffentlichen Gut beitragen möchten. In der Bestrafungsphase besteht für diese dann die Möglichkeit - basierend auf der Information über den eigenen Beitrag, den Beitrag der anderen und die Kosten sowie Auswirkung der Bestrafung für jedes der Gruppenmitglieder - andere zu bestrafen oder nicht. In den verschiedenen Entscheidungssituationen wurden die Kosten und die Auswirkung der Bestrafung zufällig aus vier vordefinierten Niveaus ausgewählt und den Teilnehmern in jeder Situation zusammen mit den anderen Informationen dargestellt.

Die Resultate zeigen, dass die Reaktionszeit ähnlich wie in Wahrnehmungs-experimenten mit der Stärke der Präferenz für oder gegen Bestrafung abnimmt. Dies deutet darauf hin, dass ein SSM-Ansatz die kognitiven Prozesse hinter pro-sozialer wie auch hinter anti-sozialer Bestrafung erklären kann. Die Unterschiede in den Reaktionszeiten zwischen dem tatsächlichen Bestrafen und dem Absehen von Bestrafung stützen diese Schlussfolgerung. Teilnehmer, die wahrscheinlicher andere bestrafen, treffen Bestrafungs-entscheidungen schneller und Teilnehmer, die weniger wahrscheinlich andere bestrafen, treffen Bestrafungsentscheidungen langsamer. Außerdem startet der *Sequential-Sampling*-Prozess näher an der Schwelle zur Nicht-

Bestrafung, wenn kooperierende Teilnehmer Trittbrettfahrer bestrafen. Wenn Trittbrettfahrer hingegen kooperierende Teilnehmer bestrafen, startet der Prozess nahe der Schwelle zur Bestrafung. Dies bedeutet, dass pro-soziale Bestrafung überlegter und anti-soziale Bestrafung intuitiver ist. Ferner wird der SSM-Ansatz durch die Reaktionszeitunterschiede zwischen Bestrafung und Nicht-Bestrafung bestätigt. Die prädiktive Güte der *out-of-sample* Vorhersagen stützt zudem die Schlussfolgerung, dass pro-soziale Bestrafung überlegterer und anti-soziale Bestrafung intuitiverer Natur ist. Im Vergleich zu einem Standard-Diffusionsansatz verbessert das Berücksichtigen der intuitiven Orientierung *out-of-sample* Vorhersagen.

Das vierte Kapitel *Response Time and Click Position: Cheap Indicators of Preferences*, (Journal of Economic Science Association, 2016) stellt eine methodisch-orientierte Untersuchung dar. Um ein besseres Verständnis ökonomischer Entscheidungsfindung zu erhalten, ist das Interesse an der Nutzung Daten, die auf kognitive Prozesse schließen lassen, stark gestiegen. Hierzu werden verschiedene Methoden wie die Analyse von Reaktionszeiten (Rubinstein 2007), Pupillenreaktionen (Wang, Spezio, und Camerer 2010), *Eye-tracking* (Reutskaja et al. 2011) oder die direkte Erfassung neuronaler Aktivität (Smith et al. 2014) herangezogen und zeigen signifikante Zusammenhänge mit menschlichem Entscheidungsverhalten auf. In diesem Kapitel untersuchen wir, wie Reaktionszeiten und die Klickposition auf einem Computer-Interface mit ökonomischen Entscheidungen in einer Aufgabe zur Erfassung der sogenannten *Social Value Orientation* (SVO) zusammenhängen (Murphy, Ackermann, and Handgraaf 2011). Die SVO-Aufgabe besteht aus mehreren Entscheidungssituationen. In jeder Situation wählen die Teilnehmer eine von mehreren verfügbaren Aufteilungen von Punkten zwischen sich selbst und einem anonymen Partner. Die SVO misst, wie stark Menschen die eigene Auszahlung im Vergleich zur Auszahlung des anderen Spielers gewichten. Dabei stellt die SVO-Aufgabe eine prototypische Situation dar, in der Menschen verschiedene Attribute einer Option bewerten und gegeneinander abwägen müssen – im Fall der SVO betrifft dies die eigene Auszahlung und die Auszahlung eines anderen Teilnehmers.

In der Analyse der SVO-Entscheidungen finden wir, dass individualistische Teilnehmer kürzere Reaktionszeiten als pro-soziale Teilnehmer aufweisen. Individualistische Teilnehmer klicken bei ihrer Entscheidung zudem öfter auf die Darstellung ihrer eigenen Auszahlung als auf die Auszahlungen anderer und sie klicken öfter auf ihre eigenen Auszahlungen als prosoziale Teilnehmer. Darüber hinaus ist die Information über die Reaktionszeit und die Information über die Klickposition komplementär in der Prognose der individuellen Präferenzen. Diese Ergebnisse zeigen,

dass nicht nur Reaktionszeiten sondern auch die Klickposition als Indikator für Entscheidungspräferenzen genutzt werden kann. Zudem lässt sich die Idee der Analyse von Klickpositionen prinzipiell auf jede Art von Wahlentscheidung zwischen Alternativen mit verschiedenen Attributen anwenden - beispielsweise auf Konsumentenentscheidungen,

bei denen sich die verschiedenen Alternativen in Preis und weiteren Markenattributen unterscheiden. Auch in diesen Arten von Entscheidungen können Daten über Reaktionszeiten und Klickpositionen zusätzliche Informationen über Präferenzen liefern und so die Prognosegüte vorhersagender Modelle und Schätzungen verbessern.

References of the summary/Literatur der Zusammenfassung

Alós-Ferrer, Carlos, and Fritz Strack. 2014. "From Dual Processes to Multiple Selves: Implications for Economic Behavior." *Journal of Economic Psychology* 41:1-11.

Andreoni, James, and John Miller. 2002. "Giving According to Garp: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica* 70 (2):737-53.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10 (1):122-42.

Bolton, Gary E., and Axel Ockenfels. 2000. "Erc: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90 (1):166-93.

Bone, Jonathan E, and Nichola J Raihani. 2015. "Human Punishment Is Motivated by Both a Desire for Revenge and a Desire for Equality." *Evolution and Human Behavior* 36 (4):323-30.

Brocas, Isabelle, and Juan D. Carrillo. 2014. "Dual-Process Theories of Decision-Making: A Selective Survey." *Journal of Economic Psychology* 41:45-54.

Cappelen, Alexander W., Ulrik H. Nielsen, Bertil Tungodden, Jean-Robert Tyran, and Erik Wengström. 2015. "Fairness Is Intuitive." *Experimental Economics*:1-14.

Cappelletti, Dominique, Werner Güth, and Matteo Ploner. 2011. "Being of Two Minds: Ultimatum Offers under Cognitive Constraints." *Journal of Economic Psychology* 32 (6):940-50.

Charness, Gary, Guillaume R Frechette, and John H Kagel. 2004. "How Robust Is Laboratory Gift Exchange?" *Experimental Economics* 7 (2):189-205.

Charness, Gary, and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117 (3):817-69.

Duffy, Sean, and John Smith. 2014. "Cognitive Load in the Multi-Player Prisoner's Dilemma Game: Are There Brains in Games." *Journal of Behavioral and Experimental Economics* 51:47-56.

Dufwenberg, Martin, and Georg Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47 (2):268-98.

Engelmann, Dirk, and Martin Strobel. 2004. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments." *American Economic Review* 94 (4):857-69.

Erlei, Mathias. 2008. "Heterogeneous Social Preferences." *Journal of Economic Behavior & Organization* 65 (3–4):436-57.

Evans, Jonathan St BT. 2008. "Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition." *Annual Review of Psychology* 59:255-78.

Falk, Armin, Ernst Fehr, and Urs Fischbacher. 2005. "Driving Forces Behind Informal Sanctions." *Econometrica* 73 (6):2017-30.

Falk, Armin, and Urs Fischbacher. 2006. "A Theory of Reciprocity." *Games and Economic Behavior* 54 (2):293-315.

Fehr, Ernst, and Urs Fischbacher. 2002. "Why Social Preferences Matter – the Impact of Non-Selfish Motives on Competition, Cooperation and Incentives." *The Economic Journal* 112 (478):C1-C33.

Fehr, Ernst, and Simon Gächter. 2002. "Altruistic Punishment in Humans." *Nature* 415 (6868):137-40.

Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl. 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation." *Quarterly Journal of Economics* 108 (2):437-59.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3):817-68.

Fiedler, Susann, Andreas Glöckner, Andreas Nicklisch, and Stephan Dickert. 2013. "Social Value Orientation and Information Search in Social Dilemmas: An Eye-Tracking Analysis." *Organizational Behavior and Human Decision Processes* 120 (2):272-84.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. "Individual Preferences for Giving." *American Economic Review* 97 (5):1858-76.

Fowler, James H, Tim Johnson, and Oleg Smirnov. 2005. "Egalitarian Motive and Altruistic Punishment." *Nature* 433 (10.1038).

Frederick, Shane. 2005. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives* 19 (4):25-42.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze. 1982. "An Experimental Analysis of Ultimatum Bargaining." *Journal of Economic Behavior & Organization* 3 (4):367-88.

Hauge, Karen Evelyn, Kjell Arne Brekke, Lars-Olof Johansson, Olof Johansson-Stenman, and Henrik Svedsäter. 2016. "Keeping Others in Our Mind or in Our Heart? Distribution Games under Cognitive Load." *Experimental Economics* 19 (3):562-76.

Hutcherson, Cendri A., Benjamin Bushong, and Antonio Rangel. 2015. "A Neurocomputational Model of Altruistic Choice and Its Implications." *Neuron* 87 (2):451-62.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*: Macmillan.

Kerschbamer, Rudolf. 2015. "The Geometry of Distributional Preferences and a Non-Parametric Identification Approach: The Equality Equivalence Test." *European Economic Review* 76:85-103.

Krajbich, Ian, Björn Bartling, Todd Hare, and Ernst Fehr. 2015. "Rethinking Fast and Slow Based on a Critique of Reaction-Time Reverse Inference." *Nature Communications* 6:7455.

Krajbich, Ian, Todd Hare, Bjoern Bartling, Yosuke Morishima, and Ernst Fehr. 2015. "A Common Mechanism Underlying Food Choice and Social Decisions." *PLoS Comput Biol* 11 (10):e1004371.

Krajbich, Ian, Bastiaan Oud, and Ernst Fehr. 2014. "Benefits of Neuroeconomic Modeling: New Policy Interventions and Predictors of Preference." *American Economic Review* 104 (5):501-6.

Levine, David K. 1998. "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics* 1 (3):593-622.

Lohse, Johannes, Timo Goeschl, and Johannes H. Diederich. forthcoming. "Giving Is a Question of Time: Response Times and Contributions to an Environmental Public Good." *Environmental and Resource Economics*.

Lotito, Gianna, Matteo Migheli, and Guido Ortona. 2013. "Is Cooperation Instinctive? Evidence from the Response Times in a Public Goods Game." *Journal of Bioeconomics* 15 (2):123-33.

Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf. 2011. "Measuring Social Value Orientation." *Judgment & Decision Making* 6 (8):771-81.

Nielsen, Ulrik H., Jean-Robert Tyran, and Erik Wengström. 2014. "Second Thoughts on Free Riding." *Economics Letters* 122 (2):136-9.

Pfattheicher, Stefan, and Simon Schindler. 2015. "Understanding the Dark Side of Costly Punishment: The Impact of Individual Differences in Everyday Sadism and Existential Threat." *European Journal of Personality* 29 (4):498-505.

Piovesan, Marco, and Erik Wengström. 2009. "Fast or Fair? A Study of Response Times." *Economics Letters* 105 (2):193-6.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (5):1281-302.

Rand, David G., Joshua D. Greene, and Martin A. Nowak. 2012. "Spontaneous Giving and Calculated Greed." *Nature* 489 (7416):427-30.

Ratcliff, Roger. 1978. "A Theory of Memory Retrieval." *Psychological Review* 85 (2):59-108.

Ratcliff, Roger, and Gail McKoon. 2008. "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks." *Neural Computation* 20 (4):873-922.

Reutskaja, Elena, Rosemarie Nagel, Colin F. Camerer, and Antonio Rangel. 2011. "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study." *American Economic Review* 101 (2):900-26.

Rubinstein, Ariel. 2007. "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal* 117 (523):1243-59.

Schulz, Jonathan F., Urs Fischbacher, Christian Thöni, and Verena Utikal. 2014. "Affect and Fairness: Dictator Games under Cognitive Load." *Journal of Economic Psychology* 41:77-87.

Smith, Alec, B. Douglas Bernheim, Colin F. Camerer, and Antonio Rangel. 2014. "Neural Activity Reveals Preferences without Choices." *American Economic Journal: Microeconomics* 6 (2):1-36.

Sylwester, Karolina, Benedikt Herrmann, and Joanna J Bryson. 2013. "Homo Homini Lupus? Explaining Antisocial Punishment." *Journal of Neuroscience, Psychology, and Economics* 6 (3):167.

Tinghog, Gustav, David Andersson, Caroline Bonn, Harald Bottiger, Camilla Josephson, Gustaf Lundgren, Daniel Vastfjall, Michael Kirchler, and Magnus Johannesson. 2013. "Intuition and Cooperation Reconsidered." *Nature* 498 (7452):E1-E2.

Wang, Joseph Tao-Yi, Michael Spezio, and Colin F. Camerer. 2010. "Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games." *American Economic Review* 100 (3):984-1007.

1 Cognitive Processes of Distributional Preferences: A Response Time Study

1.1 Introduction

A huge body of evidence indicates that people are willing to sacrifice their own material resources to benefit or hurt others. Empirical research has investigated the motives underlying this behavior and theoretical models have been developed to formalize these motives (Bolton and Ockenfels 2000, Charness and Rabin 2002, Dufwenberg and Kirchsteiger 2004, Falk and Fischbacher 2006, Fehr and Schmidt 1999, Rabin 1993b). More recently, the cognitive processes which govern people's social behavior have come into focus: How are social decisions actually made? Social decisions are particularly interesting because they can be considered as compound goods satisfying different motives. Cognitive processes are about how we deal with these conflicting motives.

There has been considerable interest in exploring the cognitive processes of social decision making using dual-process approaches which assume the existence of two qualitatively distinct processes: One is relatively automatic and intuitive, and the other is relatively controlled and deliberative (Achtziger and Alós-Ferrer 2014, Alós-Ferrer and Strack 2014, Brocas and Carrillo 2014, Chaiken and Trope 1999, Frederick 2005, Fudenberg and Levine 2006, Hauge et al. 2016, Kahneman 2003, 2011, Sloman 1996, Strack and Deutsch 2004). Relative response times (RTs) are widely used to distinguish between the intuitive and the deliberative processes since the intuitive process are executed more quickly than the deliberative process (Krajbich, Bartling, et al. 2015a). In the domain of social preferences, this raises the question whether specific motives are processed more automatically than others, in particular whether the selfish or the social motive is more intuitive. The evidence based on studies using RTs and the manipulation of cognitive process is mixed so far. Some studies find that prosocial decisions are quicker than selfish decisions and people tend to be more prosocial under time pressure or cognitive load, which indicates that the social motives are more intuitive (Cappelen et al. 2015, Cappelletti, Güth, and Ploner 2011, Cornelissen, Dewitte, and Warlop 2011, Lotito, Migheli, and Ortona 2013, Nielsen, Tyran, and Wengström 2014, Peysakhovich and Rand forthcoming, Rand, Greene, and Nowak 2012, Rubinstein 2007, Schulz et al. 2014), while other studies find that selfish motives are more intuitive (Duffy and Smith 2014, Lohse, Goeschl, and Diederich 2014, Piovesan and Wengström 2009, Tinghog et al. 2013, Verkoeijen and Bouwmeester 2014).

A different view regarding the cognitive processes of social decision making is taken by the sequential sampling models, which assume that a noisy relative decision

value is integrated at each moment in time and a choice is made when this accumulated decision value crosses a threshold. The sequential sampling models were developed for perceptual decision making (Ratcliff 1978, Ratcliff and Smith 2004) and recently adapted to the analysis of economic and in particular social decisions (Dickhaut et al. 2013, Frydman and Nave forthcoming, Hutcherson, Bushong, and Rangel 2015, Krajbich, Armel, and Rangel 2010, Krajbich, Bartling, et al. 2015a, Krajbich, Oud, and Fehr 2014). These studies indicate that there exists a common computational mechanism underlying economic (social) decision making and perceptual decision making. Importantly, sequential sampling models assume that there is only one single deliberative process governing social decisions but not competing processes. The strength of preference which is based on the utility difference between choice options determines RTs and stronger preference over one option requires shorter RTs. Moreover, the RT of selfish decisions is not significantly different from that of social decisions after controlling for the strength of preference (Krajbich, Bartling, et al. 2015a).

Which explanation of the cognitive process is correct has major implications for understanding human nature, and from a practical point of view for modeling the cognitive process to predict human behavior or designing institutions to promote prosocial behavior (Bear and Rand 2016, Cone and Rand 2015, Krajbich, Hare, et al. 2015, Krajbich, Oud, and Fehr 2014). Therefore, it is crucial to identify the cognitive processes of social decision making. This chapter studies this topic in depth by taking the heterogeneity of preferences into account and analyzing RTs in distribution decisions. It is a well-established fact that people are heterogeneous in the relevant motives and in the strength of preferences (Andreoni and Miller 2002, Engelmann and Strobel 2004b, Erlei 2008, Fisman, Kariv, and Markovits 2007b, Kerschbamer 2015). For instance, some people care more about efficiency, while others care more about fairness (Murphy, Ackermann, and Handgraaf 2011). We explicitly take individual heterogeneity into account, not only with respect to what kind of social motives people care about, but also with respect to the strength of people's selfish motives. Our experiment includes two types of binary three-person dictator games: the third-party (TP) dictator game and the second-party (SP) dictator game. In TP games, the dictator's payoffs are the same in the two allocations, while in SP games, the dictator's payoffs differ between the two allocations. Decisions in TP games allow us to identify subjects' social motives. Based on this identification, we can study how people with different social motives react to various decision situations in SP games using RT analysis. In addition, we can study which kind of motive, the selfish or the social motives, can be considered as more intuitive.

We classify subjects into three norm types which differ with respect to the relevant social motives based on the decisions for TP games. Based on this identification, a within-subject analysis on the SP games shows that the RT increases with the number of conflicts between individually relevant motives and decreases with the utility difference between choice options. These results are in line with the predictions of sequential sampling models. A between-subject analysis reveals the heterogeneity with respect to whether the selfish or the social motive is more intuitive. It turns out that the selfish motive is more intuitive for subjects who are more selfish. We also show that the heterogeneity in social motives and selfishness is reflected in process differences. Our findings demonstrate that the cognitive processes of social decision making comply with both, sequential sampling models and dual-process theory. And our findings also show that it is important to take the heterogeneity of preferences into account when investigating the cognitive processes of social decision making.

Our study contributes to the emerging and conflicting literature on the cognitive underpinnings of social decision making. The experiment allows identifying both the heterogeneity in social motives and the strength of selfishness. Based on this identification, we are not only able to provide evidence for sequential sampling models, but we can also show the heterogeneity in whether the selfish or the social motive is related to the intuitive process. This could help to reconcile the mixed results on the correlations between RT and prosociality, and resolve the debate between the dual-process explanation and the sequential sampling explanation of the cognitive processes of social decision making. We argue that conflicts between individually relevant motives, the strength of preference, and the intuition or the deliberation of selfishness all contribute to the variations of RT. Thus, it is crucial to take these factors and the heterogeneity of preferences into account when investigate the cognitive processes of social decision making.

1.2 Experimental Design and Procedures

In this section, we first describe the game that we use in our experiment and then provide a detailed description of our experimental procedures.

1.2.1 Experimental Design

The experiment consists of a series of 64 binary three-person dictator games.[1] In each of these dictator games, a subject (dictator) decides between two predefined allocations (A_1, A_2, A_3) and (B_1, B_2, B_3), which determine how money is distributed between herself (player 2) and the other two subjects in her group. The other two subjects have no choice. 32 of the 64 games are third-party (TP) dictator games, in which the dictator's payoff does not differ between the two options. In TP games we can assess the importance of different social motives – unaffected by the selfish motive – and define a 'personal norm', which refers to a person's purely social motive. These social motives include efficiency, maximin, envy, disadvantageous inequality aversion (FS-α) and advantageous inequality aversion (FS-β). The social motives are defined as follows: Efficiency maximizes the total payoff of all subjects in the group. Maximin maximizes the minimum payoff of all subjects in the group, guaranteeing that no one is left in a very bad position. Envy is to minimize the difference between one's payoff and the highest payoff of others (Engelmann and Strobel 2004b). The minimization of disadvantageous inequality corresponds to a high value of the α-term in the model by Fehr and Schmidt (1999) and the minimization of advantageous inequality corresponds to a high value of the β-term in the model by Fehr and Schmidt (1999). The conflict between the selfish and the social motives is studied in another 32 second-party (SP) dictator games in which the payoff of the dictator varies. Table 1.1 gives examples for the two types of the binary three-person dictator games.

Table 1.1 Examples of the two types of games

Player	Third party (TP) game		Second party (SP) game	
	Option A	Option B	Option A	Option B
1	$A_1 = 5$	$15 = B_1$	$A_1 = 7$	$18 = B_1$
2 (Dictator)	$A_2 = 14$	$14 = B_2$	$A_2 = 14$	$13 = B_2$
3	$A_3 = 19$	$19 = B_3$	$A_3 = 13$	$20 = B_3$

We chose the 64 games systematically in such a way that different combinations of the previously suggested motives are represented in the games. We presented the six payoffs that describe each decision situation in a numeric form as well as a graphical form, in order to make them quickly accessible. The screen layout is shown in Figure 1.16 in Appendix.

[1] Table 1.6 lists the 64 games.

1.2.2 Procedural Details

The experiment was computerized and conducted with the software "z-Tree" (Fischbacher 2007). We conducted four sessions in October and November 2013 at the Lakelab of the University of Konstanz. 105 subjects participated in the experiment.[2] Subjects were recruited using the online recruitment system "ORSEE" (Greiner 2015). Each session lasted about 50 minutes and none of the subjects participated in more than one session. Upon entering the laboratory, subjects were randomly assigned to a PC terminal and were given a copy of instructions (Appendix 1.6). A set of control questions was provided to ensure the understanding of the decision task. Questions were answered individually at subjects' own seats. The experiment did not start until all subjects had answered all questions correctly. At the end of a session, subjects were asked to fill out a socio-economic questionnaire.

To avoid order effects, we randomized the sequence of the 64 games for each subject. This means in particular that we also mixed the SP and TP games. In addition, Option A and Option B of each game were also randomly reshuffled. All subjects made their decisions as dictators by pressing key "F" or "J" on the keyboard. We recorded RTs on the server. The RT measured the time between when the allocation was sent to the client and when the server received the message that the key was pressed.[3] After each decision, subjects saw a waiting screen and were required to press the "Spacebar" to advance to the next decision. At the end of the experiment, the roles of the three players in each group were randomly determined. One of the 64 games was randomly selected and paid out according to their random roles. On average, each subject earned 9.92 Euros which included a show-up fee of 3 Euros.

1.3 Main Results

In this section, we first characterize the heterogeneity in social preferences by conducting finite mixture analysis on subjects' decisions in TP games. Then, we analyze the RTs of SP decisions to investigate the cognitive processes of distributional preferences. Finally, we show that the heterogeneity in preferences is also reflected in the differences of cognitive processes.

[2] In order to address heterogeneity, our sample size is sufficiently larger than Hutcherson, Bushong, and Rangel (2015), who have 51 subjects actively making decisions.
[3] This causes some delay in comparison to the pure response time. However, the delay is completely uncorrelated to the decision situation because their sequence was randomized.

1.3.1 Heterogeneity of Social Preferences

In order to deal with the heterogeneity and to determine individual social preferences, we analyze TP decisions using finite mixture analysis.[4] Moreover, in the finite mixture analysis we use four different structural models to check the robustness of our estimation. The results of the finite mixture analysis are robust. All four models show that the optimal number of types equals three according to the Normalized Entropy Criterion (Celeux and Soromenho 1996), and all four models create the same classification. In addition, all subjects can be assigned to one distinct type with high posterior probability. These clean classifications suggest that our analysis is able to capture the distinctive characteristics of each preference type. The results of the finite mixture analysis are shown in Appendix 1.5.2. In the following, we will base our analysis on model IV in Table 1.7, which performs best according to the BIC criterion. However, the main results are robust with respect to the choice of the model. We identify subjects' social motives according to whether the coefficient of the motive is significant or not in Table 1.7. The relevant motives for each norm type in Model IV are summarized in Table 1.2.

Table 1.2 Pure social motives of each norm type

Norm type	Social motives	Number of subjects
I	Efficiency, Maximin	10
II	Envy, Maximin	16
III	Efficiency, Maximin, FS-β	79

1.3.2 Response Time Analysis

In this section we investigate the RTs of SP decisions based on the identification of heterogeneity in preferences[5]. We first show that the cognitive processes of distributional preferences comply with the sequential sampling models. Then we show that the selfish motive is more intuitive for subjects who are more selfish.

1.3.2.1 Evidence for Sequential Sampling Models

In this section, we present evidence that the cognitive processes of distribution decisions comply with sequential sampling models (Ratcliff 1978, Ratcliff and Smith

[4] We conducted the finite mixture analysis using R package flexmix (Grun and Leisch 2008).
[5] The evolution and the distributions of RTs are shown in Figure 1.8 and Figure 1.9.

2004). Sequential sampling models originate from psychology and neuroscience, and describe the dynamics of the decision process. These models assume that there is a value representing the collected evidence in favor of a particular decision. This value is stochastically sampled over time, and the choice is executed when the value hits one of two pre-specified boundaries. Sequential sampling models predict that the probability of choosing an option is a sigmoidal function of the value (utility) difference between the two options, and RT increases with the difficulty (measured by the value difference between the two options) of the decision.

In this part, we focus on the question that how cognitive conflicts and the utility difference between the two options affect RT. We use the relevant motives identified before and define decision situation as conflicting if there is a conflict between these motives. And we use the latent variable of choosing Option A assessed by regressions as a measure of the utility difference. The utility difference between choice options measures the strength of preferences.

First, we compare the RT of conflict and consistent decisions in SP games. Conflict and consistent decisions are defined according to whether the subject's behavior is consistent with all her own motives or not. *Conflict decisions* are decisions for the situation in which the subject's behavior is consistent with some of her motives but in conflict with the other motives. *Consistent decisions* are decisions for the situation in which the subject's behavior is consistent with all her motives (Fig. 1.1). We expect that conflict decisions take longer than consistent decisions. First, the conflicts between individually relevant motives make the decision situation more complex. Second, the average utility difference between the two options in conflict situations is smaller than that in consistent situations. According to sequential sampling models, it takes more time to accumulate evidence to reach the threshold and make a decision if the utility difference is small. The left panel of Figure 2 displays the mean log(RT) of conflict and consistent decisions in SP games.[6] Each point represents one subject. Most of the data points are above the 45° line. That is, the RTs of conflict decisions are longer than the RTs of consistent decisions for most of the subjects (Wilcoxon sign-rank test, $p < 10^{-9}$).

[6] To test the robustness of our results, we also conducted the analysis using untransformed RT, which essentially leads to the same results.

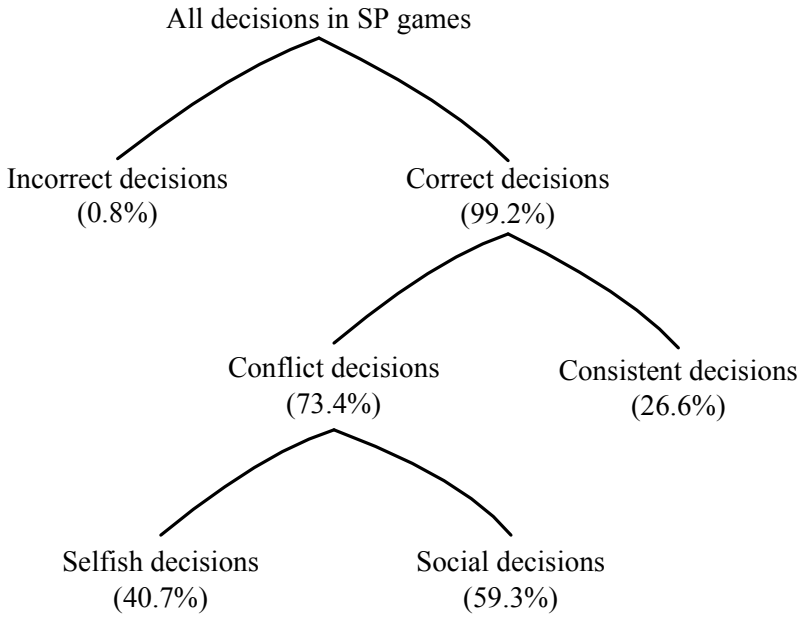


Figure 1.1 Classification of all decisions in SP games
A decision is incorrect if it neither follow the personal norms nor the selfish motives. If the personal norms and the selfish motives are in line, it is considered as consistent otherwise as a conflict decision. The conflict decision is selfish or social decision depending on which motive people follow in that decision. More details about how the decisions are classified are shown in Table 1.9.

Apart from the binary measure of decision conflict, we also use *the number of conflicts*. We get the number of conflicts by considering the pairwise comparisons of the individually relevant motives. The number of conflicts for each type of subjects in SP decisions are shown in Table 1.9. In conflict situations this number always equals one if two motives are involved, and always equals two if three motives are involved. Thus, the number of conflicts only varies for subjects of Type III, which is the largest group. In these decisions, the number of conflicts can be three or four. The right panel of Figure 1.2 displays the mean log(RT) of decisions with three conflicts and four conflicts for each subject. It shows that most of the subjects are above the 45° line. That is, the decisions with four conflicts take longer than the decisions with three conflicts (Wilcoxon sign-rank test, $p < 10^{-7}$).

Turning to the utility difference between choice options, the sequential sampling models predict that the closer the utilities of the two options, the longer time subject needs to make a decision. We construct a measure of utility difference by calculating the latent variable of choosing Option A. To do an out-of-sample analysis, we conducted regressions for the *even* trials of SP decisions with the dummy *Decision* as the dependent variable explained by variables reflecting the selfish incentive and the latent variables

derived from TP decisions.[7] Then we use the coefficients of these regressions to calculate the latent variable of choosing Option A for the *odd* trials of SP decisions and take the latent variable as the measure of the utility difference between choice options. However, the results are similar if we use the data of the odd trials to predict the RT in the even trials, which are shown in Appendix 1.5.4.

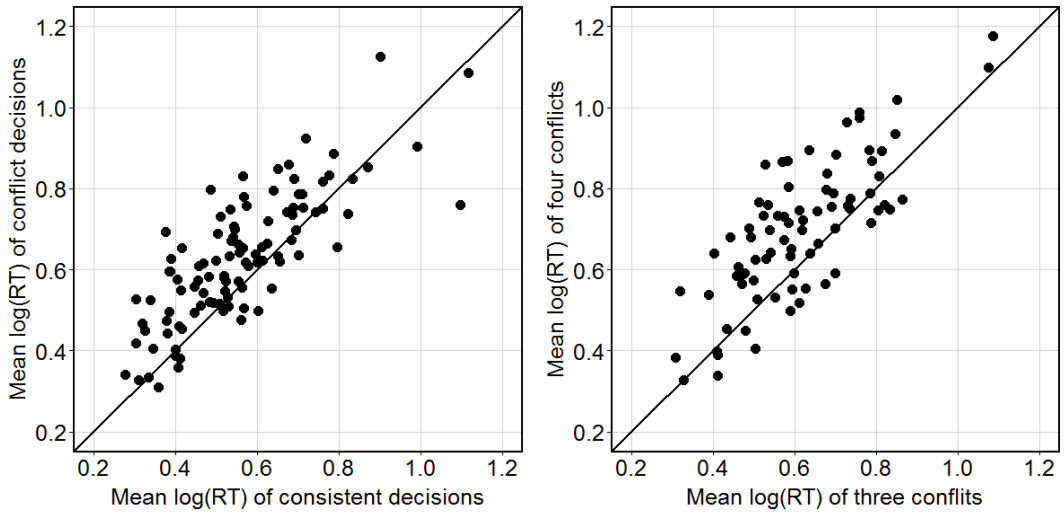


Figure 1.2 RTs of decisions with different number of conflicts in SP games

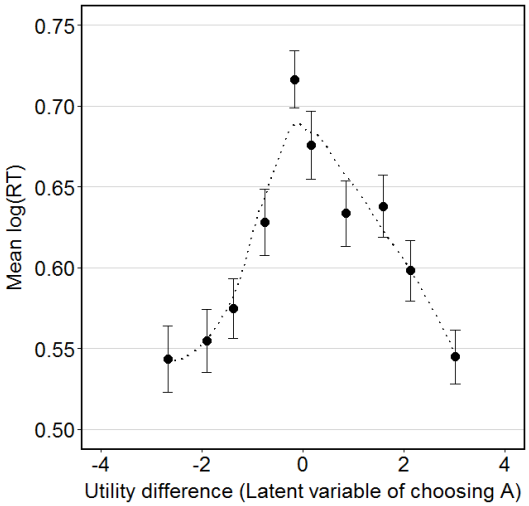


Figure 1.3 Mean log(RT) and utility difference in the odd trials of SP decisions
The dotted line is the smoothing line using method "loess".

Figure 1.3 displays the relationship between the mean log(RT) and the utility difference in the odd trials of SP decisions. All the data in the odd trials of SP decisions is divided into 10 bins of equal size. Each dot represents one bin and the solid line is the

[7] The regression results are shown in Table C2.

standard error of the mean log(RT). Figure 1.3 shows that the mean log(RT) decreases as the utility difference increases. The curve peaks at the utility difference of around 0, and falls off steadily as the utility difference increases in either direction. The correlation is highly significant (Pearson's correlation test, $r = -0.205$, $p < 10^{-15}$). But since the observations are not independent we test the correlation at the individual level. The Pearson correlation coefficient between the mean log(RT) and the absolute value of the utility difference is positive for 77 of 105 subjects, which is significantly different from the chance level of 50% (two-sided binomial test $p < 10^{-5}$).

The dependence of RTs on the number of conflicts and the utility difference can also be confirmed by mixed-effects regressions. The regression results are shown in Table 1.3. The dependent variable is *log*(*RT*). The coefficient of the *conflict decision* dummy is positive and significant in regressions (1) and (3), and the coefficient of the *number of conflicts* is positive and significant in regressions (2) and (4). This indicates that the cognitive conflict between individually relevant motives has significantly positive effects on the RTs. With respect to the utility difference, the coefficients are negative and highly significant in regressions (3) and (4). That is, the RT decreases as the utility difference increases. All these results indicate that the RT of distribution decisions is in line with the predictions of sequential sampling models. We now summarize the evidence for sequential sampling models.

Table 1.3 Mixed-effects regressions of RTs

	All SP decisions		The odd trials of SP decisions	
	All subjects	Type III	All subjects	Type III
	(1)	(2)	(3)	(4)
Constant	0.659***	0.504***	0.772***	0.709***
	(0.017)	(0.038)	(0.023)	(0.059)
Conflict decision	0.076***		0.040***	
	(0.008)		(0.012)	
Decision number	-0.003***	-0.004***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Number of conflicts		0.076***		0.036**
		(0.010)		(0.014)
abs(Utility difference)			-0.052***	-0.047***
			(0.006)	(0.008)
AIC	-976.750	-427.512	-451.559	-205.120
BIC	-946.198	-399.982	-419.047	-176.279
Log Likelihood	493.375	218.756	231.780	108.560
Num. obs.	3332	1822	1671	908
Num. groups	105	79	105	79

Notes. The dependent variable is log(RT). ***p < 0.01, **p < 0.05, *p < 0.1.

Result 1. *The cognitive processes of distributional preferences comply with the sequential sampling models. Specifically, the response time increases with the number of conflicts between individually relevant motives and decreases with the utility difference between the choice options.*

1.3.2.2 The Intuition and Deliberation of Selfishness

Dual-process theories postulate that two interacting processes produce decisions. One of the processes is more automatic and intuitive, and the other is more controlled and deliberative. Thus, people's reasoning does not consistently conform to a rationality norm. This section studies the relevance of dual-process models and tests whether selfishness or prosociality can be associated with the intuitive process when people decide in situations in which the selfish motive conflicts with their social motives.

According to dual-process theories, decisions which are more associated with the intuitive process should be quicker than decisions that are more associated with the deliberative process. However, evidence on which motive, the selfish or the social motive, is more related to the intuitive process is still disputed. Evans, Dillon, and Rand (2015) and Krajbich, Bartling, et al. (2015a) have shown that the identification of the process using RT is difficult because it interacts with the strength of conflict and the utility difference between choice options. Therefore, we will control for the strength of conflict and the utility difference in our analysis. There is a second reason for the conflicting evidence on what is the intuitive process. People may differ with respect to what is the intuitive response to them. Our experimental design allows us to take into account both the heterogeneity in social motives and the heterogeneity in selfishness.

To compare the RTs of selfish and social decisions, we exclude all consistent decisions and focus only on decisions in which the selfish motive is in conflict with the social motive. We classify all conflict decisions in SP games into *selfish decisions* and *social decisions* according to whether the decision is consistent with the selfish motive or not (Fig. 1.1). According to dual-process theories, the selfish decisions should be quicker than the social decisions if the selfish motive is more intuitive and the social decisions should be quicker than the selfish decisions if the social motive is more intuitive. To test these predictions, we conducted a mixed-effects regression with $log(RT)$ as the dependent variable. The regression result is shown in regression (1) of Table 1.4. *Decision number* controls for learning, and *Utility difference* as well as the variable *Conflict within norms* (dummy which indicates whether there is a conflict within the social motives) control for the strength of preference and the strength of conflict of the decision. The results show

that the RT does not differ between selfish and social decisions across all subjects if we control for variables *Utility difference* and *Conflict within norms*. This indicates there is no evidence that the selfish or the social decision is more intuitive if we ignore the heterogeneity in selfishness.

In the next step we take the heterogeneity of selfishness into account and study whether the selfish motive is more intuitive for some people and deliberative for others. Specifically, we study how the strength of selfishness influences subject's intuition towards the selfish or the social motive. We calculate the proportion of selfish decisions in the *even* trials of conflict decisions for each subject, and take this proportion as the measure of the strength of selfishness. Then we study the relationship between the RTs of selfish and social decisions with the strength of selfishness in the *odd* trials of conflict decisions. Figure 1.4 shows the distribution of the strength of selfishness in the even trials of conflict decisions. We expect that the selfish motive is more intuitive for the subjects who are more selfish. More specifically, we expect that more selfish subjects are quicker in making selfish decisions than social decisions.

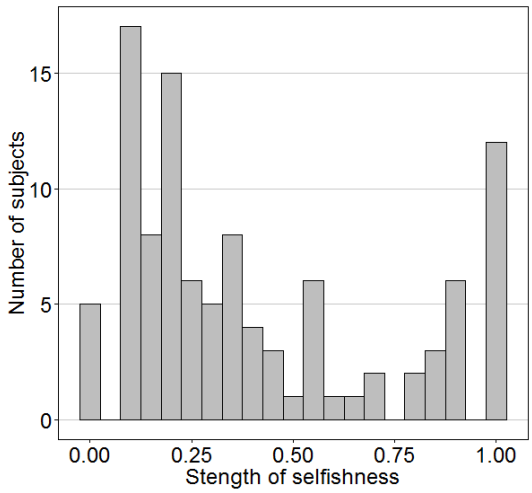


Figure 1.4 Distribution of the strength-of-selfishness (even trials of conflict decisions)

The left panel of Figure 1.5 displays the relationship between the RTs of selfish (social) decisions and the strength of selfishness in the odd trials of conflict decisions. It shows that subjects who are more selfish are quicker in making selfish decisions (Pearson's correlation test, $r = -0.391$, $p < 10^{-15}$). But subjects who are more selfish are not significantly slower in making social decisions (Pearson's correlation test, $r = 0.059$, $p = 0.117$). The right panel of Figure 1.5 displays the time difference between selfish and social decisions. It shows that the selfish decisions are quicker than social decisions for

subjects who are more selfish and the selfish decisions are slower than social decisions for subjects who are more prosocial (Pearson's correlation test, $r = -0.450$, $p < 10^{-5}$). These results suggest that the selfish motive is more intuitive for subjects who are more selfish, and the selfish motives are more deliberative for subjects who are more prosocial.

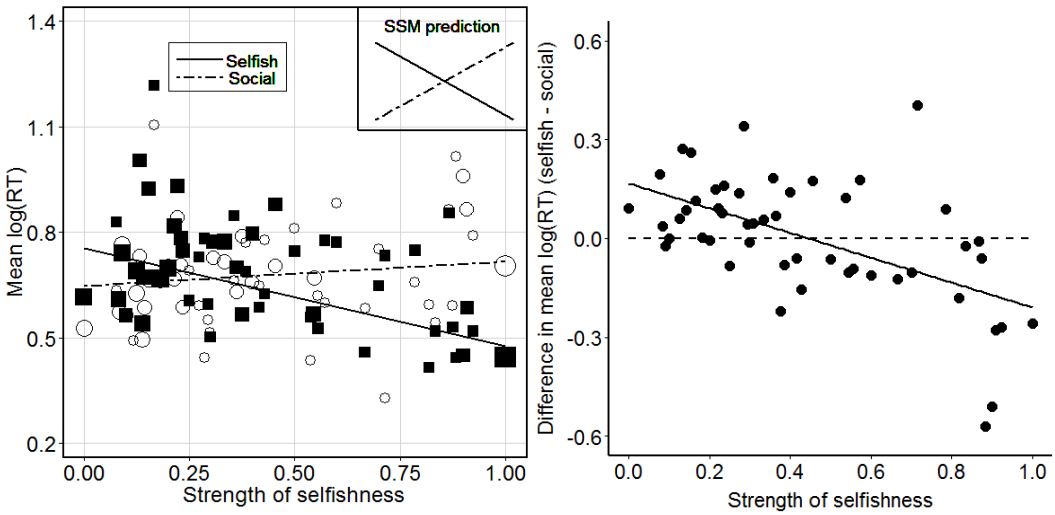


Figure 1.5 RTs of selfish and social decisions in the odd trials of conflict decisions
The size of the point or the square indicates the number of subjects in that point or square. The solid and dotted lines are regression lines.

The econometric analysis in Table 1.4 corroborates these results. Regression (3) in Table 1.4 shows that the RTs of selfish decisions significantly decrease with the strength of selfishness. However, the coefficient of strength of selfishness is not significant in regression (3). That is, the strength of selfishness has no significant effects on the RTs of social decisions. In a recent study, Rand et al. (2016) show that the intuition favors prosociality for women but not for men. We also control the gender in our analysis. As shown in regressions (2) and (3), the RTs of selfish or social decisions of male are not significantly different from those of female.

Table 1.4 Regressions of RTs in the odd trials of conflict decisions

	Mixed-effects regression	OLS regressions	
	(1)	(2)	(3)
Constant	0.787***	0.825***	0.780***
	(0.021)	(0.031)	(0.034)
Selfish decision	-0.011	-0.018	0.127***
	(0.016)	(0.029)	(0.041)
Strength of selfishness		-0.134***	0.103
		(0.047)	(0.067)
abs(Utility difference)	-0.045***	-0.035***	-0.045***
	(0.007)	(0.010)	(0.010)
Conflict within norms	0.044***	0.055***	0.038***
	(0.013)	(0.016)	(0.016)
Male		0.029	0.025
		(0.033)	(0.031)
Decision number	-0.003***	-0.003***	-0.004***
	(0.000)	(0.000)	(0.000)
Selfish decision × Male		-0.057	0.041
		(0.039)	(0.035)
Selfish decision × Strength-of-selfishness			-0.373***
			(0.068)
Num. obs.	1207	1207	1207
Num. groups	105	105	105

Notes. The dependent variable is *log*(*RT*). *Selfish Decision* is a dummy variable which indicate the decision is a selfish decision or social decision. *Male* is a dummy variable which indicate the gender. The robust standard errors for regressions (2) and (3) are clustered on subjects and reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Our results show that selfishness interacts with the speed of selfish vs. social decisions. We control for the strength of the conflict and the strength of preference, which deals with the argument brought up by Krajbich, Bartling, et al. (2015a) and Evans, Dillon, and Rand (2015). On the other hand, sequential sampling models also predict that selfish decisions is quicker for subjects who are more selfish since more selfish people put higher weight on the selfish dimension (Dickhaut et al. 2013, Frydman and Nave forthcoming, Hutcherson, Bushong, and Rangel 2015, Krajbich, Armel, and Rangel 2010, Krajbich, Bartling, et al. 2015a, Krajbich, Oud, and Fehr 2014). However, sequential sampling models predict that social decisions are slower for subjects who are more selfish as well. That is, the relationship between the RTs of selfish decisions and the strength of selfishness should be symmetric with the relationship between the RTs of social decisions and the strength of selfishness, as shown in the upper right corner of the left panel of Fig. 1.5 (SSM prediction). Obviously, the data is different from the prediction by the sequential sampling models. Of course, the control is incomplete because the utility difference and the strength of the conflict are constructed at the norm type level but not at the individual level. Eye tracking provides an independent method to assess what motive is more intuitive. Subject for which selfishness is intuitive should look at the own payoff. We have not collected eye tracking data ourselves but we analyze the eye-tracking data from

Fiedler et al. (2013) and the results show that at the beginning of the decision process, selfish subjects put more attention (80.1%) on their own payoffs compared to prosocial subjects (60.1%)(Fig. 1.13).[8] To look at each fixation individually, we also plot the proportion of subjects who fixate on their own payoffs in each fixation. It shows that 78.6% of selfish subject who fixate on their own payoffs in the first fixation, while 58.5% of prosocial subjects who fixate on their own payoffs in the first fixation. Over time people care about both payoffs equally. From the fifth fixation, the proportion of subjects who fixate on the own payoff is around 50% both for selfish and prosocial subjects (Fig. 1.14). We now summarize our evidence for the intuition and deliberation of selfish motives.

Result 2. *The direction and extent of the intuition towards selfishness depends on the selfishness of the subjects. The selfish motive is more intuitive for subjects who are more selfish and the selfish motive is more deliberative for subjects who are less selfish.*

1.3.3 Why Heterogeneity is Important

Individual heterogeneity in social preferences is not at all controversial. But data limitations have often forced us to assume homogeneity. We take into account both the heterogeneity in social motives and the heterogeneity in selfishness. In this section, we present evidence that the heterogeneity in preferences is reflected in the corresponding heterogeneity of cognitive processes. The results in the previous section show that the heterogeneity in selfishness explains different RT patterns. In Table 1.4, we have shown that if we assume homogeneity, selfish and social decisions seem to have similar RTs. However, the heterogeneity in selfishness is also reflected in the heterogeneity of the underlying decision processes – reflected in RT.

In the following, we show that taking heterogeneity into account improves process predictions. We compare the explanatory power of the variable *utility difference* on the RTs of the odd trials of SP decisions. The utility difference based on the standard Logit model is calculated using the coefficients in Tables 1.12 and 1.13.[9] Since the standard Logit model neglects heterogeneity, we expect that the utility difference based on the finite mixture model has a higher explanatory power.

The results are shown in Table 1.5. Since there is only one parameter in regressions (1) and (2), the R^2 in regressions (1) and (2) can be compared directly. The explanatory

[8] The results are shown in 1.5.6 in Appendix.
[9] The calculation is similar to the utility difference based on heterogeneity.

power of the utility difference based on the finite mixture model (adjusted R^2 = 0.042) is higher than the explanatory power of the utility difference based on the standard Logit model (adjusted R^2 = 0.038). If the two variables are simultaneously included in the regression, the coefficient of the utility difference based on the finite mixture model is more robust than the coefficient based on the standard Logit model, and the utility difference based on the standard Logit model is not significant at all. We now summarize our third result.

Result 3. *The analysis based on the heterogeneity outperforms the analysis based on the homogeneity in explaining response times. The heterogeneity in preferences is reflected in the differences of cognitive processes.*

Table 1.5 OLS regressions of RT on the utility difference (odd trials of SP decisions)

	(1)	(2)	(3)
Constant	0.690***	0.693***	0.702***
	(0.021)	(0.018)	(0.019)
abs(Utility difference FMM)	-0.053***		-0.034*
	(0.009)		(0.018)
abs(Utility difference Logit)		-0.065***	-0.032
		(0.008)	(0.020)
R^2	0.042	0.039	0.046
Adj. R^2	0.042	0.038	0.045
Num. obs.	1683	1683	1683

Notes. The dependent variable is *log*(*RT*). *Utility difference based on FMM* is the utility difference based on finite mixture model. *Utility difference based on Logit* is the utility difference based on standard Logit model. The robust standard errors are clustered on subjects and reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

1.4 Conclusion

This chapter studies the cognitive mechanism of distributional preferences by investigating subjects' RTs in a series of binary three-person dictator games. Our experiment takes into account, both, the heterogeneity in the relevant social motives as well as the heterogeneity in the strength of selfishness. We find evidence for both, the sequential sampling models and the dual-process theories. First, the results show that the RT increases with the number of conflicts between individually relevant motives and decreases with the strength of preferences, which is predicted by the sequential sampling models. Second, the more selfish subjects are quicker in making selfish decisions than the less selfish subjects. This is in line with a dual-process approach with heterogeneity in whether the selfish motive is intuitive or deliberative: the selfish motive is intuitive for subjects who are more selfish and deliberative for subjects who are more prosocial. Our

study provides an explanation for the conflicting results concerning the automaticity of the selfish motive observed in the previous literature, as well as the debate between the dual-process explanation and the sequential sampling explanation of the cognitive processes of social decision making.

It is undisputed that people are heterogeneous in their preferences but explicitly taking heterogeneity into account is still rare. Our study not only shows the analysis based on the heterogeneity outperforms the analysis based on the homogeneity, but also shows that the heterogeneity in preferences is reflected in the differences of the cognitive processes. In particular, people differ in what their automatic response is. Thus, in order to identify the correct process model, taking the heterogeneity into account can be indispensable. For example, if our data were analyzed assuming homogeneity, no evidence for a dual-process model could be detected. Thus, the heterogeneity is not only crucial for the determination of the parameters but even for the choice of the model.

We have to admit that the RT analysis does not allow to draw causal inference. However, the sequential sampling models make clear predictions that conflicts between motives increases RT and less utility difference between choice options need more time. This prediction could clearly be confirmed. The predictions of dual-process models are less clear, in particular if one assumes heterogeneity in the process. Nevertheless, it is reasonable to assume the motive that is more relevant behaviorally, is also more intuitive in the sense of dual-process models. We indeed find evidence that the processing of the selfish motive is more intuitive for more selfish subjects. Nevertheless, causal tests of dual-process models have to rely on intervention methods like time pressure or cognitive load.

1.5 Appendix

1.5.1 The 64 Games

Table 1.6 Sixty-Four Games in the Experiment

Type	Game ID	Option A			Option B			Signs					
		A1	A2	A3	B1	B2	B3	Selfishness	Efficiency	Maximin	Envy	FS-α	FS-β
TP	1	6	14	7	6	14	19	0	-1	0	1	1	-1
	2	7	18	9	7	18	20	0	-1	0	1	1	-1
	3	11	16	14	13	16	18	0	-1	-1	1	1	-1
	4	10	14	13	12	14	18	0	-1	-1	1	1	-1
	5	5	11	20	15	11	16	0	-1	-1	-1	0	-1
	6	6	13	19	16	13	16	0	-1	-1	-1	0	-1
	7	9	14	19	12	14	18	0	-1	-1	-1	-1	-1
	8	5	9	18	10	9	16	0	-1	-1	-1	-1	-1
	9	6	12	19	16	12	19	0	-1	-1	0	1	-1
	10	5	14	19	15	14	19	0	-1	-1	0	1	-1
	11	11	13	18	15	13	17	0	-1	-1	-1	1	-1
	12	8	9	19	15	9	17	0	-1	-1	-1	1	-1
	13	5	16	20	19	16	19	0	-1	-1	-1	1	-1
	14	5	15	18	17	15	17	0	-1	-1	-1	1	-1
	15	13	16	17	17	16	17	0	-1	-1	0	1	-1
	16	9	14	16	15	14	16	0	-1	-1	0	1	-1
	17	7	10	20	8	10	8	0	1	-1	-1	-1	1
	18	8	11	20	9	11	9	0	1	-1	-1	-1	1
	19	7	9	8	7	9	20	0	-1	0	1	1	-1
	20	14	15	14	14	15	19	0	-1	0	1	1	-1
	21	6	20	19	8	20	8	0	1	-1	0	0	1
	22	8	19	18	9	19	9	0	1	-1	0	0	1
	23	12	18	19	13	18	14	0	1	-1	-1	-1	1
	24	7	15	19	10	15	11	0	1	-1	-1	-1	1
	25	5	11	20	11	11	11	0	1	-1	-1	-1	-1
	26	6	11	19	7	11	14	0	1	-1	-1	-1	-1
	27	11	15	19	12	15	17	0	1	-1	-1	-1	-1
	28	6	13	20	12	13	12	0	1	-1	-1	-1	-1
	29	11	16	15	20	16	20	0	-1	-1	1	1	-1
	30	8	16	18	20	16	20	0	-1	-1	1	1	-1
	31	12	15	17	16	15	18	0	-1	-1	1	1	-1
	32	8	9	15	12	9	19	0	-1	-1	1	1	-1
SP	33	7	11	9	7	13	19	-1	-1	0	1	1	0
	34	10	15	11	10	17	19	-1	-1	0	1	1	-1
	35	9	16	14	11	17	19	-1	-1	-1	1	1	-1
	36	8	12	11	10	13	19	-1	-1	-1	1	1	-1
	37	3	14	20	15	12	15	1	-1	-1	-1	0	-1
	38	5	14	20	16	13	16	1	-1	-1	-1	0	-1
	39	7	13	18	13	12	15	1	-1	-1	-1	-1	-1
	40	2	13	20	13	11	15	1	-1	-1	-1	-1	-1
	41	7	8	16	17	10	18	-1	-1	-1	0	1	-1
	42	6	8	12	11	10	14	-1	-1	-1	0	1	-1
	43	12	13	19	17	14	18	-1	-1	-1	-1	1	-1
	44	6	7	15	15	8	15	-1	-1	-1	-1	1	-1
	45	5	17	20	17	15	17	1	-1	-1	-1	1	-1
	46	5	17	20	18	16	18	1	-1	-1	-1	1	-1
	47	6	13	16	13	12	15	1	-1	-1	0	1	-1
	48	5	16	20	15	14	18	1	-1	-1	0	1	-1
	49	12	19	20	13	16	14	1	1	-1	-1	-1	-1
	50	8	17	18	9	13	11	1	1	-1	-1	-1	-1
	51	10	12	11	10	16	17	-1	-1	0	1	1	1
	52	8	9	8	8	10	19	-1	-1	0	1	1	0
	53	5	18	17	6	20	6	-1	1	-1	0	0	1
	54	7	19	18	8	20	8	-1	1	-1	0	0	1
	55	10	17	18	11	18	12	-1	1	-1	-1	-1	1
	56	8	18	19	9	20	10	-1	1	-1	-1	-1	1
	57	3	14	20	12	12	12	1	1	-1	-1	-1	-1
	58	5	12	17	10	11	12	1	1	-1	-1	-1	-1
	59	6	13	20	11	12	14	1	1	-1	-1	-1	-1
	60	5	12	20	11	11	13	1	1	-1	-1	-1	-1
	61	10	13	10	16	12	16	1	-1	-1	1	1	-1
	62	7	14	13	18	13	20	1	-1	-1	1	1	-1
	63	5	16	19	19	14	20	1	-1	-1	1	1	-1
	64	5	16	20	18	14	19	1	-1	-1	1	1	-1

Notes. In the columns of signs, 1 indicates that the motive favors Option A, -1 indicates the motives favors Option B, and 0 means the two options are indifferent for that motive. For example, in the fourth game, the dictator could choose between allocations of Option A (10, 15, 11) and Option B (10, 17, 19), where the parameters in each option refers to the payoffs for the first players, the dictator and the third player. The *efficiency* motive favors Option B, the *envy* motive favors Option A, and the two options are indifferent for the *maximin* motive.

1.5.2 Finite Mixture Estimation

In this part, we first present the social motives that we use in the finite mixture model. Then, we introduce the estimation procedure for the finite mixture analysis.

1.5.2.1 The Decision Motives

People differ in the nature of their distributional preferences. The TP games allow to deal with the heterogeneity of distributional preferences, even for selfish subjects. In each situation of the experiment, the subject made a binary decision by choosing Option A or Option B according to her personal norm. We use a logit model to capture the importance of potential social motives in the estimation of the finite mixture model. The dependent variable of the logit model is the dummy variable *Decision* which indicates whether the subject chose Option A or not. The independent variables are the differences between the strength of different motives and the signs of these differences. The differences between Option A and Option B on each motive are calculated as follows:

$DiffEfficiency = (A_1 + A_2 + A_3) - (B_1 + B_2 + B_3)$;

$DiffEnvy = [\max(B_1, B_2, B_3) - B_2] - [\max(A_1, A_2, A_3) - A_2]$;

$DiffMaximin = \min(A_1, A_2, A_3) - \min(B_1, B_2, B_3)$;

$DiffFS\text{-}\beta = 1/2^{*}\{[\max(B_2 - B_1, 0) + \max(B_2 - B_3, 0)] - [\max(A_2 - A_1,0) + \max(A_2 - A_3,0)]\}$;

$DiffFS\text{-}\alpha = 1/2^{*}\{[\max(B_1 - B_2, 0) + \max(B_3 - B_2, 0)] - [\max(A_1 - A_2,0) + \max(A_3 - A_2,0)]\}$;

$DiffSelfish = A_2 - B_2$.

A_i and B_i are the payoffs for Player *i* in Option A and Option B. We take the signs of these difference as the signs of motives. That is, the signs of motives are discrete variables (-1, 0, 1).

In order to check the robustness of our estimation, we use four different structural models in the finite mixture analysis:

Model I: $Decision \sim SignEfficiency + DiffEfficiency + SignEnvy + DiffEnvy + SignMaximin + DiffMaximin + SignFS\text{-}\alpha + SignFS\text{-}\beta + DiffFS\text{-}\beta$;

Model II: $Decision \sim SignEfficiency + SignEnvy + SignMaximin + SignFS\text{-}\alpha + SignFS\text{-}\beta$;

Model III: $Decision \sim SignEfficiency + DiffEnvy + SignMaximin + DiffMaximin + SignFS\text{-}\beta$;

Model IV: $Decision \sim SignEfficiency + DiffEnvy + SignMaximin + SignFS\text{-}\beta$.

In the first model, we include all the variables of signs and differences on social motives. However, we have to leave out the variable *DiffFS-α* in order to avoid the collinearity problem.[10] The second model includes the signs of all the social motives. The independent variables of the third model are the subset of all variables which fit the data best according to Akaike Information Criterion (AIC) at the aggregate level. And the independent variables of the fourth model are the subset of all variables which fit the data best according to Bayesian Information Criterion (BIC) at the aggregate level.

1.5.2.2 Finite Mixture Analysis

To explore the heterogeneity in distributional preferences in the sense that there may exist distinct type of subjects that differ in their social motives, we estimate a finite mixture model (Breitmoser 2013, Bruhin, Fehr-Duda, and Epper 2010, Houser, Keane, and McCabe 2004). The finite mixture model assumes that the sample consists of C different preference types. The model consists of the estimation of the parameters of the C different types and the estimation of individual probabilities that a subject belongs to one of the C preference types. The finite mixture model's log likelihood,

$$\ln L(\Theta; X) = \sum_{i=1}^{N} \ln \sum_{c=1}^{C} \pi_c f(\theta_c; x_i),$$

weights the individual type-specific likelihood contributions $f(\theta_c; x_i)$ - here, the densities of the structural decision model with preference type parameters θ_c - by the proportions π_c of the C different types in the sample. Maximizing $\ln L(\Theta; X)$ yields the maximum likelihood estimates for the preference type parameters $\widehat{\theta_c}$ and the corresponding relative type sizes $\widehat{\pi_c}$. Once we obtain the type-specific parameters, we can calculate the posterior probability that an individual i is of type c using Bayes' rule,

$$\tau_{ic} = \frac{\widehat{\pi_c} f(\widehat{\theta_c}; x_i)}{\sum_{m=1}^{C} \widehat{\pi_m} f(\widehat{\theta_c}; x_i)}.$$

Then we classify each individual into the preference type with the highest posterior probability.

We use Normalized Entropy Criterion (NEC) (Celeux and Soromenho 1996) to determine the optimal number of types C^* by estimating mixture models with varying C.

[10] Due to the definitions of differences, there exists a linear relationship between *DiffFS-α, DiffFS-β* and *DiffEfficiency*, that is, *DiffEfficiency = 2*(*DiffFS-β – DiffFS-α*).

NEC is based on the ex-post probabilities of type membership and directly reflects the model's ability to provide a clean classification:

$$NEC(C) = \frac{E(C)}{L(C) - L(1)} \ ,$$

in which $L(C)$ is the log likelihood of the finite mixture model with C types, $L(1)$ is the log likelihood at the aggregate level, and $E(C)$ is the entropy which measures the ambiguity of the classification,

$$E(C) \ = \ - \sum_{c=1}^{C} \sum_{i=1}^{N} \tau_{ic} \ln \tau_{ic} \ .$$

The entropy is low if all τ_{ic} are either close to 1 or close to 0. And the entropy is high if many τ_{ic} are close to $1/C$, meaning that the classification of subjects into preference types is ambiguous. Thus, we determine the optimal number of types by minimizing NEC with respect to C.

1.5.2.3 Heterogeneity of Social Preferences

In order to deal with the heterogeneity and to determine individual distributional preferences, we analyze SP decisions using a finite mixture model. The results of the finite mixture analysis are robust. All four models show that the optimal number of types equals three according to NEC (Figure 1.6), and all four models create the same classification.

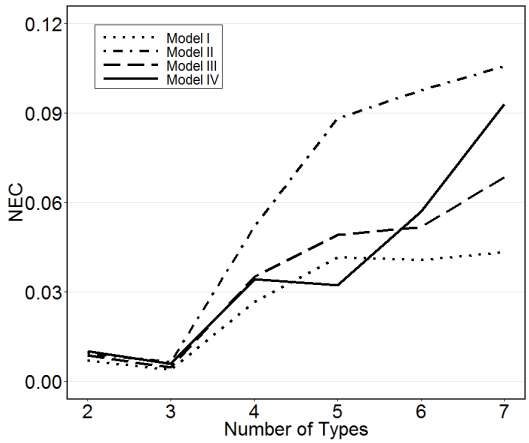


Figure 1.6 Optimal number of types

In addition, all subjects can be assigned to one distinct type with high posterior probability. Figure 1.7 shows the posterior probabilities which are larger than 0.01. A peak at probability 1 indicates that a preference type is well separated from the other

types, and no significant mass in the middle of the unit interval indicates clean classification.[11] These clean classifications suggest that our analysis is able to capture the distinctive characteristics of each preference type. We call the type based on the classification of the personal norms 'Norm Type'.

The regression results of the finite mixture analysis are shown in Table 1.7. We identify subjects' social motives according to whether the coefficient of the motive is significant or not. One exception is the coefficient of *SignEnvy* for subjects of the first norm type in Model I.[12] The relevant motives for the personal norms of each type are summarized in Table 1.8. Apart from some minor differences, all four models identify almost identical norm types. In this chapter, we base our analysis on model IV, which performs best according to the BIC criterion. However, the main results are robust with respect to the choice of the model.

[11] In Model I, all subjects can be classified into their types with the probabilities of greater than 0.93, and 93.3% of all subjects are classified into their types with the probabilities of greater than 0.99. In Model II, only one subject is classified into her type with the posterior probability of less than 0.90 (0.895), and 91.4% of all subjects are classified into their types with the posterior probabilities of greater than 0.99. In Model III, only one subject is classified into her type with the probability of less than 0.90 (0.88), and 92.4% of all subjects are classified into their types with the probabilities of greater than 0.99. In Model IV, two subjects are classified into their types with the probabilities of less than 0.90 (0.79 and 0.86), and 95.2% of all subjects are classified into their types with the probabilities of greater than 0.99.

[12] Envy is inequality aversion toward the person with the highest income. If people care about efficiency, they might like situations that envious people dislike.

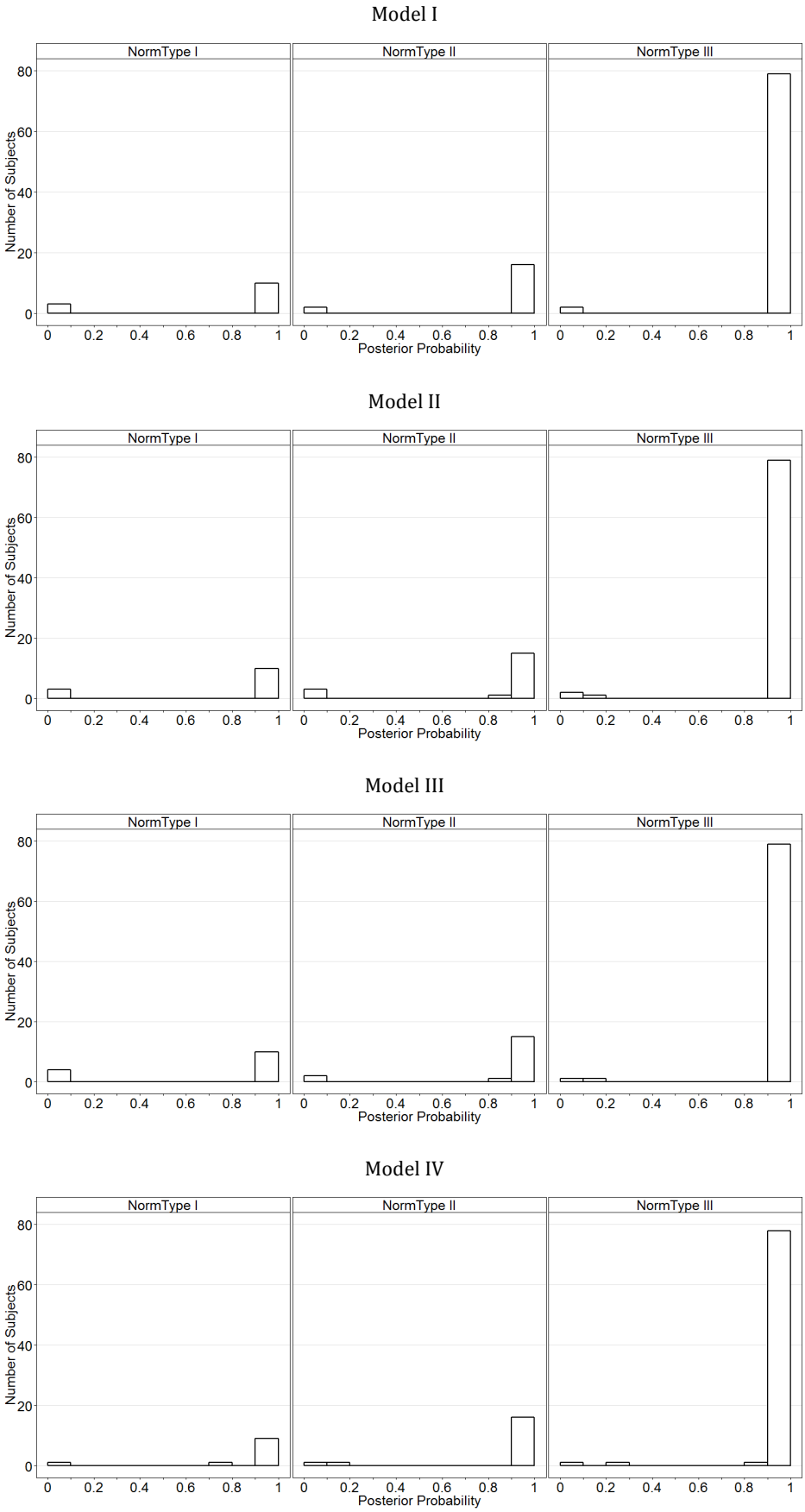


Figure 1.7 Posterior probabilities for the four models (p > 0.01)

Table 1.7 Results of the finite mixture model

	Model I			Model II			Model III			Model IV		
Num. subjects	10	16	79	10	16	79	10	16	79	10	16	79
	NT I	NT II	NT III	NT I	NT II	NT III	NT I	NT II	NT III	NT I	NT II	NT III
SignEfficiency	0.588*	-4.960	1.581***	0.542**	-3.018	2.585***	0.639***	-5.185	2.627***	0.613***	-4.612	2.395***
	(0.312)	(47.541)	(0.302)	(0.270)	(11.103)	(0.319)	(0.235)	(55.662)	(0.220)	(0.237)	(28.857)	(0.176)
DiffEfficiency	0.091	0.014	0.084									
	(0.112)	(0.107)	(0.246)									
SignEnvy	-0.658**	-0.213	0.379	-0.327	0.972***	0.435						
	(0.291)	(0.335)	(0.636)	(0.214)	(0.246)	(0.502)						
DiffEnvy	0.152	0.342**	-0.077				-0.020	0.293***	-0.039	-0.010	0.286***	-0.048
	(0.118)	(0.134)	(0.248)				(0.040)	(0.053)	(0.032)	(0.040)	(0.050)	(0.033)
SignMaximin	0.193	0.396	0.829***	0.486***	1.330***	1.435***	0.273	0.454*	0.442**			
	(0.291)	(0.301)	(0.314)	(0.186)	(0.173)	(0.211)	(0.231)	(0.252)	(0.218)			
DiffMaximin	0.065	0.287***	0.345***				0.032	0.266***	0.493***	0.073*	0.383***	0.552***
	(0.077)	(0.091)	(0.070)				(0.053)	(0.079)	(0.056)	(0.041)	(0.055)	(0.050)
SignFS-α	0.113	0.114	-0.897	-0.013	0.070	-0.353						
	(0.302)	(0.352)	0.658	(0.270)	(0.377)	(0.555)						
SignFS-β	-0.044	5.547	-0.178	-0.057	3.532	0.940***	-0.117	5.620	0.521***	-0.089	5.024	0.539***
	(0.294)	(47.540)	(0.193)	(0.218)	(11.088)	(0.108)	(0.221)	(55.661)	(0.115)	(0.222)	(29.857)	(0.119)
DiffFS-β	-0.225	-0.083	0.414									
	(0.246)	(0.239)	(0.513)									
Constant	0.195	-0.258*	-0.093	0.205	-0.136	-0.034	0.163	-0.269**	-0.032	0.136	-0.297**	-0.058
	(0.133)	(0.139)	(0.110)	(0.129)	(0.127)	(0.103)	(0.129)	(0.137)	(0.108)	(0.127)	(0.135)	(0.108)
Num. obs	320	512	2528	320	512	2528	320	512	2528	320	512	2528
AIC		1793.316			1942.759			1798.439			1801.709	
BIC		1989.146			2065.153			1920.833			1905.743	
Log Likelihood		-864.658			-951.380			-879.219			-883.854	

Notes. The dependent variable is *Decision*. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 1.8 The identified social motives for each norm type

Classification	Norm type I	Norm type II	Norm type III
Model I	Efficiency	Envy, Maximin	Efficiency, Maximin
Model II	Efficiency, Maximin	Envy, Maximin	Efficiency, Maximin, FS-β
Model III	Efficiency	Envy, Maximin	Efficiency, Maximin, FS-β
Model IV	Efficiency, Maximin	Envy, Maximin	Efficiency, Maximin, FS-β

1.5.3 Data of RT and Decisions

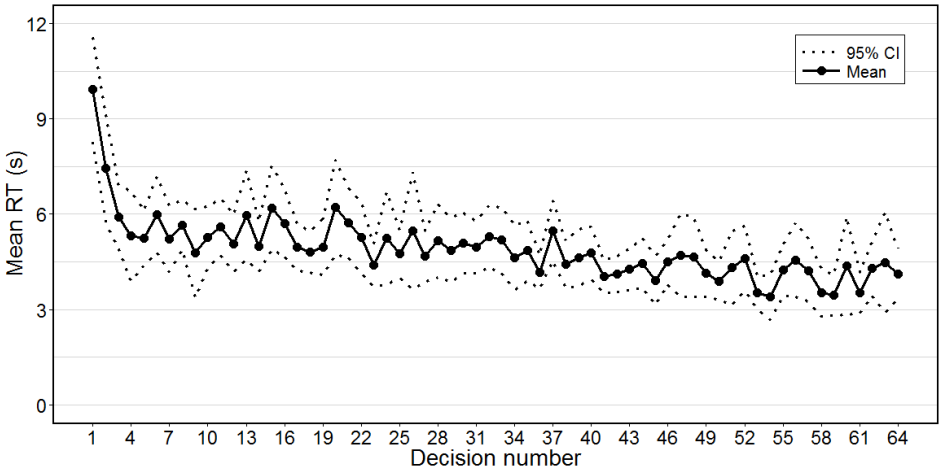


Figure 1.8 Evolution of the mean RT

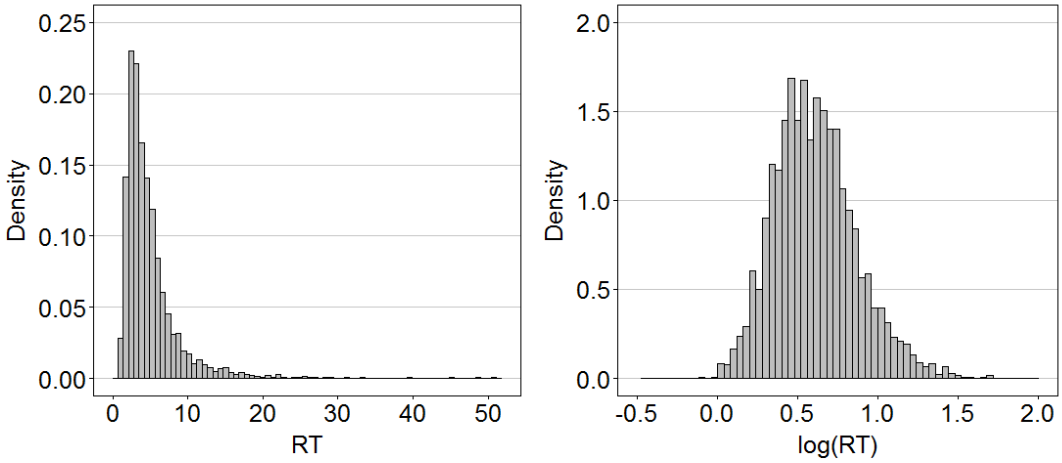


Figure 1.9 Distributions of RTs and log(RT) in SP decisions

Table 1.9 Classification of SP decisions

Norm type	Situation number	In line with					Decision type				Num. conflicts	Num. decisions
							Correct			Incorrect		
		Selfishness	Maximin	Envy	Efficiency	FS-β	Conflict		Consistent			
							Selfish	Social				
I	1	1	0		0		1				2	91
	2	1	1		0		1				2	30
	3	1	0		1		1				2	49
	4	1	1		1				1		0	88
	5	0	0		0					1		9
	6	0	1		0			1			2	14
	7	0	0		1			1			2	10
	8	0	1		1			1			2	29
II	1	1	0	0			1				2	53
	2	1	1	0			1				2	61
	3	1	0	1			1				2	24
	4	1	1	1					1		0	96
	5	0	0	0						1		15
	6	0	1	0				1			2	46
	7	0	0	1				1			2	25
	8	0	1	1				1			2	192
III	1	1	0		0	0	1				3	220
	2	1	0		1	0	1				4	287
	3	1	1		0	1	1				3	0
	4	1	1		0	0	1				4	101
	5	1	0		1	1	1				3	0
	6	1	0		0	1	1				4	0
	7	1	1		1	0	1				3	79
	8	1	1		1	1			1		0	702
	9	0	0		0	0				1		4
	10	0	0		1	0		1			3	0
	11	0	1		0	1		1			4	191
	12	0	1		0	0		1			3	1
	13	0	0		1	1		1			4	215
	14	0	0		0	1		1			3	0
	15	0	1		1	0		1			4	0
	16	0	1		1	1		1			3	728

Notes. In *Consistency* columns, 1 means the decision is consistent with the motive, 0 means the decision is not consistent with the motive. In *Decision type* columns, 1 indicates the decisions is classified into that type. *Number of conflicts* is the number of conflicts by pairwise comparisons between all the relevant social motives and selfish decisions.
For instance, for a subject of Norm type III, in situation No. 2, the decision that the subject made is in line with the selfishness and efficiency motives, but not in line with the maximin and FS-β motives. The decision is a conflict decision, since the efficiency and selfishness motives conflict with the maximin and FS-β motives when she made the decision. The number of conflicts is 4 (efficiency vs maximin, efficiency vs FS-β, selfishness vs maximin, selfishness vs FS-β).

1.5.4 Utility Calculation

Table 1.10 Logit regression for the even trials of SP decisions

	Norm type I	Norm type II	Norm type III
Constant	-0.101	0.119	0.232***
	(0.190)	(0.146)	(0.063)
Latent FMM based on TP decisions	0.577**	0.129***	0.359***
	(0.285)	(0.042)	(0.026)
DiffSelfish	0.298	0.338	0.620***
	(0.186)	(0.211)	(0.091)
SignSelfish	1.106*	-0.229	0.180
	(0.591)	(0.381)	(0.212)
AIC	166.139	350.757	1135.501
BIC	178.563	365.061	1156.018
Log Likelihood	-79.070	-171.378	-563.750
Deviance	158.139	342.757	1127.501
Num. obs.	165	264	1248

Notes. The dependent variable is *Decision*. The robust standard errors are clustered on subjects and reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 1.11 Logit regression for the odd trials of SP decisions

	Norm type I	Norm type II	Norm type III
Constant	-0.037	0.065	0.254***
	(0.172)	(0.149)	(0.077)
Latent FMM based on TP decisions	0.272	0.147***	0.397***
	(0.278)	(0.041)	(0.033)
DiffSelfish	0.205	0.509***	1.039***
	(0.223)	(0.179)	(0.107)
SignSelfish	1.191*	-0.462	-0.502**
	(0.672)	(0.393)	(0.204)
AIC	155.546	324.700	1038.934
BIC	167.720	338.754	1059.552
Log Likelihood	-73.773	-158.350	-515.467
Pseudo R^2	0.310	0.079	0.419
Num. obs.	155	248	1280

Notes. The dependent variable is *Decision*. The robust standard errors are clustered on subjects and reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 1.12 Logit regression of TP decisions

Constant	-0.095
	(0.069)
DiffEnvy	0.042***
	(0.013)
DiffMaximin	0.259***
	(0.036)
SignEfficiency	1.404***
	(0.143)
SignFS-β	0.380***
	(0.073)
AIC	2377.581
BIC	2408.179
Pseudo R^2	0.492
Log Likelihood	-1183.790
Num. obs.	3360

Notes. The dependent variable is *Decision*. The robust standard errors are clustered on subjects and reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 1.13 Logit regression of the even trials of SP decisions

Constant	0.201***
	(0.053)
Latent variable based on TP decisions	0.527***
	(0.043)
DiffSelfish	0.383***
	(0.087)
SignSelfish	0.359**
	(0.176)
AIC	1760.487
BIC	1782.186
Log Likelihood	-876.243
Deviance	1752.487
Num. obs.	1677

Notes. The dependent variable is *Decision*. The robust standard errors are clustered on subjects and reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 1.14 Logit regression of the odd trials of SP decisions

Constant	0.227***
	(0.063)
Latent variable based on TP decisions	0.598***
	(0.047)
DiffSelfish	0.614***
	(0.110)
SignSelfish	0.029
	(0.175)
AIC	1610.849
BIC	1632.563
Pseudo R^2	0.313
Log Likelihood	-801.425
Num. obs.	1683

Notes. The dependent variable is *Decision*. The robust standard errors are clustered on subjects and reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

1.5.5 Using Data of Odd Trials to Predict the RT of Even Trials

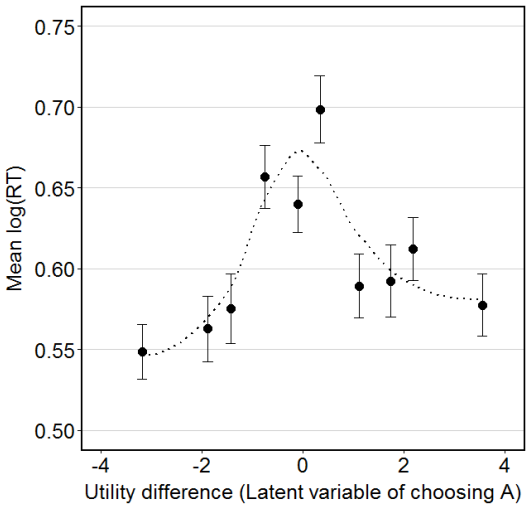


Figure 1.10 Mean log(RT) and utility difference in the even trials of SP decisions
The dotted line is the smoothing line using method "loess". The log(RT) is significantly negatively related to the absolute value of the utility difference (Pearson correlation test, r = -0.146. $p < 10^{-8}$). The log(RT) has negative relationship with the absolute value of the utility difference for 79 of 105 subjects, which is significantly different from the chance level of 50% (two-sided binomial test, $p < 10^{-6}$).

Table 1.15 Mixed-effects regressions of RTs (even trials of SP decisions)

	All SP decisions		The even trials of SP decisions	
	All subjects	Norm type III	All subjects	Norm type III
	(1)	(2)	(3)	(4)
Constant	0.659***	0.504***	0.730***	0.608***
	(0.017)	(0.038)	(0.022)	(0.065)
Conflict decision	0.076***		0.035***	
	(0.008)		(0.012)	
Decision number	-0.003***	-0.004***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Number of conflicts		0.076***		0.053***
		(0.010)		(0.016)
abs(Utility difference)			-0.036***	-0.022***
			(0.005)	(0.007)
AIC	-976.750	-427.512	-424.249	-112.853
BIC	-946.198	-399.982	-391.772	-83.972
Log Likelihood	493.375	218.756	218.124	62.427
Num. obs.	3332	1822	1661	914
Num. groups	105	79	105	79

Notes. The dependent variable is *log*(*RT*). ***p < 0.01, **p < 0.05, *p < 0.1. The coefficients of the absolute value of the utility difference are significantly negatively in regressions (3) and (4).

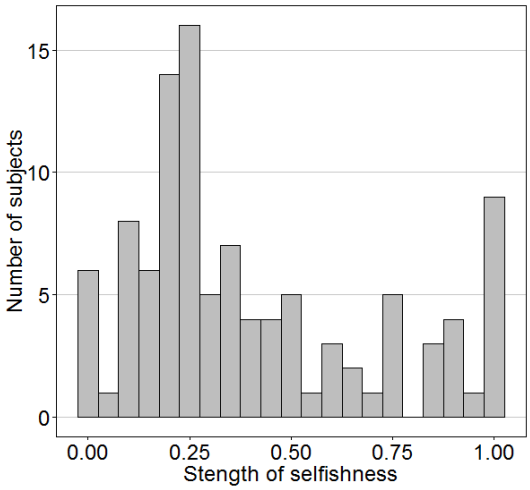


Figure 1.11 Distribution of the strength of selfishness (odd trials of conflict decisions)

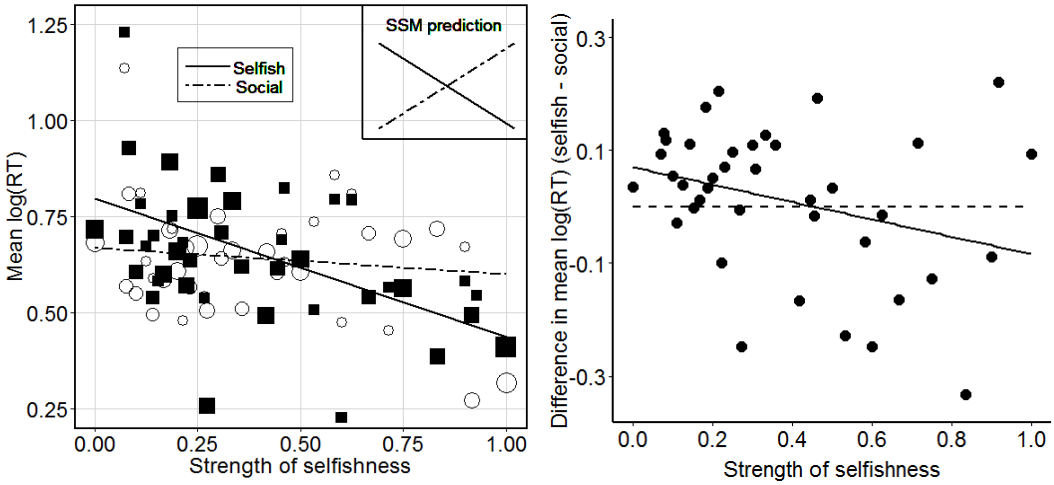


Figure 1.12 RTs of selfish and social decisions in the even trials of conflict decisions
The size of the point or the square indicates the number of subjects in that point or square. The solid and dotted lines are regression lines. The left panel displays the relationship between the RTs of selfish (social) decisions and the strength of selfishness. Subjects who are more selfish are quicker in making selfish decisions (Pearson's correlation test, r = -0.441, $p < 10^{-15}$). But subjects who are more selfish are not significantly slower in making social decisions (Pearson's correlation test, r = -0.048, p = 0.195). The right panel shows that the time difference between selfish and social decisions is negatively related with the strength of selfishness (Pearson's correlation test, r = -0.285, $p < 0.007$).

Table 1.16 Regressions of RTs in the even trials of conflict decisions

	Mixed-effects regression	OLS regressions	
	(1)	(2)	(3)
Constant	0.737***	0.826***	0.790***
	(0.022)	(0.036)	(0.037)
Selfish decision	0.011	0.023	0.122***
	(0.016)	(0.028)	(0.043)
Strength of selfishness		-0.287***	-0.090
		(0.048)	(0.079)
abs(Utility difference)	-0.028***	-0.023***	-0.029***
	(0.006)	(0.008)	(0.008)
Conflict within norms	0.049***	0.053***	0.042***
	(0.013)	(0.014)	(0.015)
Male		0.044	0.025
		(0.035)	(0.034)
Decision number	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)
Selfish decision × Male		-0.013	0.025
		(0.038)	(0.036)
Selfish decision × Strength of selfishness			-0.297***
			(0.085)
Num. obs.	1239	1239	1239
Num. groups	105	105	105

Notes. The dependent variable is *log*(*RT*). *Selfish Decision* is a dummy variable which indicate the decision is a selfish decision or social decision. *Male* is a dummy variable which indicate the gender. The robust standard errors for regressions (2) and (3) are clustered on subjects and reported in parentheses. Regression (1) and (2) show that the RTs of selfish decisions are not significantly different from the RTs of social decisions when we control the utility difference between choice options and the conflicts between norms. Regression (3) shows that the RTs of selfish decisions decrease with the strength of selfishness, but the strength of selfishness has no significant effect on the RTs of social decisions. Both regressions (2) and (3) show that there is no significant gender difference on the RTs of selfish and social decisions.
***p < 0.01, **p < 0.05, *p < 0.1.

Table 1.17 OLS regressions of RT on the utility difference (even trials of SP decisions)

	(1)	(2)	(3)
Constant	0.660***	0.664***	0.670***
	(0.020)	(0.019)	(0.019)
abs(Utility difference based on FMM)	-0.034***		-0.023*
	(0.008)		(0.014)
abs(Utility difference based on Logit)		-0.041***	-0.021
		(0.008)	(0.014)
R^2	0.021	0.019	0.024
Adj. R^2	0.021	0.019	0.023
Num. obs.	1677	1677	1677

Notes. The dependent variable is log(RT). The robust standard errors are clustered on subjects and reported in parentheses. Regressions (1) and (2) shows the utility difference based FMM has a higher explanatory power (adjusted R^2 = 0.021) than the utility difference based on the Logit model (adjusted R^2 = 0.019). Regression (3) shows that the coefficient of the utility difference based on FFM is more robust than the coefficient of the utility difference based on the Logit model, since the coefficient of the utility difference based on the Logit model is not significant at all if we simultaneously include the two variables in the regression.
***p < 0.01, **p < 0.05, *p < 0.1.

1.5.6 Analysis of Eye-Tracking Data from Fiedler et al, 2013

In the paper of Fiedler et al. (2013), subjects make decision for both Social Value Orientation (SVO) Ring Measure Task and SVO Slider Task. To do an out-of-sample analysis, we access subjects' prosociality using their decisions for SVO Slider Task, and analyze their fixation behavior in the SVO Ring Measure Task.

Using decisions for SVO Slider Task, all subjects can be classified into three types: selfish, prosocial and competitive types. Since there is only one subject who is competitive type. In the following, we focus our analysis on selfish and prosocial subjects. Figure 1.13 displays the proportion of the fixation on own payoffs and how it changes with time. We divide each decision (process) into five bins with equal time period. It shows that the selfish subjects put more attention (80.1%) on their own payoffs at the beginning of the decision process compare to prosocial subjects (60.1%). The proportion of the fixation on own payoffs decreases with time for both selfish and prosocial subjects, and becomes stable at the end of the decision process.

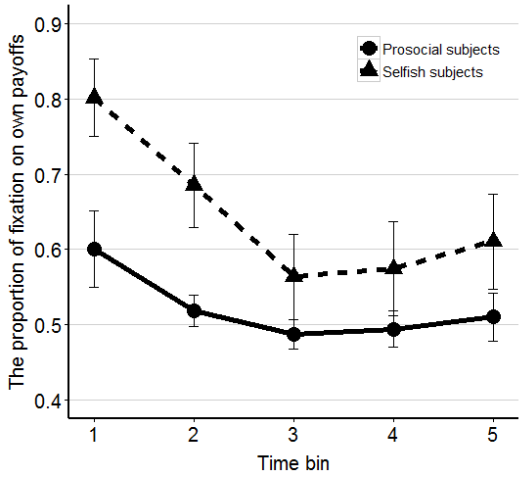


Figure 1.13 Proportion of the fixation on own payoffs

Figure 1.14 displays the proportion of subjects who fixate on their own payoffs in each fixation. It shows that 78.6% of selfish subjects who fixate on their own payoffs in the first fixation, while 58.5% of prosocial subjects who fixate on their own payoffs in the first fixation. The proportion of subjects who fixate on own payoffs becomes stable and around 50% from the fifth fixation for both selfish and prosocial types.

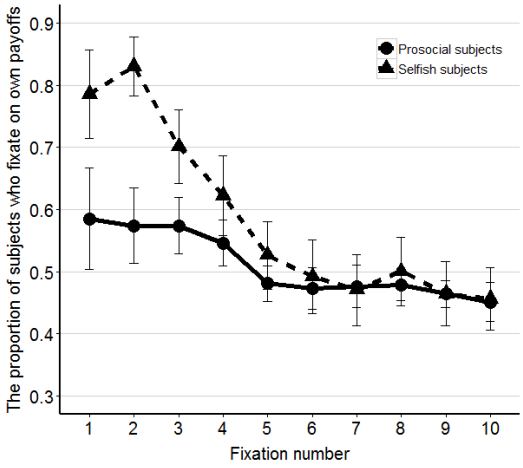


Figure 1.14 Proportion of subjects who fixate on own payoffs in each fixation

1.6 Instruction

General Instructions

Today you are participating in an economic experiment. If you read the following instructions carefully, you can – depending on your decisions – earn money in addition to the show-up fee of 3 Euros. Therefore, it is important that you read these instructions carefully.

During the whole experiment, it is not allowed to communicate with other participants. We, therefore, ask you to turn off the cell phone and not to speak with each other. If you do not understand something, please consult the instructions again. If you still have questions, please raise your hand. We will come to you and answer your questions individually.

In this experiment, you will need to decide for different situations. At the end, one of the situations will be randomly drawn and paid out. You will receive your payment in accordance with the decisions in this relevant situation.

In the instructions we do not speak of Euro, but points. The points you earn during the experiment will be converted into Euros in the following rate:

1 Point = 50 Cents

That is, you get 50 cents per point in the relevant situation. Of course, you will also receive a show up fee of 3 Euros.

On the following pages we will explain the exact course of the experiment. First, we will familiarize you with the decision situation. When you finish reading the instructions, on your screen you will find control questions which will help you to understand the situations. The experiment only begins when all participants are completely familiar with the course of the experiment.

The Experiment

All the participants in the laboratory are randomly divided into groups of three. Each group consists of Participant I, Participant II and Participant III. In each situation, two point distributions which relate to the three members of the group are available. Participant II can decide which of the two distributions is selected. Since only Participant II makes decisions, in the following we explain the experiment from the perspective of Participant II (Important: Each participant can be Participant II). In the experiment, each participant makes decisions as Participant II. Which person is Participant I, II or III in the group will be randomly drawn at the end of the experiment. In addition, one of the decisions of Participant II will be randomly drawn to be implemented at the end of the experiment.

Display on the Screen

Figure 1.15 Keys for making decisions

This experiment consists of a series of 64 decision situations in which you can choose one of two point distributions as Participant II. The following screenshot shows an example. In the left option, Participant I receives 9 points, you, as Participant II, receive 12 points and Participant III receives 16 points. The height of bars on the left corresponds to the corresponding amounts. “Your” bar is always shown in the middle and in white color. In the right option, Participant I receives 10 points, you, as Participant II, receive 17 points and Participant III receives 19 points. The height of the bars on the right also

corresponds to these amounts. You make your decisions with the help of the keyboard. For the left option, you press the key “F” and for the right option you press the key “J” (see Figure 1.15). Which key to press is also displayed at the bottom of the screen. Therefore, in this example, if you press “F”, you receive 12 points, Participant I receives 9 points and Participant III receives 16 points. If you press “J”, you receive 17 points, Participant I receives 10 points and Participant III receives 19 points. After each decision you have to press the ‘Spacebar’ to continue.

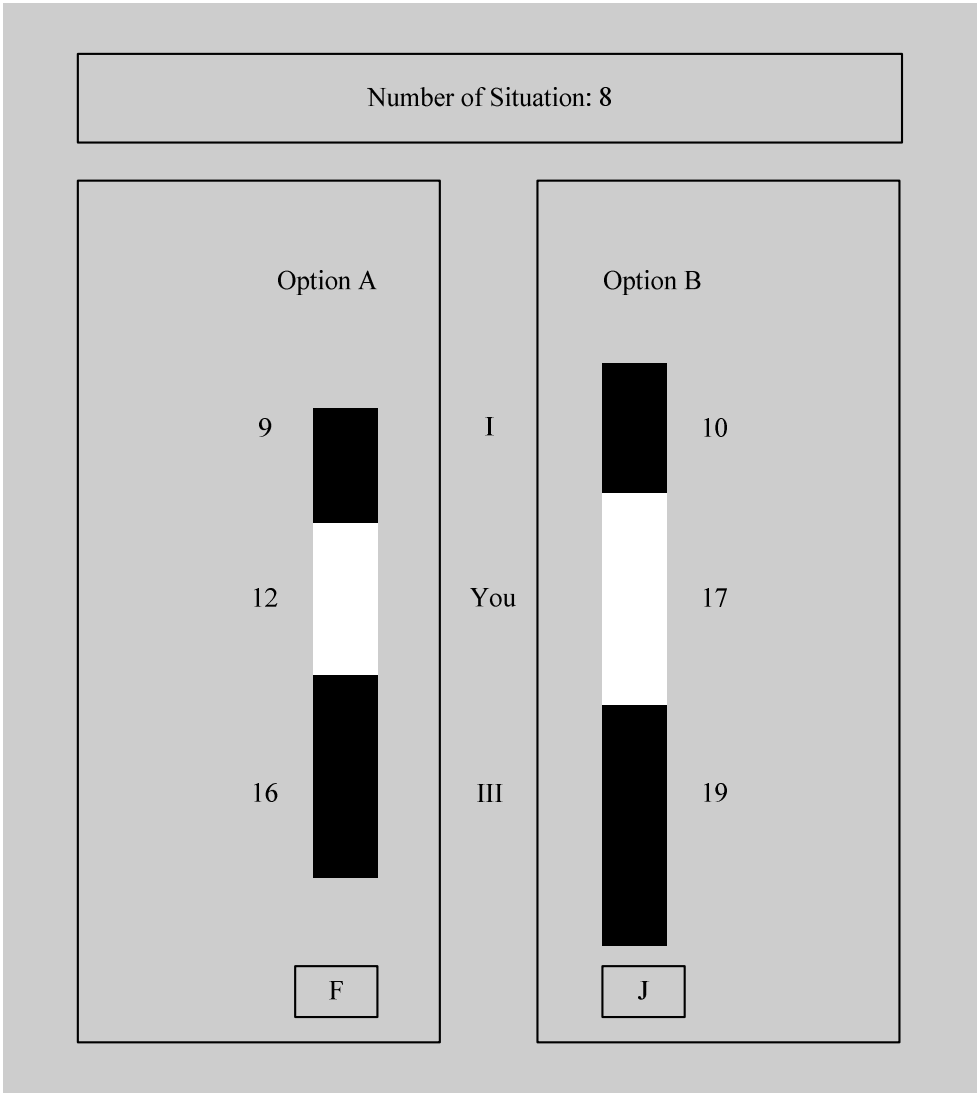


Figure 1.16 Screen layout

Payment

At the end of the experiment, it will be randomly drawn which one of the 64 situations will be paid and who is Participant I, II and III. The draw will be made with a

die by the participant at the 13th place. Then it takes about one minute to display all your decision situations and your income in the experiment.

If you have understood the instructions, please answer the control questions on the screen.

1.7 References of Chapter 1

Achtziger, Anja, and Carlos Alós-Ferrer. 2014. "Fast or Rational? A Response-Times Study of Bayesian Updating." *Management Science* 60 (4):923-38.

Alós-Ferrer, Carlos, and Fritz Strack. 2014. "From Dual Processes to Multiple Selves: Implications for Economic Behavior." *Journal of Economic Psychology* 41:1-11.

Andreoni, James, and John Miller. 2002. "Giving According to Garp: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica* 70 (2):737-53.

Bear, Adam, and David G. Rand. 2016. "Intuition, Deliberation, and the Evolution of Cooperation." *Proceedings of the National Academy of Sciences* 113 (4):936-41.

Bolton, Gary E., and Axel Ockenfels. 2000. "Erc: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90 (1):166-93.

Breitmoser, Yves. 2013. "Estimation of Social Preferences in Generalized Dictator Games." *Economics Letters* 121 (2):192-7.

Brocas, Isabelle, and Juan D. Carrillo. 2014. "Dual-Process Theories of Decision-Making: A Selective Survey." *Journal of Economic Psychology* 41:45-54.

Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper. 2010. "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion." *Econometrica* 78 (4):1375-412.

Cappelen, Alexander W., Ulrik H. Nielsen, Bertil Tungodden, Jean-Robert Tyran, and Erik Wengström. 2015. "Fairness Is Intuitive." *Experimental Economics*:1-14.

Cappelletti, Dominique, Werner Güth, and Matteo Ploner. 2011. "Being of Two Minds: Ultimatum Offers under Cognitive Constraints." *Journal of Economic Psychology* 32 (6):940-50.

Cary, Frydman, and Gideon Nave. 2015. "Extrapolative Beliefs in Perceptual and Economic Decisions: Evidence of a Common Mechanism." *Working Papaer*.

Celeux, Gilles, and Gilda Soromenho. 1996. "An Entropy Criterion for Assessing the Number of Clusters in a Mixture Model." *Journal of Classification* 13 (2):195-212.

Chaiken, Shelly, and Yaacov Trope. 1999. *Dual-Process Theories in Social Psychology*: Guilford Press.

Charness, Gary, and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117 (3):817-69.

Cone, Jeremy, and David G. Rand. 2015. "Time Pressure Increases Cooperation in Competitively Framed Social Dilemmas." *PLoS ONE* 9 (12):e115756.

Cornelissen, Gert, Siegfried Dewitte, and Luk Warlop. 2011. "Are Social Value Orientations Expressed Automatically? Decision Making in the Dictator Game." *Personality and Social Psychology Bulletin* 37 (8):1080-90.

Dickhaut, John, Vernon Smith, Baohua Xin, and Aldo Rustichini. 2013. "Human Economic Choice as Costly Information Processing." *Journal of Economic Behavior & Organization* 94:206-21.

Duffy, Sean, and John Smith. 2014. "Cognitive Load in the Multi-Player Prisoner's Dilemma Game: Are There Brains in Games." *Journal of Behavioral and Experimental Economics* 51:47-56.

Dufwenberg, Martin, and Georg Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47 (2):268-98.

Engelmann, Dirk, and Martin Strobel. 2004. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments." *American Economic Review* 94 (4):857-69.

Erlei, Mathias. 2008. "Heterogeneous Social Preferences." *Journal of Economic Behavior & Organization* 65 (3–4):436-57.

Evans, Anthony M, Kyle D Dillon, and David G Rand. 2015. "Fast but Not Intuitive, Slow but Not Reflective: Decision Conflict Drives Reaction Times in Social Dilemmas." *Journal of Experimental Psychology: General* 144 (5):951-66.

Falk, Armin, and Urs Fischbacher. 2006. "A Theory of Reciprocity." *Games and Economic Behavior* 54 (2):293-315.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3):817-68.

Fiedler, Susann, Andreas Glöckner, Andreas Nicklisch, and Stephan Dickert. 2013. "Social Value Orientation and Information Search in Social Dilemmas: An Eye-Tracking Analysis." *Organizational Behavior and Human Decision Processes* 120 (2):272-84.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2):171-8.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. "Individual Preferences for Giving." *American Economic Review* 97 (5):1858-76.

Frederick, Shane. 2005. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives* 19 (4):25-42.

Fudenberg, Drew, and David K. Levine. 2006. "A Dual-Self Model of Impulse Control." *American Economic Review* 96 (5):1449-76.

Greiner, Ben. 2015. "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." *Journal of the Economic Science Association* 1 (1):114-25.

Grun, Bettina, and Friedrich Leisch. 2008. "Flexmix Version 2: Finite Mixtures with Concomitant Variables and Varying and Constant Parameters." *Journal of Statistical Software* 28 (4):1-35.

Hauge, Karen Evelyn, Kjell Arne Brekke, Lars-Olof Johansson, Olof Johansson-Stenman, and Henrik Svedsäter. 2016. "Keeping Others in Our Mind or in Our Heart? Distribution Games under Cognitive Load." *Experimental Economics* 19 (3):562-76.

Houser, Daniel, Michael Keane, and Kevin McCabe. 2004. "Behavior in a Dynamic Decision Problem: An Analysis of Experimental Evidence Using a Bayesian Type Classification Algorithm." *Econometrica* 72 (3):781-822.

Hutcherson, Cendri A., Benjamin Bushong, and Antonio Rangel. 2015. "A Neurocomputational Model of Altruistic Choice and Its Implications." *Neuron* 87 (2):451-62.

Kahneman, Daniel. 2003. "A Perspective on Judgment and Choice: Mapping Bounded Rationality." *American Psychologist* 58 (9):697-720.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*: Macmillan.

Kerschbamer, Rudolf. 2015. "The Geometry of Distributional Preferences and a Non-Parametric Identification Approach: The Equality Equivalence Test." *European Economic Review* 76:85-103.

Krajbich, Ian, Carrie Armel, and Antonio Rangel. 2010. "Visual Fixations and the Computation and Comparison of Value in Simple Choice." *Nature Neuroscience* 13 (10):1292-8.

Krajbich, Ian, Bjorn Bartling, Todd Hare, and Ernst Fehr. 2015. "Rethinking Fast and Slow Based on a Critique of Reaction-Time Reverse Inference." *Nature Communication* 6:7455.

Krajbich, Ian, Todd Hare, Bjoern Bartling, Yosuke Morishima, and Ernst Fehr. 2015. "A Common Mechanism Underlying Food Choice and Social Decisions." *PLoS Comput Biol* 11 (10):e1004371.

Krajbich, Ian, Bastiaan Oud, and Ernst Fehr. 2014. "Benefits of Neuroeconomic Modeling: New Policy Interventions and Predictors of Preference." *American Economic Review* 104 (5):501-6.

Lohse, Johannes, Timo Goeschl, and Johannes Diederich. 2014. "Giving Is a Question of Time: Response Times and Contributions to a Real World Public Good." *University of Heidelberg Department of Economics Discussion Paper Series*, (566).

Lotito, Gianna, Matteo Migheli, and Guido Ortona. 2013. "Is Cooperation Instinctive? Evidence from the Response Times in a Public Goods Game." *Journal of Bioeconomics* 15 (2):123-33.

Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf. 2011. "Measuring Social Value Orientation." *Judgment & Decision Making* 6 (8):771-81.

Nielsen, Ulrik H., Jean-Robert Tyran, and Erik Wengström. 2014. "Second Thoughts on Free Riding." *Economics Letters* 122 (2):136-9.

Peysakhovich, Alexander, and David G. Rand. forthcoming. "Habits of Virtue: Creating Norms of Cooperation and Defection in the Laboratory." *Management Science*.

Piovesan, Marco, and Erik Wengström. 2009. "Fast or Fair? A Study of Response Times." *Economics Letters* 105 (2):193-6.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (5):1281-302.

Rand, David G, Victoria Brescoll, Jim AC Everett, Valerio Capraro, and Helene Barcelo. 2016. "Social Heuristics and Social Roles: Intuition Favors Altruism for Women but Not for Men." *Forthcoming in Journal of Experimental Psychology: General*.

Rand, David G., Joshua D. Greene, and Martin A. Nowak. 2012. "Spontaneous Giving and Calculated Greed." *Nature* 489 (7416):427-30.

Ratcliff, Roger. 1978. "A Theory of Memory Retrieval." *Psychological Review* 85 (2):59-108.

Ratcliff, Roger, and Philip L. Smith. 2004. "A Comparison of Sequential Sampling Models for Two-Choice Reaction Time." *Psychological Review* 111 (2):333-67.

Rubinstein, Ariel. 2007. "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal* 117 (523):1243-59.

Schulz, Jonathan F., Urs Fischbacher, Christian Thöni, and Verena Utikal. 2014. "Affect and Fairness: Dictator Games under Cognitive Load." *Journal of Economic Psychology* 41:77-87.

Sloman, Steven A. 1996. "The Empirical Case for Two Systems of Reasoning." *Psychological Bulletin* 119 (1):3-22.

Strack, Fritz, and Roland Deutsch. 2004. "Reflective and Impulsive Determinants of Social Behavior." *Personality and Social Psychology Review* 8 (3):220-47.

Tinghog, Gustav, David Andersson, Caroline Bonn, Harald Bottiger, Camilla Josephson, Gustaf Lundgren, Daniel Vastfjall, Michael Kirchler, and Magnus Johannesson. 2013. "Intuition and Cooperation Reconsidered." *Nature* 498 (7452):E1-E2.

Verkoeijen, Peter P. J. L., and Samantha Bouwmeester. 2014. "Does Intuition Cause Cooperation?" *PLoS ONE* 9 (5):e96654.

2 Sequential Sampling with Intuition

2.1 Introduction

Traditionally, economists have focused on "black box" models of decision making, neglecting the underlying process which leads to decisions. In recent years, research has increasingly focused on the cognitive process underlying human behavior since process data, such as response times (RTs), allow for a deeper understanding of decision making (Chen and Fischbacher 2016, Clithero 2016b, Frydman and Krajbich 2016, Fudenberg, Strack, and Strzalecki 2015, Konovalov and Krajbich 2015, Krajbich, Oud, and Fehr 2014, Rubinstein 2007, 2016, Spiliopoulos and Ortmann 2016, Woodford 2014). However, results of studies using RT to investigate the cognitive process of social decision making are still disputed and inconclusive.

One debate is whether people have a spontaneous tendency to be prosocial or selfish. This line of studies investigates the cognitive process of social decision making using a dual-process framework (Alós-Ferrer and Strack 2014, Brocas and Carrillo 2014, Evans 2008, Frederick 2005, Kahneman 2011). Some studies find that people's intuition favors prosocial behavior (Cappelen et al. 2015, Cappelletti, Güth, and Ploner 2011, Lotito, Migheli, and Ortona 2013, Nielsen, Tyran, and Wengström 2014, Rand, Greene, and Nowak 2012, Schulz et al. 2014), while other studies find that people's intuition favors the egoistic choice (Duffy and Smith 2014, Fiedler et al. 2013, Hauge et al. 2016, Lohse, Goeschl, and Diederich forthcoming, Piovesan and Wengström 2009, Tinghog et al. 2013).

Another line of research has argued that the cognitive process of social decision making is better explained by sequential sampling models (SSMs), which assume that information is sampled continuously until sufficient evidence is accumulated for one of the available options (Ratcliff 1978, Ratcliff and McKoon 2008). In particular, these papers have argued that variations in RT are due to the strength of preferences, not to parallel competing processes (Hutcherson, Bushong, and Rangel 2015, Krajbich, Bartling, et al. 2015a, Krajbich, Hare, et al. 2015, Krajbich, Oud, and Fehr 2014).[13]

The purpose of this chapter is to (1) resolve these debates and reconcile the conflicting results. In particular, we seek to understand why time pressure favors prosocial behavior in some cases and selfishness in others; (2) propose a model to describe the cognitive process underlying social decision making by integrating the notion of intuition into a SSM framework. To do so, we use a series of mini-dictator games in which there is always a conflict between selfishness and prosociality. Subjects make

[13] The strength of preferences is usually measured by the value difference or the utility difference between choice options in value-based or economic decision making.

binary decisions for these games under time-free, time-pressure and time-delay conditions in the experiment. The experiment allows us to identify whether subjects have an early bias to be prosocial or selfish by comparing how their preferences change across different time conditions. The experiment also allows us to test whether the cognitive process of social decision making complies with SSMs, even under time pressure and time delay. Going a step further, we integrate the early bias (or 'intuition') into the starting point of the SSM process, and propose a SSM with intuition (SSMi) to model the cognitive process. Then we validate the model with out-of-sample predictions.

The experimental results show that subjects are heterogeneous in whether their intuition favors prosociality or selfishness. Selfish subjects become more selfish under time pressure and become more prosocial under time delay compared to the time-free condition, while prosocial subjects become more prosocial under time pressure and more selfish under time delay. This indicates that the intuition favors the selfish option for selfish subjects, and the intuition favors the prosocial option for prosocial subjects.

The experiment also provides strong evidence that social decision making is in line with SSMs. That is, RTs decrease with strength of preference in all three conditions. The out-of-sample prediction results reveal that the SSMi improves the predictive power compared to the standard SSM. More interestingly, the SSMi outperforms standard logistic choice models, similar to other recent results in this domain (Clithero 2016a, Krajbich, Hare, et al. 2015). We also propose and test alternative ways to capture intuition with SSMs, comparing biases in drift rates to biases in starting points. The results show that intuition is better explained by the starting point.

Our study is related to the studies mentioned above which use RT to investigate whether people are intuitive in selfishness or prosociality. Different from these studies which use the relative RT to distinguish the intuitive and the deliberative choices, our experiment provides causal evidence about the heterogeneity in whether the intuition favors prosociality or selfishness using time manipulation. This study is also related to the studies mentioned above which investigate the cognitive process of social decision making using SSMs. However, these studies do not take intuition into account. To the best of our knowledge, our study is the first study to integrate the intuition of selfishness or prosociality into SSMs to model the cognitive process of social decision making.

In addition to the papers mentioned above, our study is highly related to theoretical studies which combine dual-process theories and SSMs (Alós-Ferrer forthcoming, Caplin and Martin 2016). Caplin and Martin (2016) assume that decisions are the interaction results of two processes, one is the automatic process as that in dual-process theories,

and the other is the controlled process which complies with the standard drift-diffusion model. Alós-Ferrer (forthcoming) assumes decisions arise from the interplay of the utility process and the heuristic process, and both processes comply with SSMs. Our model departs from both of these models by instead assuming a single decision process that incorporates intuition as a prior (i.e. starting point) in that process.

This chapter is organized as follows. Section 2 presents the experimental design and procedures. Section 3 presents the experimental evidence about the cognitive process of social decision making. Section 4 describes the SSMi and reports the model estimation. Section 5 validates the model with out-of-sample predictions. Section 6 tests alternative versions of the model. Section 7 concludes.

2.2 Experimental Design and Procedures

2.2.1 Experimental Design

The experiment consists of 200 mini-dictator games, in which the subject has to make a binary decision about how money is distributed between herself (dictator) and the other participant (receiver) in her group. In each game, there is always a conflict between selfishness and advantageous inequality aversion (Fehr and Schmidt 1999). Figure 2.1 shows an example screenshot of a mini-dictator game in the experiment. In this game, the subject can either choose the selfish option (Option B) which has a higher payoff for herself (95 vs. 85) or choose the prosocial option (Option A) which has a higher payoff for the other participant in her group (21 vs. 18).

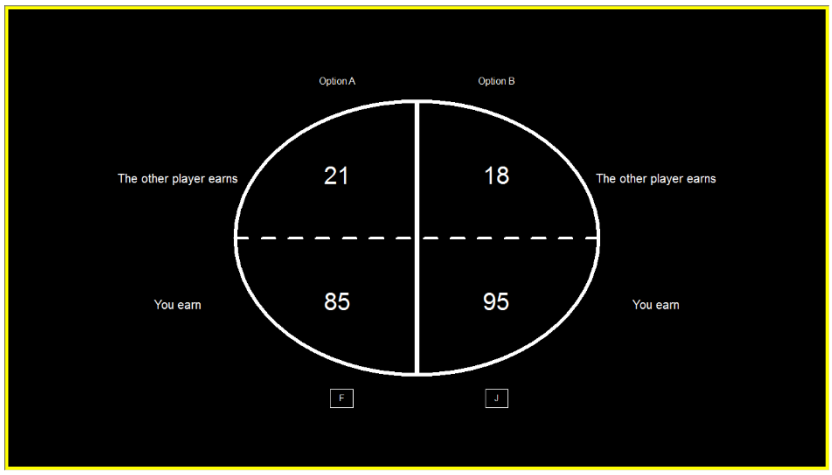


Figure 2.1 A screenshot of a mini-dictator game in the experiment

We divide the whole experiment into four parts with 50 games in each. In one part, subjects have to make each of their decisions within 2 seconds (time-pressure condition). In another part, subjects have to make each of their decisions after the decision situation has been displayed for 10 seconds (time-delay condition). In the other two parts, subjects can take as long as they want for each decision (time-free condition). The first part and the fourth part of the experiment are time-free conditions. The second and the third parts are time-pressure and time-delay conditions, but the order of these two conditions is counterbalanced across subjects.

How to enforce time constraints in these experiments is an important problem. We choose to allow subjects to respond at any time, to avoid missing observations. However, if the time-constraint is disobeyed, we impose a probabilistic penalty such that a dictator's decision would only be implemented with a probability of 10%. In the other 90% of cases, the dictator earns 0 and the receiver earns the lower of the two amounts that she could earn in the current game. We choose this scheme relying on IIA (independence and irrelevant axiom) to ensure that subjects would not choose a different option outside of the time constraints, and trying to minimize the chance that subjects would prefer the penalized outcome (i.e. worst possible outcome for self, worst available outcome for the other). Subjects are notified with a warning message if they were too slow under time pressure or too fast under time delay.

The games under different time conditions have the same properties but minor differences in payoffs. Specifically, the differences between the dictators' payoffs in the two options are 2, 4, 6, 8, and 10, and the differences between the receivers' payoffs are from 3 to 57, in step of 6. We first fix the parameters for 50 games under time-free condition (Games 1-50). Then we decrease (increase) all the parameters by 1 for one half of these games, and increase (decrease) all the parameters by 1 for the other half of games to get the 50 games for time-pressure (-delay) condition. Finally, we decrease all the parameters by 2 for one half of the games and increase all the parameters by 2 for the other half of games to get the other 50 time-free trials (Games 51-100). Therefore, games under different time conditions have the same difference between dictator and receiver's payoffs.

2.2.2 Procedures

We programed the experiment using z-Tree (Fischbacher 2007) and recruited subjects using ORSEE (Greiner 2015). In total we had 102 subjects. 18 subjects took part

in an initial experiment at the Ohio State University in March 2016, followed by 84 subjects at the University of Konstanz from April to August 2016.

At the beginning of each session, we randomly matched subjects into two-person groups. To make sure they understand the decision situation, we provided a set of control questions at the beginning of the first three parts. We randomized the order of the games under different time conditions for each group. Subjects made decisions by pressing the "F" or "J" keys to select the left or right option, respectively. They saw a waiting screen after each decision and were required to press "Spacebar" to advance to the next trial. At the end of the experiment, we randomly selected one trial for each group and paid them according to one of their decisions (randomly selected). On average, subjects earned 20 dollars at the Ohio State University and 16 Euros at the University of Konstanz (including show-up fees).

2.3 Cognitive Processes underlying Decision Making

We first report the behavioral and RT data in the experiment. Subjects choose the selfish option in 42.3% of all games in the time-free condition, they choose the selfish option in 50.7% of all games in the time-pressure condition, and they choose the selfish option in 40.3% of all games in the time-delay condition.[14] The median RTs for decisions in the time-free, time-pressure, and time-delay conditions are 2.156 s ($sd = 2.797$), 0.937 s ($sd = 0.393$), and 11.470 s ($sd = 1.642$), respectively.[15] Subjects violate the time constraints in 1.6% of time-pressure trials and 1.0% of time-delay trials.

2.3.1 Heterogeneity in Intuition towards Selfishness or Prosociality

Dual-process theories (models) have been widely used to describe the process underlying human behavior in a number of domains (Evans 2008, Weber and Johnson 2009). These theories share the fundamental assumption that decisions can result from two types of processes, which have been labelled differently by different authors.[16] One is the intuitive process, which is fast, effortless, and requires no conscious control. The

[14] The distributions of the frequency of selfish choice under different time conditions are shown in Figure 2.10 in Appendix.
[15] The distributions of RTs under different time conditions are shown in Figure 2.11 in Appendix.
[16] The two types of processes have also been labelled as "automatic vs. controlled" (Schneider and Shiffrin 1977), "impulsive vs. reflective" (Strack and Deutsch 2004), "experiential vs. rational" (Epstein 1994), and "System I vs. System II" (Kahneman and Frederick 2002).

other is the deliberative process, which is slow, controlled, and consumes cognitive resources.

Here we study whether people might be intuitively prosocial or selfish. There is still a large debate on this topic and findings favor both hypotheses. We conjecture that if people have such intuitions, they are likely heterogeneous. That is, some people are intuitively prosocial, while others are intuitively selfish. We test this hypothesis by comparing how subjects' preferences change across different time conditions in our experiment.

According to these dual-process theories, decisions under time pressure should rely more on subjects' intuition and decisions under time delay should rely more on deliberation. Therefore, subjects should become more prosocial in the time-pressure (-free) condition compared to the time-free (-delay) condition if they are intuitively prosocial. On the contrary, subjects should become more selfish in the time-pressure (-free) condition compared to the time-free (-delay) condition if they are intuitively selfish.

Consistent with prior work in this literature, the behavioral results show that our subjects generally do care about both players' payoffs (e.g. Charness and Rabin 2002, Fehr and Schmidt 1999). We employ the inequality aversion model proposed by Fehr and Schmidt (1999) to model subjects' behavior and estimate their preferences. Since the dictator's payoff is always higher than the receiver's payoff in each option, we do not have to consider disadvantageous inequality aversion in the model. Thus, a subject's utility for each option in the mini-dictator game is given by

$$\mathrm{U}(u_d, u_r) = u_d - \beta(u_d - u_r) \tag{1}$$

in which u_d is the dictator's payoff and u_r is the receiver's payoff. The parameter β indicates the subject's social preference, with higher β indicating more prosociality.[17] Using maximum likelihood estimation (MLE), we estimate β for each subject under time-free (β_f), time-pressure (β_p), and time-delay (β_d) conditions separately (we estimate β_f using only Games 1-50).[18]

We find that the effect of decision time differs substantially across subjects. Subjects with higher β_f (prosocial subjects) generally become more prosocial under time

[17] $\beta = 0$ indicates that the subject is completely selfish and only cares about her own payoff, $\beta > 0$ indicates that the subject is inequality averse, and $\beta < 0$ indicates that the subject is inequality seeking.

[18] We get similar results if we estimate β_f using decisions for Games 51-100 or using decisions for Games 1-100. We estimate β_f using Games 1-50 here since we will estimate our model using decisions for Games 1-50 to do out-of-sample predictions for Games 51-100.

pressure, while subjects with lower β_f (selfish subjects) become more selfish under time pressure. Similarly, prosocial subjects generally become less prosocial under time delay, while selfish subjects become less selfish under time delay, though these effects are less pronounced, as shown in Figure 2.2.

To statistically test these effects, we split subjects into two types according to the median β_f. Wilcoxon signed rank tests show that $\beta_p > \beta_f$ (p = 0.025), $\beta_f > \beta_d$ (p = 0.027), and $\beta_p > \beta_d$ (p = 0.001) for subjects with higher β_f, and $\beta_p < \beta_f$ (p = 0.111), $\beta_f < \beta_d$ (p = 0.001), and $\beta_p < \beta_d$ (p < 0.001) for subjects with lower β_f. In other words, subjects are heterogeneous in their tendency to choose the selfish or prosocial options, time pressure magnifies those tendencies, while time delay reduces them.

2.3.2 Sequential Sampling Model of Social Decision Making

Prior studies have shown that behavior in social decision making is in line with SSM predictions (Hutcherson, Bushong, and Rangel 2015, Krajbich, Bartling, et al. 2015a, Krajbich, Hare, et al. 2015, Krajbich, Oud, and Fehr 2014). Specifically, RT decreases with strength of preference when people make social decisions. Our experiment not only allows us to test this hypothesis, but also allows us to test whether this hypothesis still holds under time pressure and time delay. In particular, if decisions under time pressure exclusively (or preferentially) rely on an intuitive process, then we might expect no (or a greatly reduced) relationship between RT and strength of preference.

To test this, we calculate the utility difference between the two options in the mini-dictator games using model (1) and the advantageous inequality aversion parameters $(\beta_f, \beta_p, \beta_d)$ estimated before.[19] We standardize the utility differences to $[-1, 1]$ at the individual level, and divide all the data under each time condition into 11 bins with equal size.

Figure 2.3 plots the correlations between RT and the utility difference under different time conditions. Obviously, RT decreases with strength of preference under time-free condition, as shown in panel (a). More interestingly, there also exists an inverted-U shape between RT and utility difference under time pressure and time delay (panels (b) and (c)). Mixed-effects regressions with log(RT) as the dependent variable show that RT is negatively related with the absolute value of utility difference under all the three time conditions (t(4997) = -14.059, p < 0.001 for time-free condition, t(4997) =

[19] We get similar results if we calculate the utility difference using β_f in all three time conditions. The results are shown in Figure 2.12 in Appendix.

-9.954, p < 0.001 for time-pressure condition, and t(4997) = -2.539, p = 0.011 for time-delay condition). Therefore, we conclude that social decision making complies with the SSM predictions, even under time pressure and time delay.

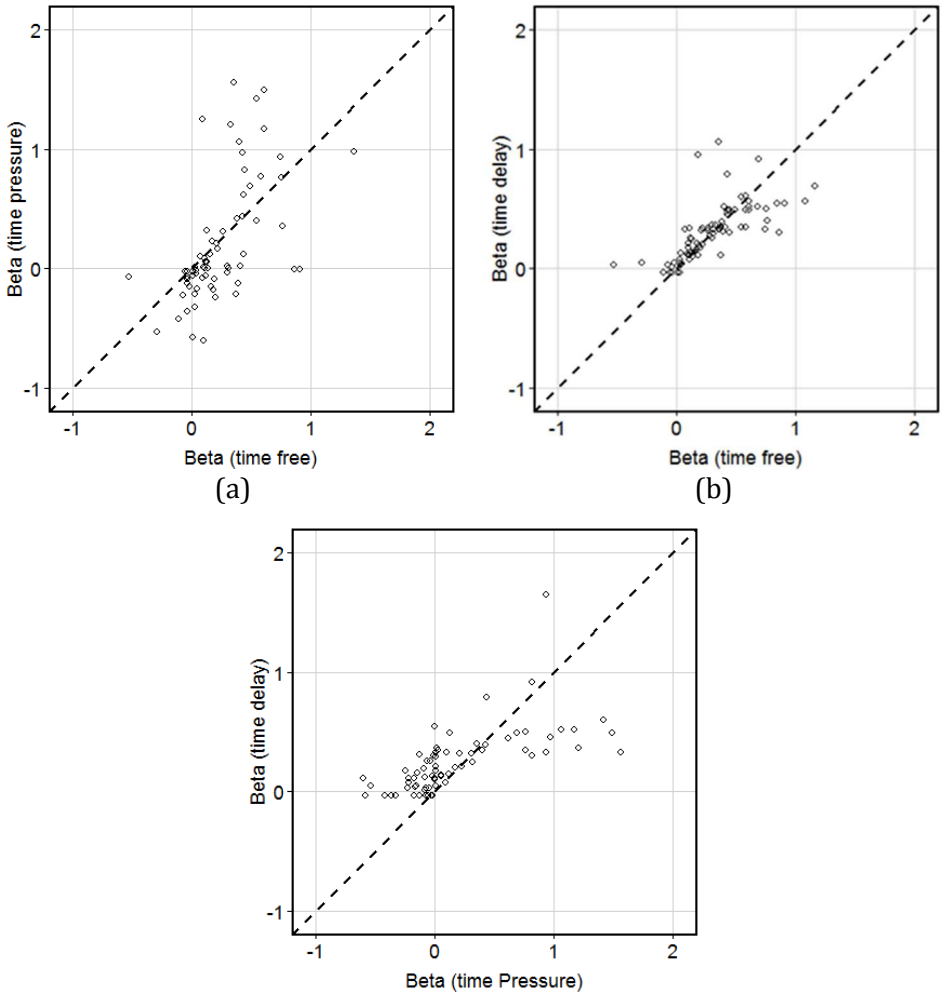


Figure 2.2 Advantageous inequality aversion under different time conditions
Each dot represents one subject's degree of prosociality (β), comparing (a) time-free and time-pressure conditions, (b) time-free and time-delay conditions, and (c) time-pressure and time-delay conditions[20]

[20] Figure 2 does not include 12 subjects whose β_f is outside of (-1, 2), 26 subjects whose β_p is outside of (-1, 2), and 8 subjects whose β_d is outside of (-1, 2). However, we never exclude any subjects in the following analysis.

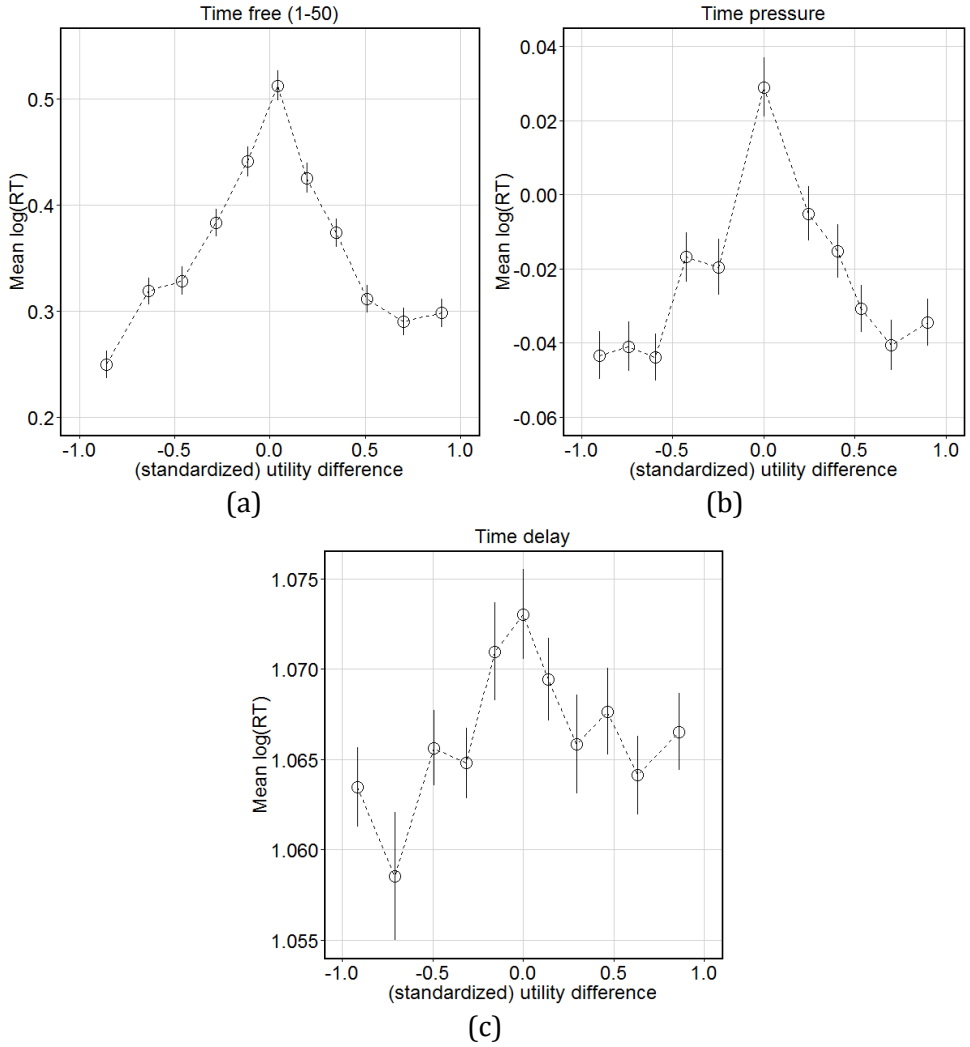


Figure 2.3 Correlations between log(RT) and utility difference
The x-axis is the average utility difference of each bin, and the y-axis is the average log(RT) of each bin

2.4 Modeling the Cognitive Process of Social Decision Making

The experimental evidence above paints a complex picture. On the one hand, the relationship that we observe between RTs and strength of preference in all conditions is consistent with a single SSM process underlying these decisions. On the other hand, the magnification of preferences under time pressure (and shrinking under time delay) is the opposite of what one would expect from time pressure in a simple SSM with no biases in starting point (i.e. priors). Therefore, next we investigate whether biases in the starting point of the SSM process can account for the data. In doing so, we provide a way to operationalize intuition in a SSM framework.

2.4.1 The Modeling Framework

The SSM with intuition (SSMi) is based on the drift-diffusion model (Ratcliff 1978, Ratcliff and McKoon 2008), which was originally developed for memory and perception and is increasingly used to study the cognitive process of economic decision making (Fehr and Rangel 2011, Frydman and Nave forthcoming, Krajbich, Armel, and Rangel 2010, Krajbich, Bartling, et al. 2015a). It assumes that the decision maker accumulates stochastic evidence favoring one option or the other until he/she is confident enough that one option is better. The SSMi models intuition as the starting point of accumulation, i.e. an initial bias towards selfishness or prosociality.

In our experimental setting, the SSMi assumes that decisions are initiated by a noisy process which accumulates relative evidence (R) that one option is better than the other. The relative evidence R follows a diffusion process and evolves in small time increments according to a stochastic difference equation,

$$R_{t+1} = R_t + v + s_t, \tag{2}$$

where v is the drift rate, which represents the average strength of preference for the selfish option, and s represents mean-zero Gaussian noise. A choice is made once the relative evidence R reaches one of the two thresholds, normalizing the prosocial threshold to zero and the selfish threshold to a constant, a. Another parameter in the model is the non-decision time, t_0, which denotes the time required for encoding the stimulus and motor execution.

The starting point R_0 represents the initial point of the SSM process. In most cases the literature assumes an unbiased starting point, that is, the process starts at the midpoint of the two thresholds ($R_0 = a/2$). There are however some perceptual studies which allow the starting point to deviate from $a/2$ to model a prior bias towards one option (Diederich and Busemeyer 2006, Ratcliff 1985). Here we allow the variability in starting point to represent a bias towards selfishness or prosociality, and we use these starting point biases to model the effect of intuition on the decision process.[21] Specifically, the process starts near the selfish threshold for subjects who are intuitively selfish (the red path in Figure 2.4), and the process starts near the prosocial threshold for subjects who are intuitively prosocial (the blue path in Figure 2.4). If there is no intuition bias involved, the process starts from the midpoint of the two thresholds (the black path in

[21] We discuss alternative models which capture the intuition by the drift rate or by both the starting point and the drift rate in Section 2.6.

Figure 2.4). That is, $R_0 > a/2$ for intuitively selfish subjects, $R_0 < a/2$ for intuitively prosocial subjects, and $R_0 = a/2$ for subjects who are neither.

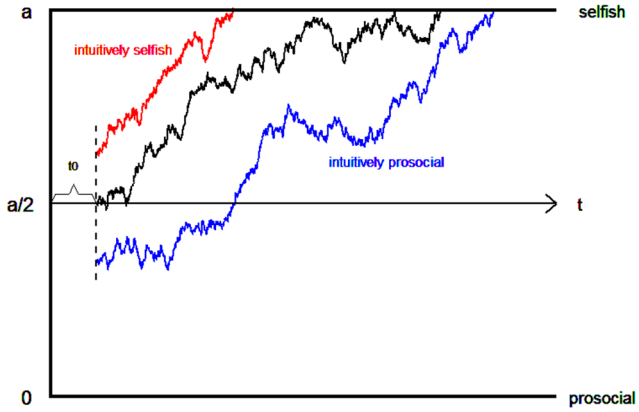


Figure 2.4 A graphical illustration of the SSMi
0 and a are the two preset thresholds for the prosocial option and the selfish option. Non-decision time, t_0, denotes the time required for encoding the stimulus and decision execution. The three paths indicate the evolution of the relative evidence (R) over time. The red path which starts near the threshold of the selfish option represents the SSMi process for intuitively selfish subjects; the blue path which starts near the threshold of the prosocial option represents the SSMi process for intuitively prosocial subjects; and the black path represents the unbiased SSM process. When R reaches one of the thresholds, the process terminates and a decision is initiated.

2.4.2 Model Estimation

Since the subjects are heterogeneous in whether their intuition favors prosociality or selfishness, to estimate the model, we split subjects into two types according to their inequality aversion parameters under time pressure (β_p) and time delay (β_d).[22] Specifically, we label subjects with $\beta_p < \beta_d$ as intuitively selfish and subjects with $\beta_p > \beta_d$ as intuitively prosocial. This classification yields 56 intuitively selfish subjects and 46 intuitively prosocial subjects.

We estimate the SSMi using the Kolmogorov-Smirnov method of Fast DM (Voss, Voss, and Lerche 2015). To do out-of-sample predictions, we estimate the model using the data of Games 1-50 and the data of Games 51-100 under time-free condition separately. In the estimation, the RT distribution is split into two parts, one for selfish choices and the other for prosocial choices. Thus, the probability of a selfish choice is implicitly represented in the RT distributions. A multidimensional search is conducted

[22] We split subjects into two types according to the inequality aversion parameter under time-pressure and time-delay conditions because we will do out-of-sample predictions for time-free condition in Section 2.5.

using a version of the simplex downhill method (Nelder and Mead 1965) to minimize the maximal vertical distance of the predicted and the empirical cumulative RT distributions.

Using this method, we estimate the various parameters of the SSMi. Most importantly, we estimate the relative starting point, z ($z \in [0,1]$). $z = 0.5$ if the starting point is in the middle of the two thresholds ($R_0 = a/2$). $z > 0.5$ if the process starts near the selfish threshold and $z < 0.5$ if the process starts near the prosocial threshold. In the estimation, we let z depend on the subjects' intuition type. The drift rate of the SSMi depends on the stimuli of the decision situation. Thus, we estimate a drift rate for each combination of the difference between dictator's and receiver's payoffs. Since we have 50 different combinations (5 differences between the dictator's payoffs and 10 differences between the receiver's payoffs) in our games, the estimation results in 50 drift rates. In the estimation, we also include the inter-trial variability of the relative starting point (szr), but we keep szr, non-decision time (t_0), and threshold (a) constant across games and subject types.

As hypothesized, the relative starting point (z) for intuitively selfish subjects is indeed larger than 0.5 (0.748 for Games 1-50 and 0.756 for Games 51-100), and for intuitively prosocial subjects is indeed less than 0.5 (0.374 for Games 1-50 and 0.373 for Games 51-100) (Table 2.1). To ensure that these results are not driven by extreme subjects, we estimate the model excluding 23 subjects whose proportion of selfish choices is higher than 0.95 or lower than 0.05, and also estimate the model excluding 28 subjects whose inequality aversion parameter (β_f) is greater than 3 or less than -3. We get similar results as in Table 2.1 in both cases.[23]

Table 2.1 Estimation results of SSMi

	Relative starting point (z)		Non-decision time (t_0)	Threshold (a)	Inter-trial variability of z (szr)
	Intuitively selfish	Intuitively prosocial			
Games 1-50	0.748 (0.005)	0.374 (0.005)	0.755 (0.028)	3.585 (0.039)	0.415 (0.034)
Games 51-100	0.756 (0.005)	0.373 (0.005)	0.688 (0.026)	3.616 (0.030)	0.376 (0.028)

Notes. The estimation is done using data for games in the time-free condition. The standard errors of the estimators are calculated using a jackknife method and reported in parentheses.

[23] The estimation results are shown in Table 2.5 and Table 2.6 in Appendix.

Moreover, similar to other studies which assume that drift rate depends on the utility difference between choice options, the drift rate in the estimation is also significantly related to the utility difference between choice options (Pearson correlation test, $r = 0.842$, $p < 0.001$ for Games 1-50, $r = 0.805$, $p < 0.001$ for Games 51-100), as shown in Fig. 2.13 in Appendix.

These estimation results are consistent with the hypothesis that intuition is captured by biased starting points in the SSM process. In the next two sections, we investigate alternative explanations for the data, comparing this model to alternative models.

2.5 Out-of-Sample Predictions

We first compare the out-of-sample predictions by the SSMi and the standard SSM. Then, we compare the out-of-sample predictions by the SSMi and the standard logistic model to test whether the process-based model outperforms the static model in describing and predicting social preferences.

2.5.1 SSMi vs. Standard SSM

The SSMi captures the intuition of selfishness or prosociality by the starting point. To study whether taking intuition into account improves predictive power, we compare the SSMi to a benchmark model in which we set the relative starting point to 0.5 for all subjects (standard SSM). More specifically, we estimate the two models using RT and choice data of Games 1-50 under time-free condition and use the estimated parameters to predict choices in Games 51-100, by simulating the model 5000 times for each game. We calculate the proportion of selfish choice in the simulations and take this proportion as the predicted probability of selfish decisions. We then compare these probabilities with the empirical probabilities of selfish choice in the experiment.

Figure 2.5 and Figure 2.6 display the out-of-sample prediction results for intuitively selfish and intuitively prosocial subjects separately. The panel (a) of each figure shows the predicted probability of selfish choice by the SSMi (y-axis) and the empirical probability of selfish choice in the experiment (x-axis). The panel (b) of each figure shows the predicted probability of selfish choice by the standard SSM (y-axis) and the empirical probability of selfish choice in the experiment (x-axis). If the model is perfectly accurate, all the points should be on the 45^o line. Obviously, the SSMi predicts more accurately than the standard SSM in both cases.

To compare the two models quantitatively, we calculate the absolute error (AE) between the predicted probability and the empirical probability for each game (Table 2.2). The SSMi has less AE than the standard SSM in 40 out of 50 games for intuitively selfish subjects and 44 out of 50 games for intuitively prosocial subjects (binomial tests, $p < 0.001$ for both intuitively selfish and prosocial subjects). Moreover, the summed AE for the SSMi is less than half of that for the standard SSM, in both cases.

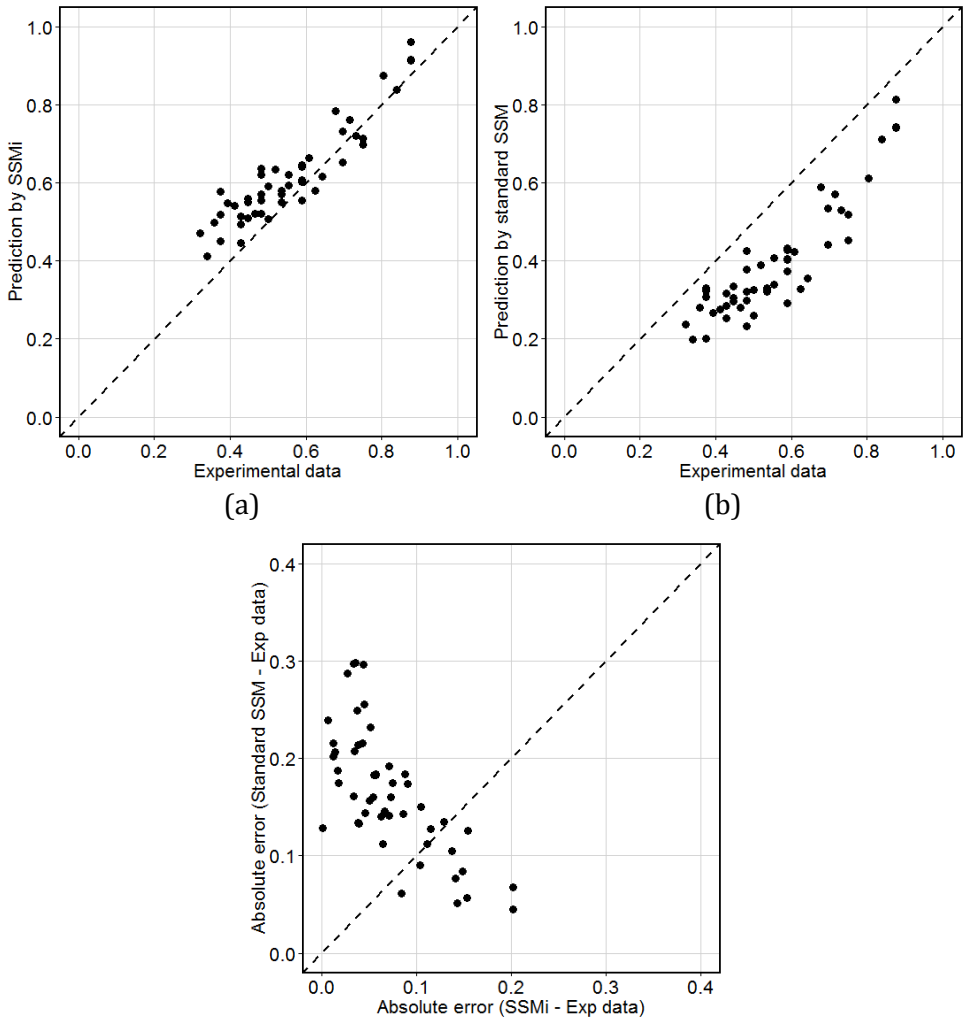


Figure 2.5 Predictions of SSMi and standard SSM for intuitively selfish subjects
Each dot represents one game

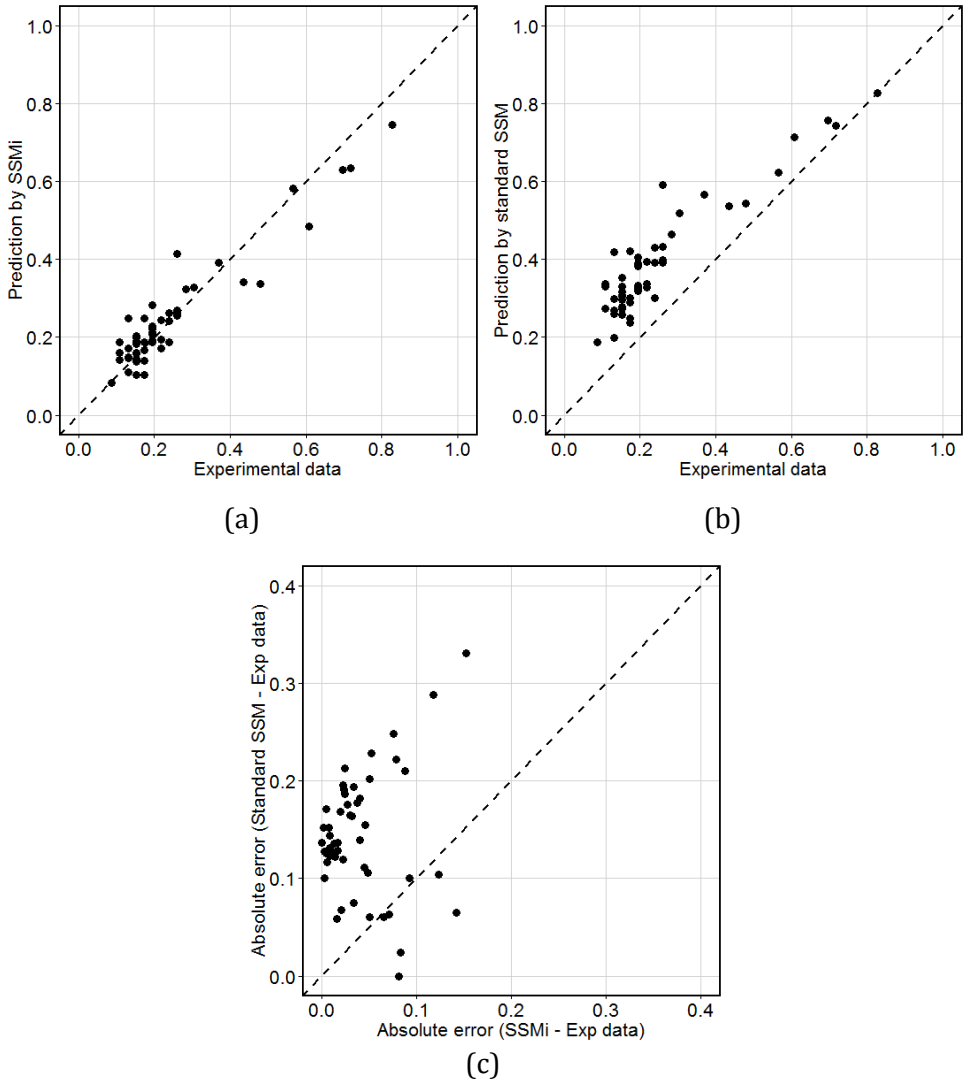


Figure 2.6 Predictions of SSMi and standard SSM for intuitively prosocial subjects
Each dot represents one game

The results in Table 2.2 are using data of Games 1-50 to predict decisions in Games 51-100. To test the robustness of the result, we do another 50 independent out-of-sample predictions. More specifically, we pool the data of Games 1-100 under time-free condition together and we randomly select half of the trials 50 times. In each time, we estimate the SSMi and the standard SSM using the selected half of the data. And then we use the estimated parameters to predict subjects' decisions in the other half of the data. In these 50 times of independent predictions, the SSMi gives more accurate predictions than the standard SSM. That is, $\sum AE_{SSMi}$ is less than $\sum AE_{SSM}$ in all these 50 predictions. And the SSMi has less AE in more than half of the games (mean = 38.200, sd = 3.470 for intuitively selfish subjects, and mean = 40.160, sd = 2.542 for intuitively prosocial subjects) than the standard SSM in all these 50 predictions. Therefore, the SSMi significantly improves the ability of out-of-sample predictions compared to the standard SSM.

Table 2.2 AE between predicted and empirical probabilities for SSMi and standard SSM

	$\sum AE_{SSMi}$	$\sum AE_{SSM}$	Num. of games in which $AE_{SSMi} < AE_{SSM}$
Intuitively selfish	3.589	8.201	40/50
Intuitively prosocial	2.052	7.181	44/50

Notes. In this table, we use the data of Games 1-50 to predict decisions in Games 51-100 under time-free condition. AE_{SSMi} is the absolute error between the predicted probability by the SSMi and the empirical probability of selfish choice in the experiment, and AE_{SSM} is the absolute error between the predicted probability of selfish choice by the standard SSM and the empirical data.

2.5.2 SSMi vs Logistic Model

The canonical static method for modelling binary decisions in economics is the logistic model (Luce 1959, Mcfadden 1974, Webb 2013). The SSMi is a dynamic equivalent to the logistic model that produces RTs in addition to choice probabilities. In this part, we study whether the process-based model, SSMi, outperforms the standard logistic model in out-of-sample predictions.

We choose the following logistic model as the benchmark,

$$Selfish = \beta_0 + \beta_1 * DictDiff + \beta_2 * ReceDiff + \beta_3 * IntuitionType + \varepsilon \quad (3)$$

In logistic model (3), the dependent variable is a dummy which indicates whether the choice is selfish or prosocial. The independent variables are the difference between the dictator's payoffs ($DictDiff$), the difference between the receiver's payoffs ($ReceDiff$), and a dummy which indicates whether the subject is intuitively selfish or prosocial ($IntuitionType$). The dummy *IntuitionType* is included to mimic the starting point bias in the SSMi, i.e. to capture a bias towards one type of response or the other. We first estimate the regression using data from Games 1-50 and then use the regression results to calculate the predicted probability of selfish choices for Games 51-100 under time-free condition.

Figure 2.7 and Figure 2.8 show the out-of-sample prediction results for intuitively selfish and intuitively prosocial subjects separately. For intuitively selfish subjects, the SSMi has less AE than the logistic model in 25 out of 50 games (panel (c) of Figure 2.7). But the sum of the AE for the SSMi is 3.589, which is higher than that for the logistic model 3.191 (Table 2.3). Therefore, the SSMi does not outperform the logistic model for intuitively selfish subjects. However, the SSMi does outperform the logistic model for intuitively prosocial subjects. The sum of the AE for the SSMi is 2.052, which is less than

that of the logistic model 3.771, and in 37 out of 50 games the SSMi is more accurate than the logistic model (binomial test, $p < 0.001$).

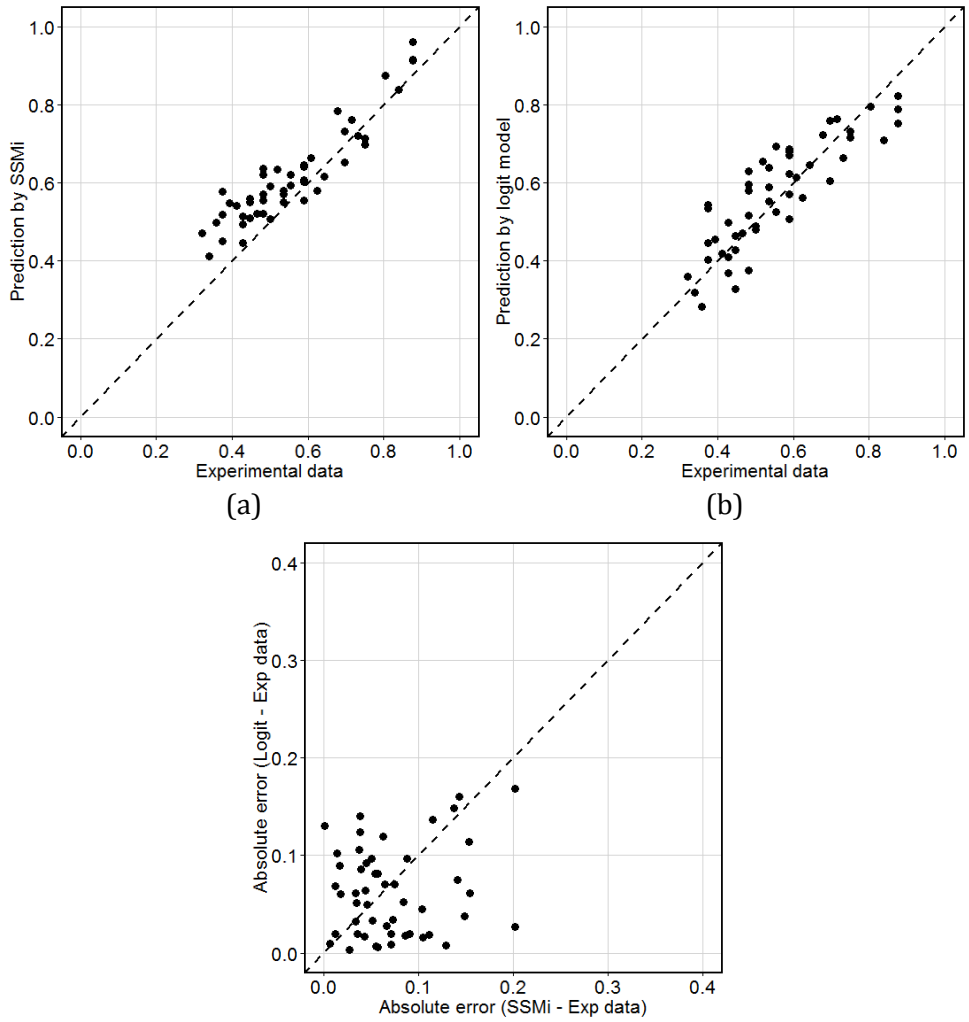


Figure 2.7 Predictions of SSMi and logistic model for intuitively selfish subjects
Each dot represents one game

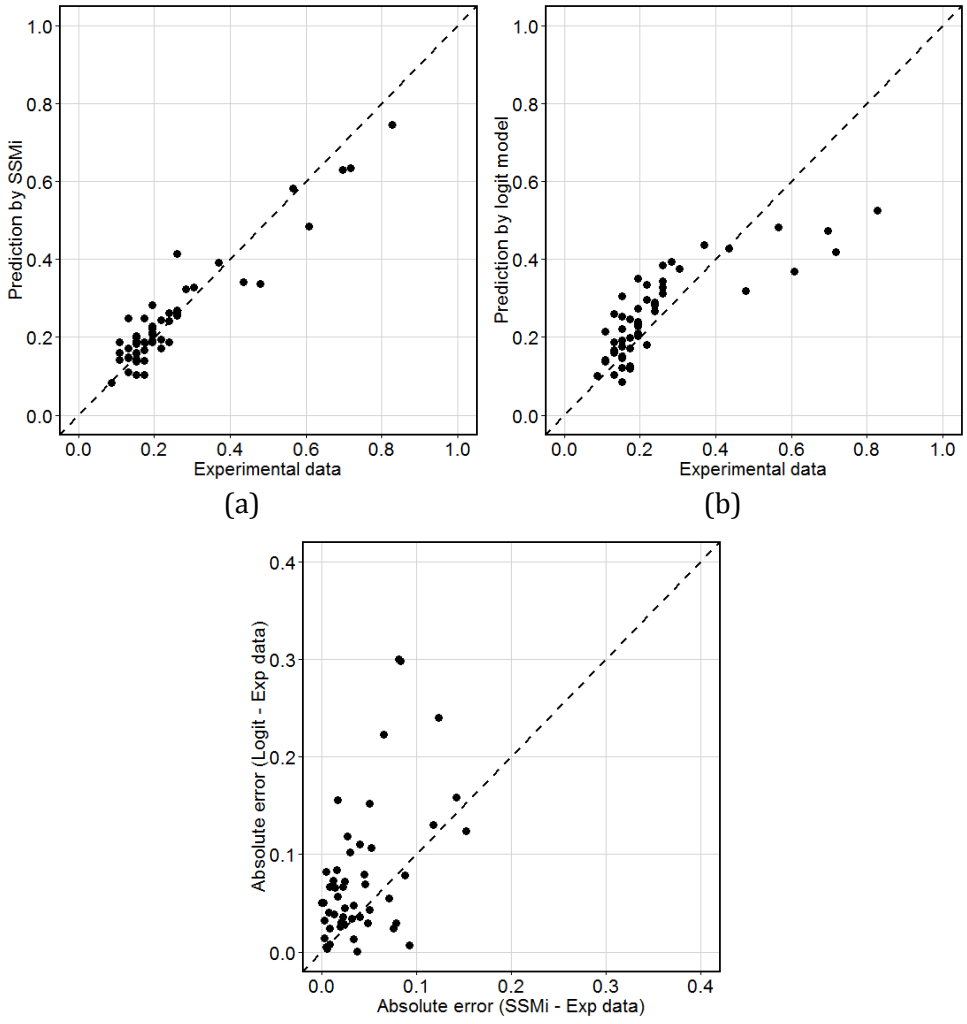


Figure 2.8 Predictions of SSMi and logistic model for intuitively prosocial subjects
Each dot represents one game

Table 2.3 AE between predicted and empirical probabilities for SSMi and logistic model

	$\sum AE_{SSMi}$	$\sum AE_{Logit}$	Num. of games in which $AE_{SSMi} < AE_{Logit}$
Intuitively selfish	3.589	3.191	25/50
Intuitively prosocial	2.052	3.771	37/50

Notes. In this table, we use the data of Games 1-50 to predict decisions in Games 51-100 under time-free condition. AE_{SSMi} is the absolute error between the predicted probability by the SSMi and the empirical data in the experiment, and AE_{Logit} is the absolute error between the predicted probability of selfish choice by the logistic model and the empirical data.

Since the SSMi outperforms the logistic model (3) only for intuitively prosocial subjects, but not for intuitively selfish subjects, next we study whether the SSMi outperforms the logistic model (3) at the aggregate level. We measure the aggregate

predictive performance using Cramer's λ (Clithero 2016a, Cramer 1999, Webb et al. 2016a) which is calculated as the following:

$$\lambda = \bar{P}^{+} - \bar{P}^{-} \tag{4}$$

where $\bar{P}^{+}$ and $\bar{P}^{-}$ denote the model's probability of choosing the selfish option on trials in which the selfish option is actually chosen and on trials in which the prosocial option is actually chosen, respectively. Thus, $\lambda \in [0,1]$ reflects how much of the choice variation across trials is captured by the model. $\lambda = 1$ indicates that the model can perfectly discriminate choice outcomes, and $\lambda = 0$ indicates that the model predicts decisions at chance. The Cramer's λ for the SSMi is 0.194, and the Cramer's λ for the logistic model (3) is 0.1637. Therefore, the SSMi has higher predictive performance than the logistic model (3) overall.

The SSMi uses both choice and RT data in the modeling, while the logistic model (3) only uses choice data. To check whether the outperformance of the SSMi is due to the additional RT data, we also compare the model with another logistic model that includes *RT*, as well as the interaction between *IntuitiveType* and *RT*,

$$Selfish = \beta_0 + \beta_1 * DictDiff + \beta_2 * ReceDiff + \beta_3 * IntuitionType + \beta_4 * RT + \beta_5 * IntuitionType * RT + \varepsilon \tag{5}$$

The Cramer's λ for the logistic model (5) is 0.1643, which is also less than that of the SSMi.[24]

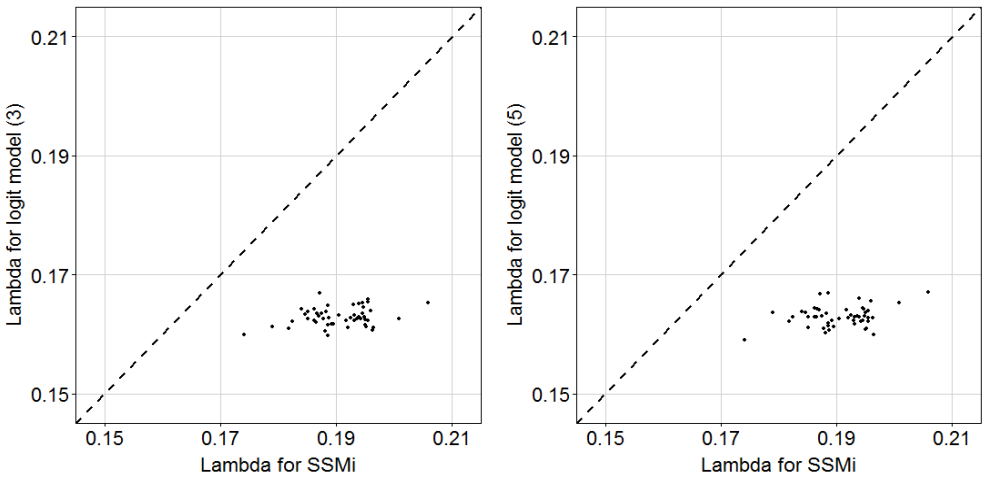


Figure 2.9 Cramer's λ of SSMi and logistic model

[24] We also calculate the Cramer's λ for the standard SSM and it is 0.085, which is less than those of the SSMi and logistic models (3) and (5).

As in the previous section, we also try pooling the data of Games 1-100, 50 times randomly select half of the trials, estimate the SSMi and the logistic model, and then use the estimated parameters to predict decisions in the other half of the data. Figure 2.9 shows the Cramer's λ for each of these 50 prediction comparisons. In every case the Cramer's λ for the SSMi ($mean = 0.191, sd = 0.006$) is greater than that for the logistic model (3) ($mean = 0.163, sd = 0.002$) and that for the logistic model (5) ($mean = 0.163, sd = 0.002$).[25] Therefore, we conclude that the SSMi significantly outperforms the standard logistic choice models in out-of-sample prediction.

2.6 Alternative Models

The SSMi captures the intuition of selfishness or prosociality by the starting point, which is akin to a bias in a Bayesian prior. However, an alternative possibility is that intuition could be reflected by a bias in the evidence accumulation process, which would be reflected in the drift rate. In perceptual decision making there is an ongoing debate about whether prior biases are encoded in the starting point or the drift rate (Gao et al. 2009, Hanks et al. 2011, Mulder et al. 2012, Rorie et al. 2010, Teodorescu and Usher 2013, van Ravenzwaaij et al. 2012, White and Poldrack 2014). In economic decision making, Frydman and Nave (forthcoming) show that the starting point is better to capture people's prior beliefs about stock performance, while Enax, Krajbich, and Weber (2016) show that nutrition information influences the drift rate of the choice process, but not the starting point.

In SSMs, time pressure is typically modeled as tighter decision boundaries. As the decision boundaries are drawn in, decisions are more determined by noise and less by drift rate, leading to a decrease in both RTs and accuracy. This is the speed-accuracy tradeoff. If intuition were captured by a bias in drift rate, then time pressure should decrease the choice bias; in fact, we observe the opposite. Nevertheless, we decide to estimate two alternative versions of the SSMi to investigate which model performs best in out-of-sample predictions.

One model captures the intuition by the drift rate (model DR), in which we let the drift rate not only depend on the combination of the difference between dictator's payoffs and the difference between receiver's payoffs, but also depend on the subjects' intuition type. The other model captures the intuition by both the starting point and the drift rate

[25] The Cramer's λ for the SSMi is also greater than that of the standard SSM ($mean = 0.086, sd = 0.003$) in the 50 independent predictions; the results are shown in Figure 2.14 in the Appendix.

(model SP+DR), in which we let both the starting point and the drift rate depend on subjects' intuition type. We compare the out-of-sample predictions of these two models with our baseline model which captures the intuition by the starting point alone (model SP).

The results of using the data of Games 1-50 to predict decisions in Games 51-100 are shown in Table 2.4. The sum of the AE for model SP is less than that of model DR and that of model SP+DR for both intuitively selfish and intuitively prosocial subjects. The model SP has more accurate predictions than the model DR and the model SP+DR in more than or equal to half of the games for both intuitively selfish and intuitively prosocial subjects.

Similar to previous analysis, we also randomly select one half of the pooled data 50 different times. Each time, we estimate the three models using the selected data and then predict decisions in the other half of the data. The out-of-sample prediction results show that the model SP gives more accurate predictions than the model DR and the model SP+DR.

Table 2.4 Out-of-sample prediction for different versions of SSMi

	Intuitively selfish	Intuitively prosocial
$\sum AE_{SP}$	3.579	2.001
$\sum AE_{DR}$	3.726	2.868
$\sum AE_{SP+DR}$	4.816	3.024
Num. of games in which $AE_{SP} < AE_{DR}$	25/50	34/50
Num. of games in which $AE_{SP} < AE_{SP+DR}$	32/50	35/50

Notes. In this table, we use the data of Games 1-50 to predict decisions in Games 51-100 in the time-free condition. AE_{SP} is the absolute error between the predicted probability by the model SP and the empirical probability of selfish choice in the experiment, AE_{DR} is the absolute error between the predicted probability of selfish choice by the model DR and the empirical data, and AE_{SP+DR} is the absolute error between the predicted probability of selfish choice by the model SP+DR and the empirical data.

Specifically, for intuitively selfish subjects, the sum of the AE of model SP ($mean = 4.164, sd = 0.557$) is less than that of model DR ($mean = 4.807, sd = 0.476$) in 41 of the 50 predictions (binomial test, $p < 0.001$), and less than that of model SP+DR ($mean = 6.278, sd = 0.664$) in all 50 predictions (Fig. 2.15 in Appendix). The model SP has more accurate predictions than the model DR for most of the games ($mean$ = 27.260, sd =

3.545) in 36 of the 50 cases (binomial test, p = 0.001), and has more accurate predictions than the model SP+DR for most of the games ($mean$ = 32.480, sd = 3.786) in 48 of the 50 predictions (binomial test, p < 0.001).

For intuitively prosocial subjects, the sum of the AE of model SP ($mean = 2.888, sd = 0.291$) is less than that of model DR ($mean = 3.436, sd = 0.378$) in all 50 predictions, and less than that of model SP+DR ($mean = 3.737, sd = 0.456$) in 45 of the 50 predictions (Fig. 2.16 in Appendix) (binomial test, p < 0.001). The model SP has more accurate predictions than the model DR for most of the games ($mean$ = 31.060, sd = 3.472) in 47 of the 50 predictions (binomial test, p < 0.001), and more accurate predictions than the model SP+DR for most of the games ($mean$ = 28.920, sd = 3.752) in 42 of the 50 predictions (binomial test, p < 0.001).[26]

2.7 Discussion and Conclusion

This chapter studies the underlying process of social decision making using mini-dictator games and time manipulation. In our experiment, subjects make binary decisions for mini-dictator games under time-free, time-pressure and time-delay conditions. The experimental evidence gives clear answers about the cognitive process of social decision making. First, subjects are heterogeneous in whether their intuition favors prosociality or selfishness. Second, the cognitive process of social decision making complies with SSMs not only under time-free condition, but also under time pressure and time delay. This result helps to resolve the debate and the conflicting results about whether intuition favors prosociality or selfishness, and also helps to resolve the debate on whether social decisions are governed by parallel competing processes or a single deliberative process.

We argue that the notion of intuition can be captured with the starting point in a SSM framework. The out-of-sample prediction results show that the SSMi improves the predictive performance compared to the standard SSM. Thus, it is necessary to take intuition into account when modeling the cognitive process of social decision making. In addition, we test alternative versions of the SSMi, and the results show that intuition is better explained by the starting point than the drift rate or both the starting point and the drift rate. Furthermore, the SSMi outperforms standard logistic choice models, which underlines the importance of investigating the cognitive process and building process-

[26] In Table 2.4, we only take the inter-trial variability of the starting point into account and we let the inter-trial variability constant across subject types. To test the robustness of these results, we also estimate the three models in which we take both the inter-trial variability of the drift rate and the starting point into account, and we let the inter-trial variability depends on subjects' type. We get similar out-of-sample prediction results, which are shown in Table 2.7 in Appendix.

based models to describe and predict human behavior. This also provides a first step in answering a common challenge posed to neuroeconomics that neuroeconomic models can uniquely help in empirical challenges of general interest to economics (Bernheim 2009).

The SSMi allows to take the heterogeneity of intuition into account. In this sense, our study provides an approach to test whether people's intuition favors prosociality or selfishness within the SSM framework in which the strength of preference is controlled. However, since we do not have enough trials at the individual level, we have to estimate the SSMi at the aggregate level. Future work can test the model at the individual level. It is surprising that we observe an inverted-U shape between RT and utility difference under time delay. One potential reason is that time delay make some subjects deliberate more, but others just wait until the 10 seconds are over and then starting making their decisions. Future studies can use eye-tracking to check what subjects actually look at under time delay and can also test the SSMi using eye-tracking data.

2.8 Appendix

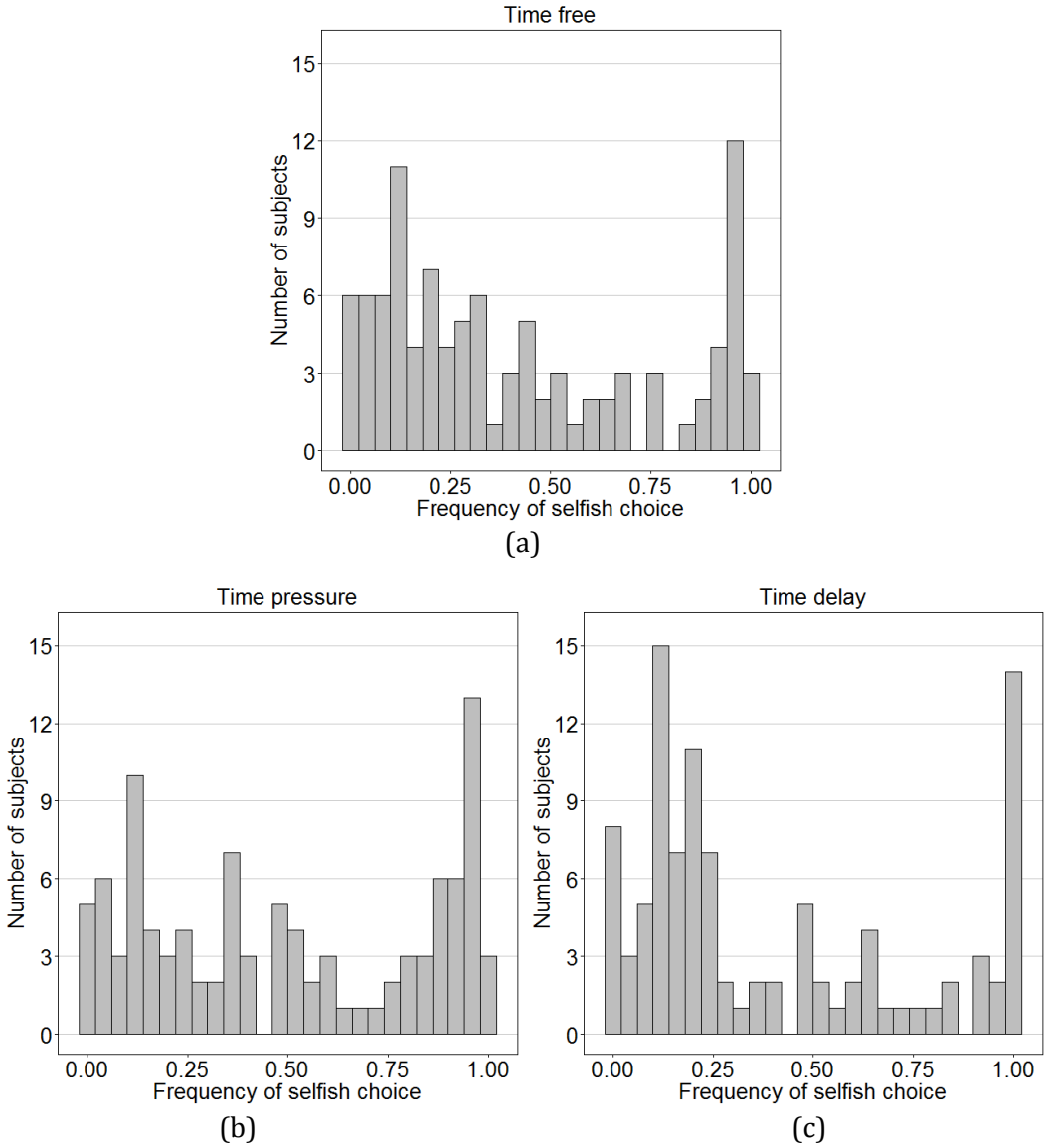


Figure 2.10 Distributions of the frequency of selfish choice

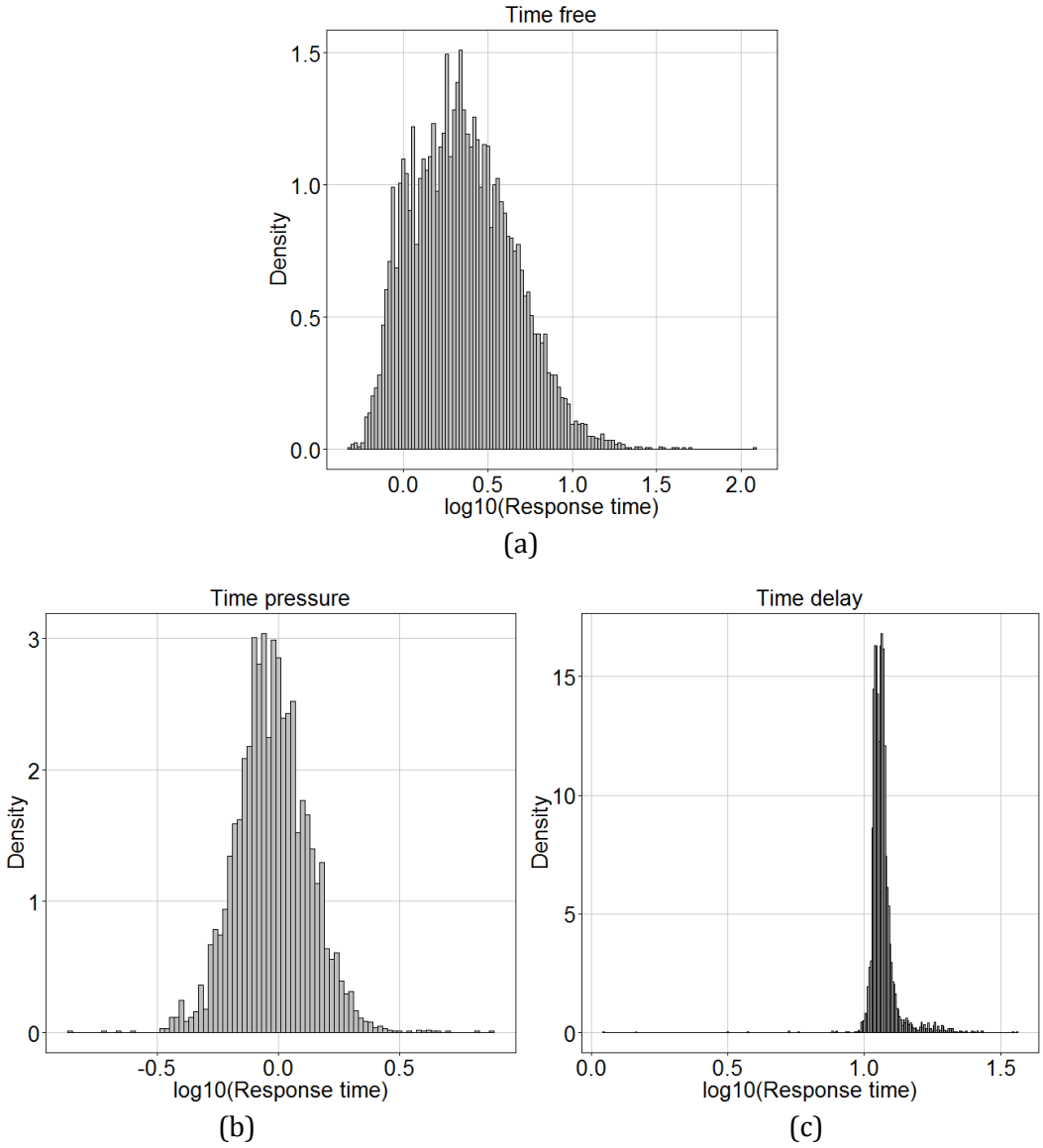


Figure 2.11 Distributions of the common logarithm of RTs

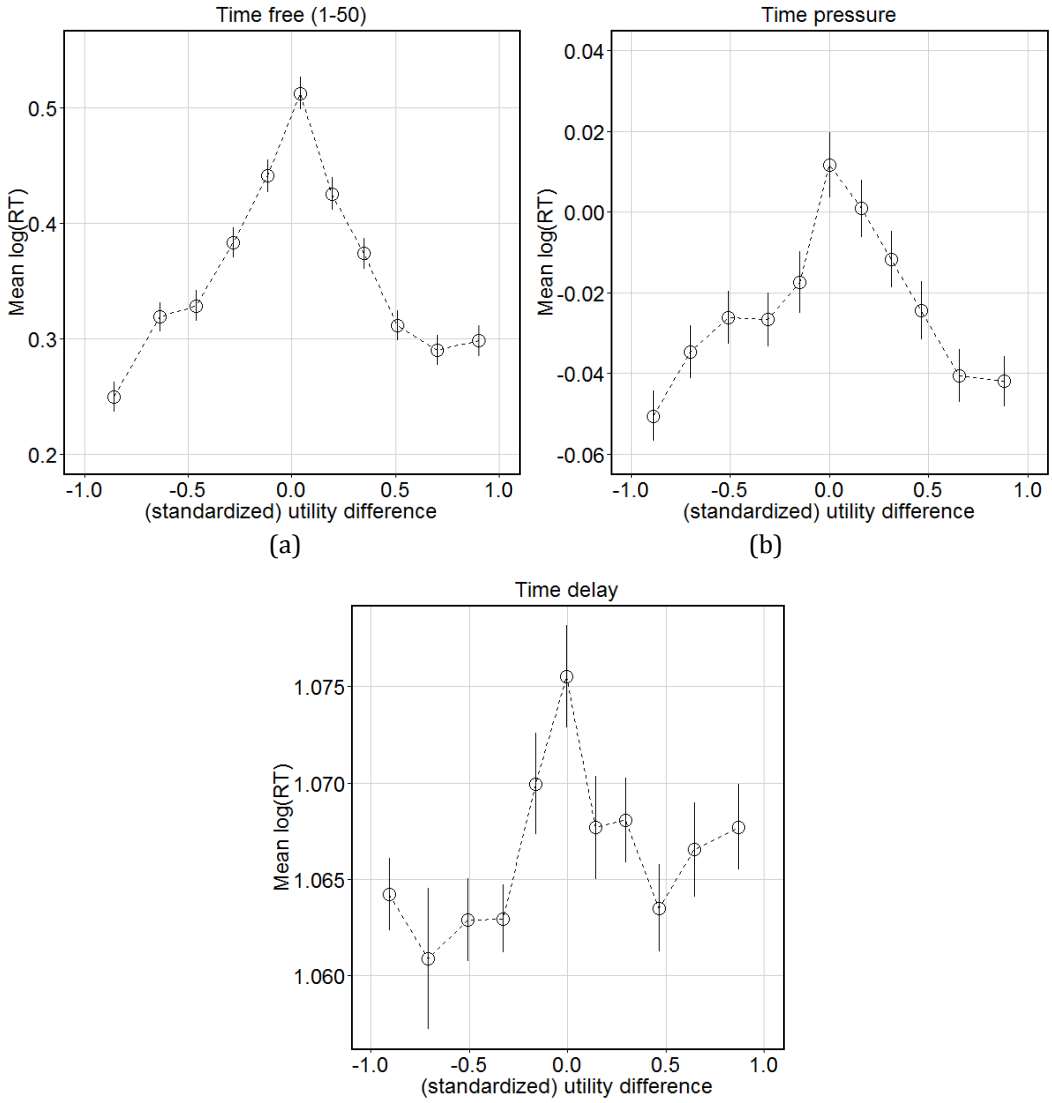


Figure 2.12 Correlation between RT and utility difference
In this figure, the utility difference under different time conditions is calculated using the same β_f. Mixed-effects regressions with log(RT) as the dependent variable show that RT is negatively related with the absolute value of the utility difference under the three time conditions($t(4997) = -14.059$, $p < 0.001$ for time-free condition; $t(4997) = -9.506$, $p < 0.001$ for time-pressure condition; $t(4997) = -0.298$, $p = 0.036$ for time-delay condition)

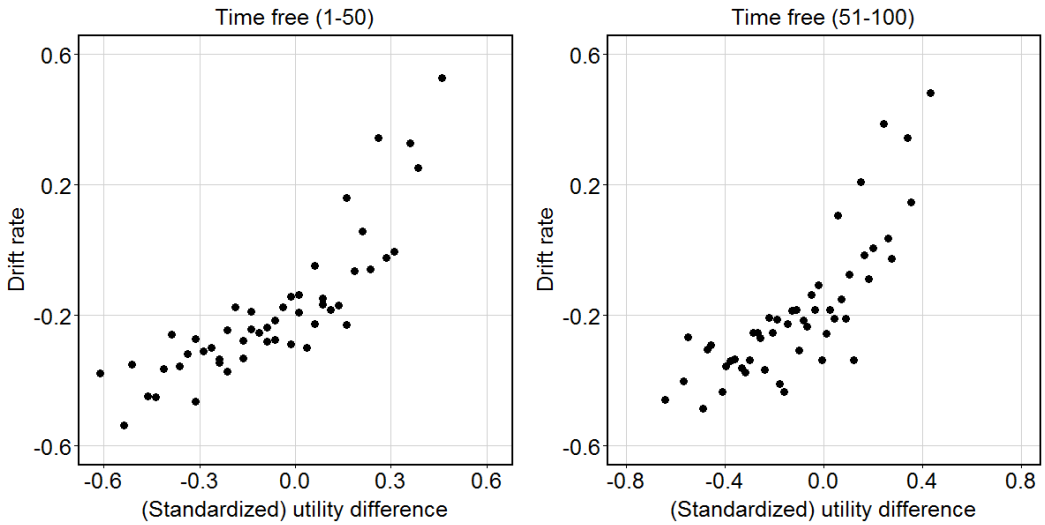


Figure 2.13 Correlation between the drift rate and the average utility difference
Each point represents one game. The correlation between the drift rate and the utility difference is 0.842 ($p = 1.821*10^{-14}$) for Games 1-50, and 0.805 ($p = 1.933*10^{-12}$) for Games 51-100 (Pearson correlation two-sided tests)

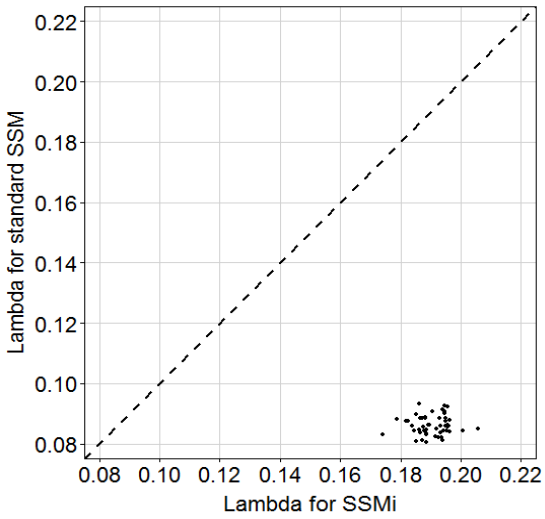


Figure 2.14 Cramer's λ for SSMi and standard SSM
The Cramer's λ for the SSMi ($mean = 0.191, sd = 0.006$) is greater than that of the standard SSM ($mean = 0.086, sd = 0.003$) in all 50 independent predictions

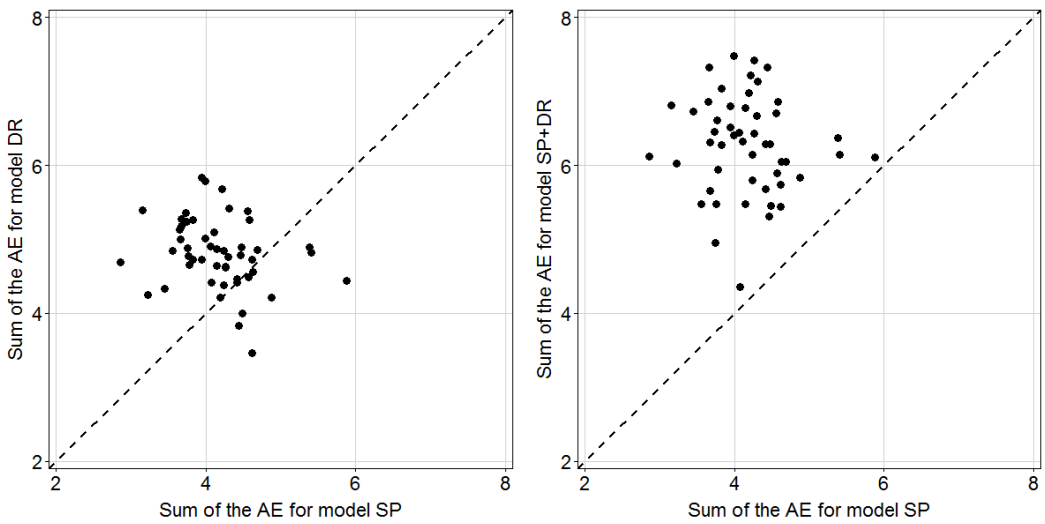


Figure 2.15 sum(AE) for different versions of SSMi (intuitively selfish subjects)
Each dot represents one out-of-sample prediction

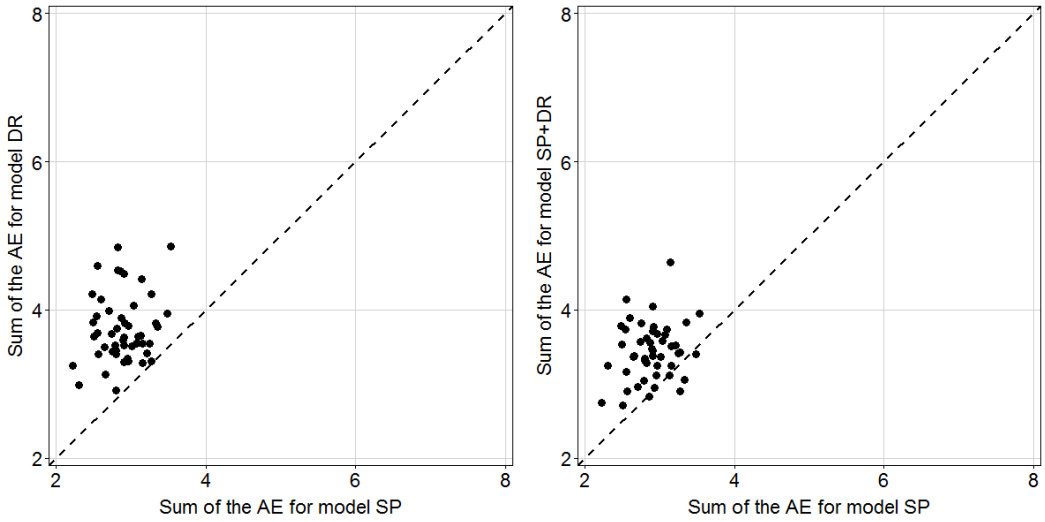


Figure 2.16 sum(AE) for different versions of SSMi (intuitively prosocial subjects)
Each dot represents one out-of-sample prediction

Table 2.5 Model estimation without extreme subjects

	Relative starting point (z)		Non-decision time (t_0)	Boundary (a)	Inter-trial variability of z (szr)
	Intuitively selfish	Intuitively prosocial			
Games 1-50	0.680	0.377	0.685	4.164	0.531
Games 51-100	0.708	0.410	0.851	3.684	0.569

Notes. In this estimation, we exclude subjects whose frequency of selfish choice is higher than 0.95 or lower than 0.05.

Table 2.6 Model estimation without subjects whose abs(β)>3

	Relative starting point (z)		Non-decision time (t_0)	Boundary (a)	Inter-trial variability of z (szr)
	Intuitively selfish	Intuitively prosocial			
Games 1-50	0.740	0.377	0.660	3.733	0.331
Games 51-100	0.743	0.348	0.739	3.763	0.448

Notes. In this estimation, we exclude subjects whose inequality aversion parameter (β) is greater than 3 or less than -3.

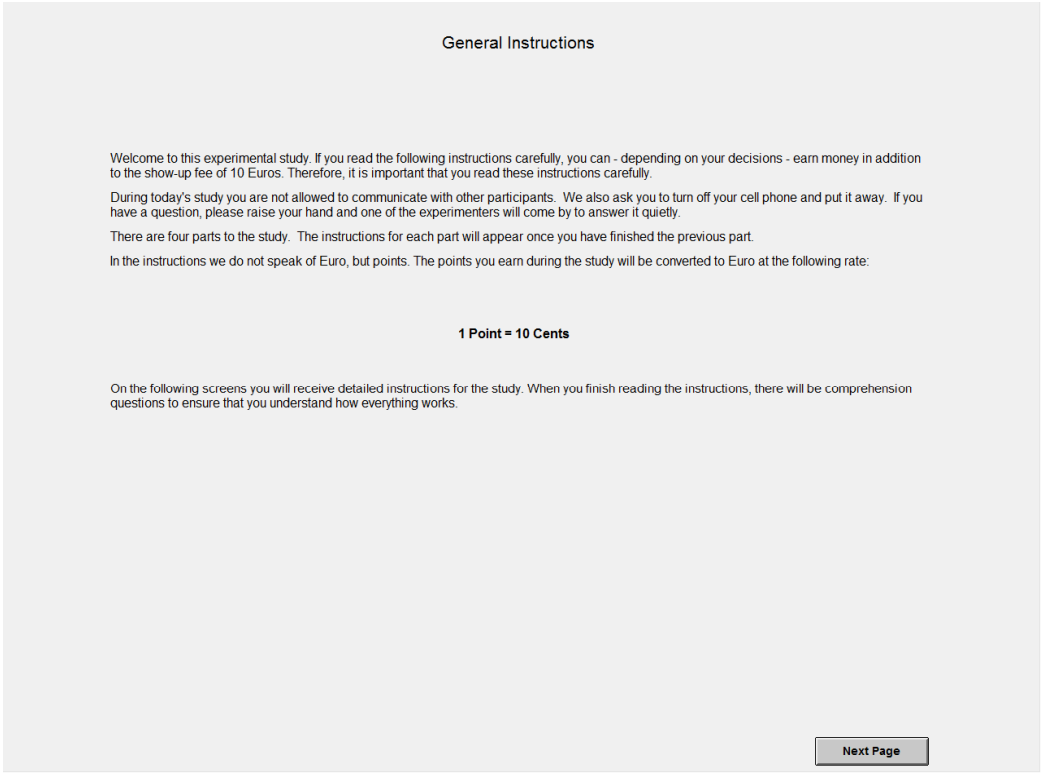
Table 2.7 Out-of-sample prediction for different versions of SSMi with inter-trial variability

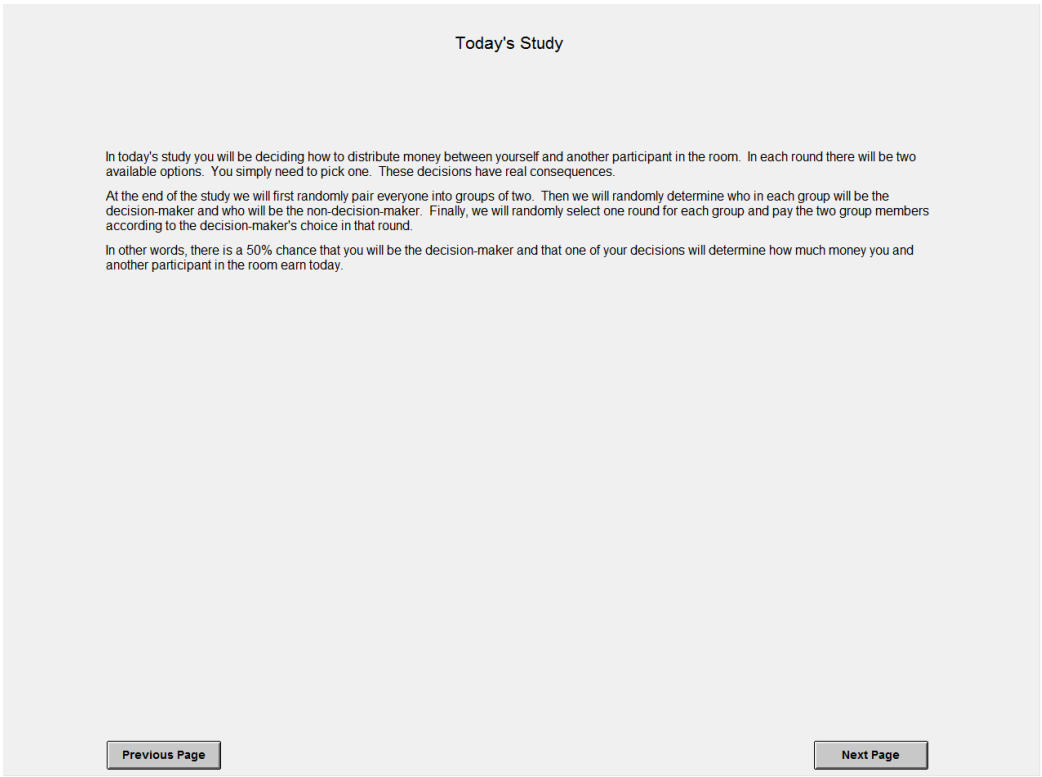
	Intuitively selfish	Intuitively prosocial
$\sum AE_{SP}$	4.518	3.466
$\sum AE_{DR}$	6.038	5.685
$\sum AE_{SP+DR}$	5.795	3.386
Num. of games in which $AE_{SP} < AE_{DR}$	31/50	37/50
Num. of games in which $AE_{SP} < AE_{SP+DR}$	27/50	29/50

Notes. In this table, we take the inter-trial variability of the drift rate and the starting point into account. We use the data of Games 1-50 to predict decisions in Games 51-100 in the time-free condition. The estimation is done using the Maximum Likelihood method of Fast DM. AE_{SP} is the absolute error between the predicted probability by the model SP and the empirical probability of selfish choice in the experiment, AE_{DR} is the absolute error between the predicted probability of selfish choice by the model DR and the empirical data, and AE_{SP+DR} is the absolute error between the predicted probability of selfish choice by the model SP+DR and the empirical data.

2.9 Instructions

In the experiment, we displayed the instructions of on the screen as the following:





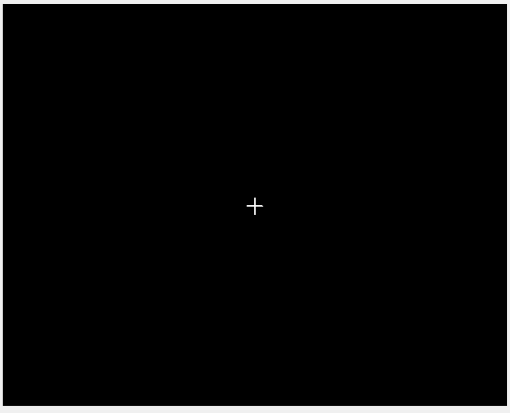
Example round

This is an example of the kind of decision you will be making in today's study. There are two possible options: the left option (Option A) and the right option (Option B). In Option A you get 65 points while the other person gets 31 points. In Option B you get 58 points while the other person gets 47 points. To choose Option A you would press the "F" key and to choose Option B you would press the "J" key. These keys are also indicated below each option, in case you forget.

Previous Page

Next Page

Between each round there will be a page that simply has a cross at the center of the screen. On this page, press the "Spacebar" to continue to the next round.



Previous Page

Continue

Q1: Suppose you press key "F" in the decision situation above,

1a) How many points do you get in this situation?

1b) How many points does the other person in your group get in this situation?

Q2: Suppose you press key "J" in the decision situation above,

2a) How many points do you get in this situation?

2b) How many points does the other person in your group get in this situation?

Instructions

OK

This is the first part of the study.

Please press "Spacebar" to continue.

You have finished the first part of the study.

Please click "Continue" to advance to the next part.

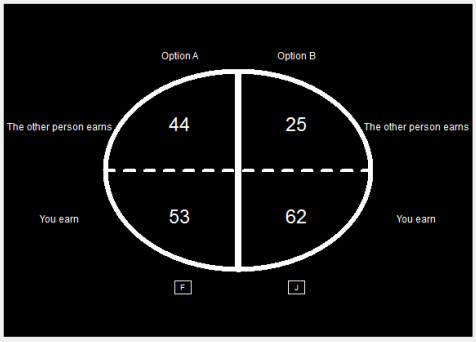
Continue

This is the second part of the study.

In this part you have to make each decision within 2 seconds.
If you do not make your decision within that time limit, you still need to indicate your choice, but then there is only a 10% chance that your decision will be implemented and a 90% chance that you will earn 0 points and the other person in your group will earn the lower of the two amounts that he/she could have earned in this round.

Please click "Continue" to answer the comprehension questions.

Continue



Q: Suppose you press key "F" in the decision situation above, but you do not make your decision within 2 seconds.

a) What is the chance that you will earn 53 points if this round is selected for payment?

- 10%
- 90%

b) What is the chance that you will earn 0 points if this round is selected for payment?

- 10%
- 90%

c) What is the chance that the other person will earn 44 points if this round is selected for payment?

- 10%
- 90%

d) What is the chance that the other person will earn 25 points if this round is selected for payment?

- 10%
- 90%

Instructions

OK

You made your decision too slowly. Please make your decisions before 2 seconds have elapsed.

Please press "Spacebar" to continue.

This is the third part of the study.

In this part you must wait 10 seconds before making your decision.
If you do not wait for the full 10 seconds then there is only a 10% chance that your decision will be implemented and a 90% chance that you will earn 0 points and the other person in your group will earn the lower of the two amounts that he/she could have earned in this round. Also, please note that you cannot speed up this part of the study by choosing quickly because no matter when you make your choice, you will always have to wait 10 seconds.

Please click "Continue" to answer the comprehension questions.

Continue

Q: Suppose you press the "F" key in the decision situation above, but you press the key before 10 seconds has elapsed.

a) What is the chance that you will earn 53 points if this round is selected for payment?

- 10%
- 90%

b) What is the chance that you will earn 0 points if this round is selected for payment?

- 10%
- 90%

c) What is the chance that the other person will earn 44 points if this round is selected for payment?

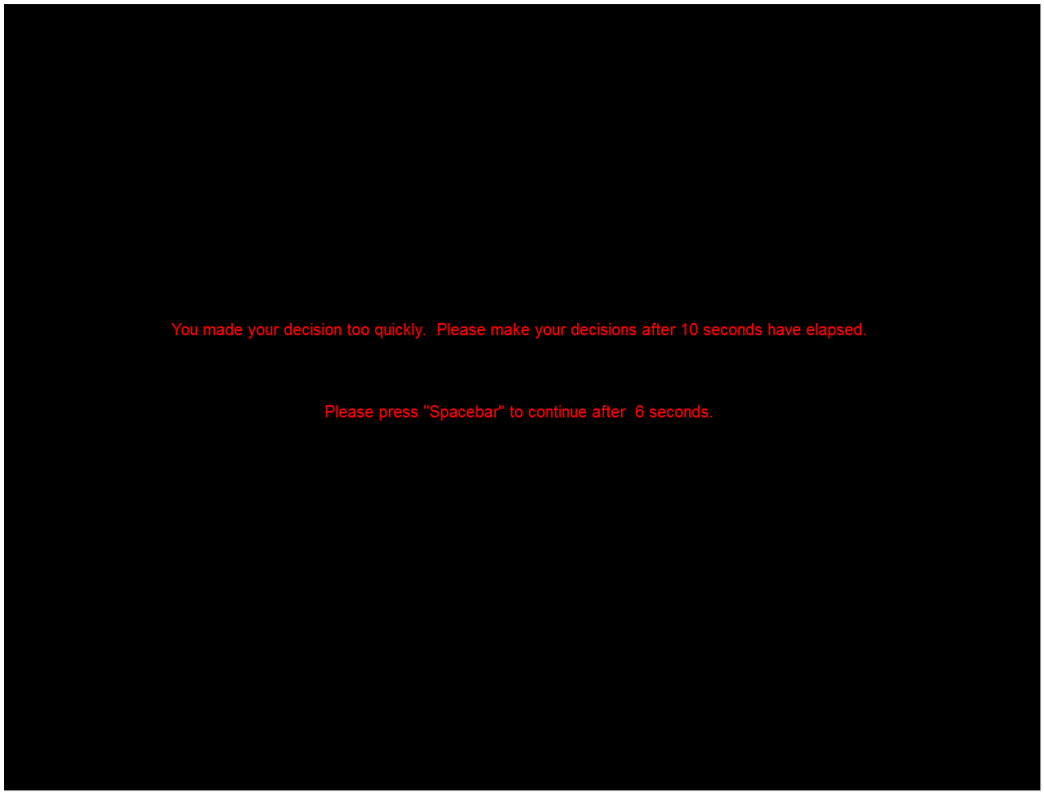
- 10%
- 90%

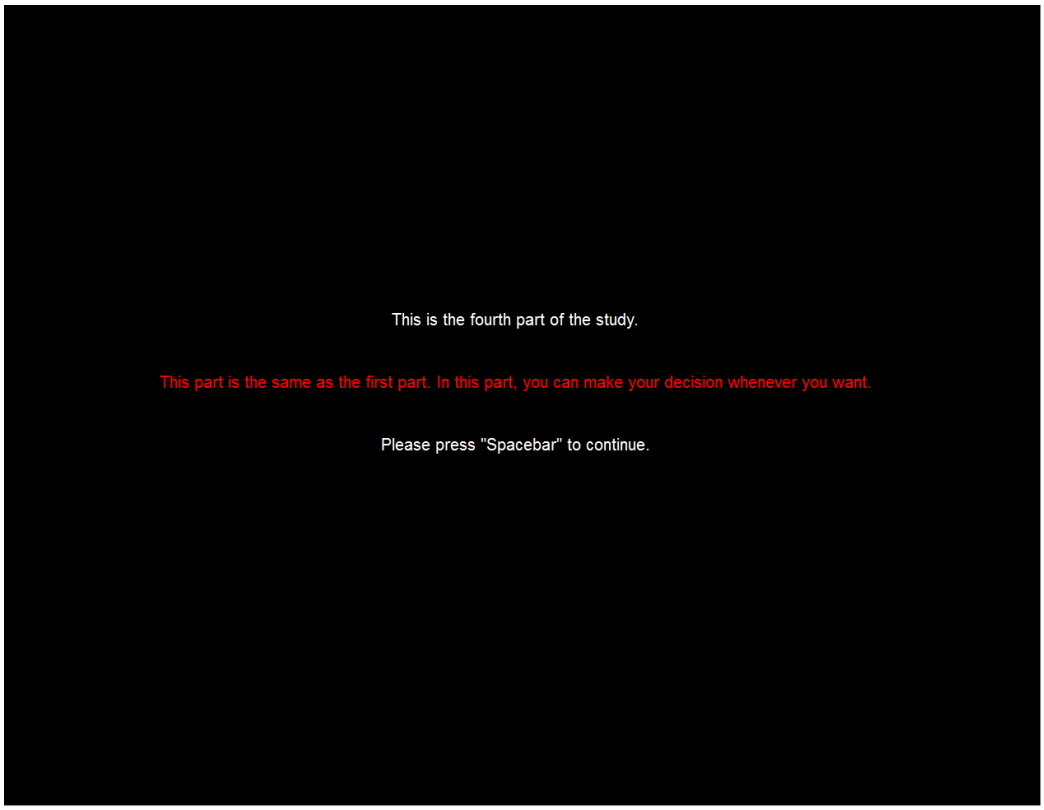
d) What is the chance that the other person will earn 25 points if this round is selected for payment?

- 10%
- 90%

Instructions

OK


You made your decision too quickly. Please make your decisions after 10 seconds have elapsed.
Please press "Spacebar" to continue after 6 seconds.


This is the fourth part of the study.
This part is the same as the first part. In this part, you can make your decision whenever you want.
Please press "Spacebar" to continue.

2.10 References of Chapter 2

Alós-Ferrer, Carlos. forthcoming. "A Dual-Process Diffusion Model." *Journal of Behavioral Decision Making*.

Alós-Ferrer, Carlos, and Fritz Strack. 2014. "From Dual Processes to Multiple Selves: Implications for Economic Behavior." *Journal of Economic Psychology* 41:1-11.

Bernheim, B Douglas. 2009. "On the Potential of Neuroeconomics: A Critical (but Hopeful) Appraisal." *American Economic Journal: Microeconomics* 1 (2):1-41.

Brocas, Isabelle, and Juan D. Carrillo. 2014. "Dual-Process Theories of Decision-Making: A Selective Survey." *Journal of Economic Psychology* 41:45-54.

Caplin, Andrew, and Daniel Martin. 2016. "The Dual-Process Drift Diffusion Model: Evidence from Response Times." *Economic Inquiry* 54 (2):1274-82.

Cappelen, Alexander W., Ulrik H. Nielsen, Bertil Tungodden, Jean-Robert Tyran, and Erik Wengström. 2015. "Fairness Is Intuitive." *Experimental Economics*:1-14.

Cappelletti, Dominique, Werner Güth, and Matteo Ploner. 2011. "Being of Two Minds: Ultimatum Offers under Cognitive Constraints." *Journal of Economic Psychology* 32 (6):940-50.

Charness, Gary, and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117 (3):817-69.

Chen, Fadong, and Urs Fischbacher. 2016. "Response Time and Click Position: Cheap Indicators of Preferences." *Journal of the Economic Science Association* 2 (2):109-26.

Clithero, John A. 2016a. "Improving out-of-Sample Predictions Using Response Times and a Model of the Decision Process." *Available at SSRN 2798459*.

Clithero, John A. 2016b. "Response Times in Economics: Looking through the Lens of Sequential Sampling Models." *Available at SSRN*.

Cramer, Jan Salomon. 1999. "Predictive Performance of the Binary Logit Model in Unbalanced Samples." *Journal of the Royal Statistical Society: Series D (The Statistician)* 48 (1):85-94.

Diederich, Adele, and Jerome R Busemeyer. 2006. "Modeling the Effects of Payoff on Response Bias in a Perceptual Discrimination Task: Bound-Change, Drift-Rate-Change, or Two-Stage-Processing Hypothesis." *Perception & Psychophysics* 68 (2):194-207.

Duffy, Sean, and John Smith. 2014. "Cognitive Load in the Multi-Player Prisoner's Dilemma Game: Are There Brains in Games." *Journal of Behavioral and Experimental Economics* 51:47-56.

Enax, Laura, Ian Krajbich, and Bernd Weber. 2016. "Salient Nutrition Labels Increase the Integration of Health Attributes in Food Decision-Making." *Judgment and Decision Making* 11 (5):460.

Epstein, Seymour. 1994. "Integration of the Cognitive and the Psychodynamic Unconscious." *American psychologist* 49 (8):709.

Evans, Jonathan St BT. 2008. "Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition." *Annual Review of Psychology* 59:255-78.

Fehr, Ernst, and Antonio Rangel. 2011. "Neuroeconomic Foundations of Economic Choice—Recent Advances." *Journal of Economic Perspectives* 25 (4):3-30.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3):817-68.

Fiedler, Susann, Andreas Glöckner, Andreas Nicklisch, and Stephan Dickert. 2013. "Social Value Orientation and Information Search in Social Dilemmas: An Eye-Tracking Analysis." *Organizational Behavior and Human Decision Processes* 120 (2):272-84.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2):171-8.

Frederick, Shane. 2005. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives* 19 (4):25-42.

Frydman, Cary, and Ian M Krajbich. 2016. "Using Response Times to Infer Others' Beliefs: An Application to Information Cascades." *Available at SSRN 2817026*.

Frydman, Cary, and Gideon Nave. forthcoming. "Extrapolative Beliefs in Perceptual and Economic Decisions: Evidence of a Common Mechanism." *Management Science*.

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki. 2015. "Stochastic Choice and Optimal Sequential Sampling." *Available at SSRN 2602927*.

Gao, Juan, KongFatt Wong-Lin, Philip Holmes, Patrick Simen, and Jonathan D Cohen. 2009. "Sequential Effects in Two-Choice Reaction Time Tasks: Decomposition and Synthesis of Mechanisms." *Neural Computation* 21 (9):2407-36.

Greiner, Ben. 2015. "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." *Journal of the Economic Science Association* 1 (1):114-25.

Hanks, Timothy D, Mark E Mazurek, Roozbeh Kiani, Elisabeth Hopp, and Michael N Shadlen. 2011. "Elapsed Decision Time Affects the Weighting of Prior Probability in a Perceptual Decision Task." *Journal of Neuroscience* 31 (17):6339-52.

Hauge, Karen Evelyn, Kjell Arne Brekke, Lars-Olof Johansson, Olof Johansson-Stenman, and Henrik Svedsäter. 2016. "Keeping Others in Our Mind or in Our Heart? Distribution Games under Cognitive Load." *Experimental Economics* 19 (3):562-76.

Hutcherson, Cendri A., Benjamin Bushong, and Antonio Rangel. 2015. "A Neurocomputational Model of Altruistic Choice and Its Implications." *Neuron* 87 (2):451-62.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*: Macmillan.

Kahneman, Daniel, and Shane Frederick. 2002. "Representativeness Revisited: Attribute Substitution in Intuitive Judgment." *Heuristics and biases: The psychology of intuitive judgment* 49.

Konovalov, Arkady, and Ian Krajbich. 2015. "Revealed Indifference: Using Response Times to Infer Preferences." *Working Paper*.

Krajbich, Ian, Carrie Armel, and Antonio Rangel. 2010. "Visual Fixations and the Computation and Comparison of Value in Simple Choice." *Nature Neuroscience* 13 (10):1292-8.

Krajbich, Ian, Bjorn Bartling, Todd Hare, and Ernst Fehr. 2015. "Rethinking Fast and Slow Based on a Critique of Reaction-Time Reverse Inference." *Nature Communications* 6:7455.

Krajbich, Ian, Todd Hare, Bjoern Bartling, Yosuke Morishima, and Ernst Fehr. 2015. "A Common Mechanism Underlying Food Choice and Social Decisions." *PLoS Comput Biol* 11 (10):e1004371.

Krajbich, Ian, Bastiaan Oud, and Ernst Fehr. 2014. "Benefits of Neuroeconomic Modeling: New Policy Interventions and Predictors of Preference." *American Economic Review* 104 (5):501-6.

Lohse, Johannes, Timo Goeschl, and Johannes H. Diederich. forthcoming. "Giving Is a Question of Time: Response Times and Contributions to an Environmental Public Good." *Environmental and Resource Economics*.

Lotito, Gianna, Matteo Migheli, and Guido Ortona. 2013. "Is Cooperation Instinctive? Evidence from the Response Times in a Public Goods Game." *Journal of Bioeconomics* 15 (2):123-33.

Luce, R.. Ducan. 1959. *Individual Choice Behavior a Theoretical Analysis*: John Wiley and sons.

Mcfadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." *Frontiers in Econometrics*.

Mulder, Martijn J, Eric-Jan Wagenmakers, Roger Ratcliff, Wouter Boekel, and Birte U Forstmann. 2012. "Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff." *Journal of Neuroscience* 32 (7):2335-43.

Nelder, John A, and Roger Mead. 1965. "A Simplex Method for Function Minimization." *The Computer Journal* 7 (4):308-13.

Nielsen, Ulrik H., Jean-Robert Tyran, and Erik Wengström. 2014. "Second Thoughts on Free Riding." *Economics Letters* 122 (2):136-9.

Piovesan, Marco, and Erik Wengström. 2009. "Fast or Fair? A Study of Response Times." *Economics Letters* 105 (2):193-6.

Rand, David G., Joshua D. Greene, and Martin A. Nowak. 2012. "Spontaneous Giving and Calculated Greed." *Nature* 489 (7416):427-30.

Ratcliff, Roger. 1978. "A Theory of Memory Retrieval." *Psychological Review* 85 (2):59-108.

Ratcliff, Roger. 1985. "Theoretical Interpretations of the Speed and Accuracy of Positive and Negative Responses." *Psychological review* 92 (2):212.

Ratcliff, Roger, and Gail McKoon. 2008. "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks." *Neural Computation* 20 (4):873-922.

Rorie, Alan E, Juan Gao, James L McClelland, and William T Newsome. 2010. "Integration of Sensory and Reward Information During Perceptual Decision-Making in Lateral Intraparietal Cortex (Lip) of the Macaque Monkey." *PloS one* 5 (2):e9308.

Rubinstein, Ariel. 2007. "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal* 117 (523):1243-59.

Rubinstein, Ariel. 2016. "A Typology of Players: Between Instinctive and Contemplative." *Quarterly Journal of Economics* 131 (2):859-90.

Schneider, Walter, and Richard M Shiffrin. 1977. "Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention." *Psychological review* 84 (1):1.

Schulz, Jonathan F., Urs Fischbacher, Christian Thöni, and Verena Utikal. 2014. "Affect and Fairness: Dictator Games under Cognitive Load." *Journal of Economic Psychology* 41:77-87.

Spiliopoulos, Leonidas, and Andreas Ortmann. 2016. "The Bcd of Response Time Analysis in Experimental Economics." *Available at SSRN 2401325*.

Strack, Fritz, and Roland Deutsch. 2004. "Reflective and Impulsive Determinants of Social Behavior." *Personality and Social Psychology Review* 8 (3):220-47.

Teodorescu, Andrei R, and Marius Usher. 2013. "Disentangling Decision Models: From Independence to Competition." *Psychological Review* 120 (1):1-38.

Tinghog, Gustav, David Andersson, Caroline Bonn, Harald Bottiger, Camilla Josephson, Gustaf Lundgren, Daniel Vastfjall, Michael Kirchler, and Magnus Johannesson. 2013. "Intuition and Cooperation Reconsidered." *Nature* 498 (7452):E1-E2.

van Ravenzwaaij, Don, Martijn Mulder, Francis Tuerlinckx, and Eric-Jan Wagenmakers. 2012. "Do the Dynamics of Prior Information Depend on Task Context? An Analysis of Optimal Performance and an Empirical Test." *Frontiers in Psychology* 3:132.

Voss, Andreas, Jochen Voss, and Veronika Lerche. 2015. "Assessing Cognitive Processes with Diffusion Model Analyses: A Tutorial Based on Fast-Dm-30." *Frontiers in Psychology* 6 (336).

Webb, Ryan. 2013. "Dynamic Constraints on the Distribution of Stochastic Choice: Drift Diffusion Implies Random Utility." *unpublished, New York University*.

Webb, Ryan, Paul W Glimcher, Ifat Levy, Stephanie C Lazzaro, and Robb B Rutledge. 2016. "Neural Random Utility: Relating Cardinal Neural Observables to Stochastic Choice Behaviour." *Working Paper*.

Weber, Elke U, and Eric J Johnson. 2009. "Mindful Judgment and Decision Making." *Annual Review of Psychology* 60:53.

White, Corey N, and Russell A Poldrack. 2014. "Decomposing Bias in Different Types of Simple Decisions." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 40 (2):385.

Woodford, Michael. 2014. "Stochastic Choice: An Optimizing Neuroeconomic Model." *American Economic Review* 104 (5):495-500.

3 Cognitive Processes of Prosocial and Antisocial Punishment

3.1 Introduction

Prosocial punishment plays an important role in promoting cooperation and is crucial for the maintenance of cooperation (Fehr and Gächter 2002, Gürerk, Irlenbusch, and Rockenbach 2006, Henrich et al. 2006), while antisocial punishment is detrimental for the efficiency of the informal punishment mechanism (Falk, Fehr, and Fischbacher 2005, Fehr, Hoff, and Kshetramade 2008, Gächter and Herrmann 2011, Herrmann, Thöni, and Gächter 2008). To better understand punishment behavior and promote cooperation efficiently, research across a wide variety of fields has studied the motivations underlying prosocial and antisocial punishment (Bone and Raihani 2015, Fehr and Gächter 2002, Fowler, Johnson, and Smirnov 2005, Pfattheicher and Schindler 2015, Sylwester, Herrmann, and Bryson 2013). However, the underlying process which governs punishment behavior remains largely uninvestigated. The underlying process sheds light on the origins and mechanisms of prosocial and antisocial punishment, and helps to understand the micro-foundation of punishment behavior.

The purpose of this chapter is to fill this gap and examine the cognitive processes underlying prosocial and antisocial punishment. First, I investigate whether the cognitive process is in line with sequential sampling models. Sequential sampling models, such as drift-diffusion model (DDM), assume that decisions are made by a noisy process in which evidence is accumulated over time until the decision maker is confident enough to make the decision (Ratcliff 1978, Ratcliff and Smith 2004). It was originally developed for perceptual decision making and then also used for value-based decision making (Krajbich, Armel, and Rangel 2010). In recent years, there is increasing evidence that the DDM is a valid approximation of the algorithm used by the brain to make simple economic or social decisions (Frydman and Nave forthcoming, Hutcherson, Bushong, and Rangel 2015, Krajbich, Bartling, et al. 2015a, Krajbich, Hare, et al. 2015). But it still remains unclear whether the DDM can explain the cognitive process underlying punishment behavior since whether to punish or not is more complicated. Additional to the sequential sampling process, I also investigate whether people have a spontaneous tendency to punish others or not. Studies have shown that social decisions are the interaction results of the intuitive process and the deliberative process (Alós-Ferrer and Strack 2014, Kahneman 2003, 2011, Strack and Deutsch 2004). But whether prosocial (antisocial) behavior is more associated with the intuitive process or the deliberative process is still disputed and inconclusive (Chen and Fischbacher 2015a, Fiedler et al. 2013, Piovesan and Wengström 2009, Rand, Greene, and Nowak 2012, Rubinstein 2007, Schulz et al. 2014, Tinghog et al. 2013).

More specifically, I study the cognitive process of prosocial and antisocial punishment using behavioral and response time (RT) data from a repeated public goods game. In the contribution stage of the public goods game, subjects decide how many points they would like to contribute to the public goods. In the punishment stage, subjects decide whether to punish each of their group members or not based on the information of the own contribution, the other's contribution, the cost and the impact of punishment. To create different situations for subjects, the cost and the impact of punishment are randomly selected from four preset levels and provided to subjects together with other information in each decision situation.

The results show that, similar to perceptual and value-based decision making, both the cognitive process of prosocial and antisocial punishment is consistent with the predictions of sequential sampling models. That is, RT decreases with the strength of preference on punishment or non-punishment. This result can be corroborated by the RT difference between punishment and non-punishment. Subjects who are more likely to punish others are quicker in making punishment decisions, and subjects who are less likely to punish others are slower in making punishment decisions. In addition, the sequential sampling process starts near the boundary of non-punishment when cooperators punish defectors, and the process starts near the boundary of punishment when defectors punish cooperators. This indicates that prosocial punishment is more deliberative and antisocial punishment is more intuitive than non-punishment. We corroborate the deliberation of prosocial punishment and the intuition of antisocial punishment by out-of-sample predictions. The DDM which takes the intuition (deliberation) into account improves the ability to make out-of-sample predictions compared to the standard DDM.

Additional to the papers mentioned above, this study is highly related to the paper by Artavia-Mora, Bedi, and Rieger (2016), who use a natural field experiment to investigate whether people are more intuitively inclined to help or punish others. They find that people punish others intuitively since the punishment rate is higher under time pressure. Another paper which is related to this study is Corgnet, Espín, and Hernán-González (2015), who show that prosocial behavior at low cost is the result of the deliberative process, while antisocial behavior is more associated with the intuitive process. Different from these studies, the experiment in this chapter creates different decision situations with different strength of preferences which allow to study whether the underlying process complies with sequential sampling models. Moreover, this chapter studies whether people have a spontaneous tendency to punish others within the

framework of sequential sampling process in which the strength of preferences is controlled. This deals with the critique proposed by Krajbich, Bartling, et al. (2015a).

This chapter proceeds as follows. Section 2 describes the experimental design and procedures in detail. Section 3 reports the data in the experiment. Section 4 describes the intuitive DDM and presents the predictions given by this model. Section 5 provides the experimental evidence about the cognitive processes of prosocial and antisocial punishment. And Section 6 concludes.

3.2 Experimental Design and Procedures

3.2.1 Experimental Design

The workhorse of the experiment is a public goods game which consists of two stages: the contribution stage and the punishment stage (Fehr and Gächter 2002). At the beginning of the contribution stage, each of the four participants in a group receives an endowment of 20 points. They decide how many points to contribute to a public project and how many to keep for themselves. Available contributions are multiples of 2 points (i.e. 0, 2, 4, ..., 20). All the points that are contributed by the members of the group are multiplied by 1.6 and equally divided among the four group members.

After all group members have made their contribution decisions simultaneously, they are informed about the outcome of the first stage and then move to the punishment stage. In the punishment stage, participants decide whether to punish each of the group members one after another based on the following information: the own contribution, the other's contribution, the cost of punishment, and the impact of punishment. The cost and the impact of punishment are randomly selected from four pre-set levels with equal probability. As shown in Table 3.1, punishment is costly in the first two levels, and each punishment decision costs the punisher 2 points and reduces the punished member 2 and 8 points respectively. In the third and fourth punishment levels, punishment is costless. The public goods game repeats for 20 periods, and the group composition is randomly changed across the 20 periods using a stranger-matching design.

Table 3.1 Four punishment levels

No.	Cost of punishment : Impact of punishment
I	2 : 2
II	2 : 8
III	0 : 2
IV	0 : 8

To control subjects' prosociality, the experiment also includes a social value orientation (SVO) task (Murphy, Ackermann, and Handgraaf 2011). In the SVO task, subjects make a series of allocation decisions. In each situation, the subject chooses one of nine allocations involving varying payoffs for herself and another anonymous participant.

3.2.2 Procedures

The experiment was computerized using z-Tree (Fischbacher 2007). A total of 132 students recruited via ORSEE (Greiner 2015) took part in the experiment between October and November 2014. The experiment took place in the Lakelab at the University of Konstanz. The subjects knew that there were two experiments in the session but they received the instructions for the SVO task first. After completing the SVO task, they received the instructions for the public goods game experiment. To ensure their understanding of the experiment, subjects had to answer several control questions before SVO task and the public goods game. The experiment started only after all participants had answered all questions correctly.

In the contribution stage of the public goods game, available contributions are displayed in small boxes and subjects make their decisions by clicking on the corresponding boxes. In the punishment stage, subjects decide whether to punish or not by clicking on boxes "Deduction" or "No Deduction". If a subject cannot make her decision within 120 seconds, 3 points will be deducted.[27] There is a waiting screen between each decision and participants move to the next situation by clicking the button "Continue". The RTs of decisions in the punishment stage are recorded. At the end of the session, one of the 15 situations in the SVO task was randomly selected to be paid out in each group according to one person's choice, which was randomly determined as well. One of the 20 periods of the public goods game was also randomly chosen to be paid out. The whole experiment lasted for about 90 minutes and participants earned 13.80 Euros on average.

3.3 Basic Results from the Experiment

I first assess subjects' SVOs using their decisions in the SVO task and then classify subjects into different types according to Murphy, Ackermann, and Handgraaf (2011).

[27] Theoretically, the income of the participant can be negative. In this case, the participant will receive an income of 0. But no subject failed to make the decision within 120 seconds in the experiment.

Subjects' decisions determine a scale of SVO (SVO angle), and the SVO angel indicates subjects' concern for the welfare of others. There exists a lot of heterogeneity in subjects' SVOs. The average SVO angle is 21.69° (*sd*. = 14.15) with the lowest observation being an angle of -16.26° and the highest observation being an angle of 52.91°. Among the 132 subjects, 68 subjects are prosocial, 63 subjects are individualistic, and 1 subject is competitive (Figure 3.6 in Appendix).

In the public goods game, I am interested in how individuals who have contributed a certain amount make punishment decisions to other group members who contributed either less, the same amount, or more than them. Similar to Herrmann, Thöni, and Gächter (2008), I label the decision situation of positive deviation between the own and the other's contribution as *prosocial situation* since the target member behaved less prosocially than the punisher and rode free on the punisher's contribution. By comparing the contributions of them, I call the target member *defector* and the punisher *cooperator*. I label the decision situation of negative or zero deviation as *antisocial situation*, since the target member (*cooperator*) behaved at least as prosocially as the punisher (*defector*). Therefore, I call the punishment in prosocial situations *prosocial punishment* and the punishment in antisocial situations *antisocial punishment*.

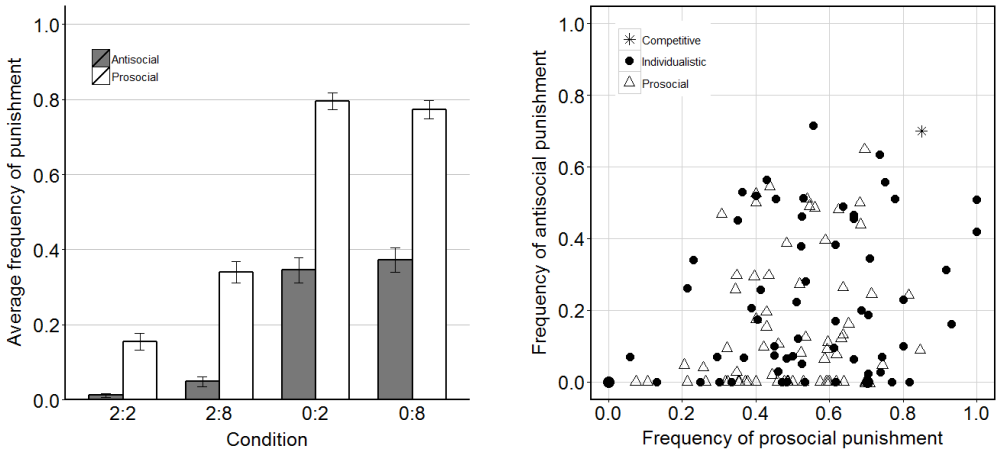


Figure 3.1 Frequency of punishment

The left panel of Figure 3.1 shows the average frequency of punishment of all subjects across the four punishment levels. Punishment behavior differs strongly across the four levels. In both prosocial and antisocial situations, the frequency of punishment increases from the punishment level 2:2 to the punishment level 0:8. The right panel of Figure 3.1 displays the frequency of prosocial and antisocial punishment at the individual level. 39 amongst 132 subjects (13 individualistic subjects and 26 prosocial subjects) who

never punish others antisocially. Some other subjects, even prosocial subjects, punish those who contribute the same or more than them as harshly as those who ride free on them.

With respect to RTs, the mean RT of decisions in the punishment stage decreases sharply at the beginning periods, following by a rather stable level throughout the remaining periods.[28] In addition, the mean RT also decreases across the three decisions within each period (Figure 3.7 in Appendix). The distributions of RTs in prosocial and antisocial situations are right skewed and the distributions of the log(RT) are more normally distributed (Figures 3.8 and 3.9 in Appendix).

3.4 The Intuitive Drift-Diffusion Model and Predictions

In this part, I first describe the intuitive DDM which takes the intuition (deliberation) into account, and then I present the predictions given by the intuitive DDM.

3.4.1 The Intuitive Drift-Diffusion Model

The intuitive DDM is an extension of the standard DDM which integrate the idea of dual process theories into the drift-diffusion process (Chen and Krajbich 2016). The standard DDM (Ratcliff 1978, Ratcliff and McKoon 2008) was originally developed to explain RTs and the accuracy of perceptual decisions in binary choice tasks, for instance, determining the right or the left side has a higher number of moving dots. Recent work has shown that the DDM can also be used to explain behavior and response time in economic or social decision making (Fehr and Rangel 2011, Krajbich, Oud, and Fehr 2014).

The basic assumption of the DDM is that incoming sensory evidence about the identity of a stimulus is noisy and time is costly. The model implements an algorithm that minimizes the decision time for a given level of accuracy (Bogacz et al. 2006). More specifically, the DDM assumes that decisions are initiated by a noisy process that accumulates relative evidence (R) over time from a starting point toward one of the two preset boundaries, as shown in Figure 3.2 (the black path). The relative evidence R follows a diffusion process:

[28] In one session, after period 16, the experimental program was stuck for a few seconds. Therefore, the mean RT in Period 17 is longer than the previous and the subsequent periods.

$$R_{t+1} = R_t + v + s_t,$$

where v is the drift rate which indicates the speed of information uptake. The drift rate represents the strength of incoming sensory information for a decision situation. When the discriminability between the two options is high, the drift rate is large; if instead, the two options are difficult to discriminate, then the drift rate is small. s represents the standard deviation of mean-zero Gaussian distributed noise. A choice is made once the relative evidence R reaches one of the two boundaries, where we normalize the lower boundary to zero and the upper boundary to a constant, a. Another parameter in the model is the non-decision time, t_0, which denotes the time required for encoding the stimulus and decision execution. The starting point of the dynamic process $z = R_0$ represents the initial point of the drift-diffusion process. The standard DDM assumes that the process starts from the middle of the two boundaries, that is, $R_0 = a/2$.

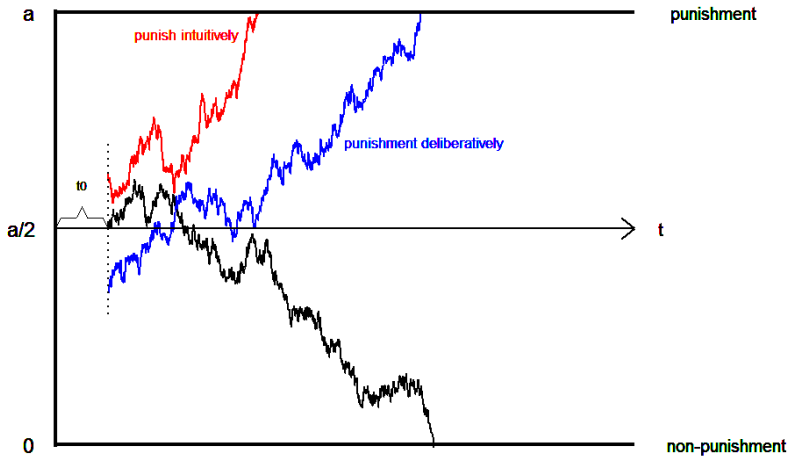


Figure 3.2 A graphical illustration of the intuitive DDM
0 and *a* are the boundaries for punishment and non-punishment. t_0 is the non-decision time, which denotes the time required for encoding the stimulus and decision execution. The red path represents the evolution of the relative evidence (R) for people who are intuitively in punishment; the blue path represents the evolution of the relative evidence (R) for people who are intuitively in non-punishment; and the black path represents the evolution of the relative evidence (R) which starts exactly from the middle of the two boundaries, as assumed by the standard DDM. When the relative evidence (R) reaches one of the boundaries, the process terminates and a decision is executed.

Different from the standard DDM, the intuitive DDM maps the intuition into the starting point of the dynamic process. The intuition on punishment makes the starting point deviate from the middle of the two boundaries and shifts that towards the boundary of punishment. And the intuition of non-punishment (or the deliberation of punishment) shifts the starting point towards the boundary of non-punishment. As shown in Figure 3.2, the red path represents the sequential sampling process for people who are

intuitively in punishment. The blue path represents the process for people who are intuitively in non-punishment.

3.4.2 Predictions

In the (intuitive) DDM, the drift rate depends on the discriminability between the two options. In economic decision making, discriminability of the decision situation can be measured by the utility difference between choice options (i.e., the strength of preferences) (Hutcherson, Bushong, and Rangel 2015, Krajbich, Armel, and Rangel 2010, Krajbich, Bartling, et al. 2015a). Thus, the larger the utility difference, the higher the drift rate, and the less time the process needs to reach one of the boundaries and make a decision. Therefore, we have the following prediction:

Prediction 1. *If the cognitive process of punishment complies with sequential sampling models, RT decreases with the strength of preferences.*

The second hypothesis is about the intuition and deliberation of punishment. The intuitive DDM captures the intuition or deliberation by the relative starting point, and Chen and Krajbich (2016) have shown that the intuition or deliberation is better encoded in the starting point than in the drift rate or both. Thus, we can study whether the punishment is more associated with the intuition or the deliberation by looking at how the starting point deviates from the middle of the two boundaries. If punishment is more intuitive, the starting point differs from the middle of the two boundaries and the process starts near the boundary of punishment. And if punishment is more deliberative, the starting point differs from the middle of the two boundaries and the process starts near the boundary of non-punishment. This leads to the second prediction:

Prediction 2. *If punishment is more intuitive, the relative starting point is greater than 0.5; and if punishment is more deliberative, the relative starting point is less than 0.5.*

3.5 Cognitive Processes underlying Punishment

In this part, I first provide evidence for the first prediction and show that the cognitive process underlying punishment is consistent with predictions by sequential

sampling models. Then I estimate the intuitive DDM and show that prosocial punishment is more deliberative while antisocial punishment is more intuitive than non-punishment. At the same time, I provide additional evidence to corroborate these two results.

3.5.1 Sequential Sampling Process of Punishment

Different from perceptual or value-based decision making, the utility difference between punishment and non-punishment in the experiment is not obvious. Therefore, I conduct logistic regressions in which the dependent variable is punishment or not and independent variables are contribution difference, impact of punishment, cost of punishment, SVO angle, as well as the interactions between contribution difference and other variables. Then I calculate the latent variable of each decision using the coefficients from the regressions. I take the latent variable of each decision as a measure of the utility difference between punishment and non-punishment (Krajbich, Bartling, et al. 2015a, Krajbich, Hare, et al. 2015, Krajbich, Oud, and Fehr 2014). To do an out-of-sample analysis, I first conduct logistic regressions on decisions of even trials in prosocial and antisocial situations (Table 3.4 in Appendix), and then I use the coefficients of the regressions to calculate the latent variable of choosing punishment in the odd trials.[29] I assume that if the latent variable is positive, subjects will choose punishment. Otherwise, subjects will choose non-punishment. If prediction 1 is correct, we would expect an inverted-U shape between RT and the utility difference. That is, if we plot RT as a function of the utility difference between punishment and non-punishment, we would find that the curve peaks at the utility difference around 0, and falls off steadily as the utility difference increases in either direction.

Figure 3.3 displays the relationship between the mean log(RT) and the utility difference. The data in the odd trials of prosocial and antisocial situations is split into 15 bins with equal size. Each dot represents one bin, and the solid line on each dot indicates the standard error of the mean log(RT). Figure 3.3 shows that, in both prosocial and antisocial situations, there exists an inverted-U relationship between RT and the utility difference. The correlation between RT and the absolute value of the utility difference is highly significant (Pearson correlation two-sided test, $r = -0.092$, $p = 1.554*10^{-4}$ for prosocial situations, and $r = -0.144$, $p = 6.469*10^{-10}$ for antisocial situations). Since the observations are not independent, I also test the correlation at the individual level. In

[29] The results of using odd trials to predict behavior and RT in the even trials (Figure 3.10, Figure 3.11 and Table 3.5 in Appendix) are essentially similar to the results of using even trials to predict behavior and RT in the odd trials.

prosocial situations, the Pearson correlation between the mean log(RT) and the absolute value of the utility difference is negative for 77 of 128 subjects, which is significantly higher than the chance level of 50% (binomial test, $p = 0.013$). In antisocial situations, the Pearson correlation is negative for 61 of 93 subjects, which is also significantly different from the chance level of 50% (binomial test, $p = 0.002$). The regression analysis also support this result. Regressions (1) and (4) in Table 3.2 show that the coefficients of the absolute value of the utility difference are significantly negative. Therefore, we can conclude that both the cognitive process of prosocial and antisocial punishment are in line with the predictions of sequential sampling models.

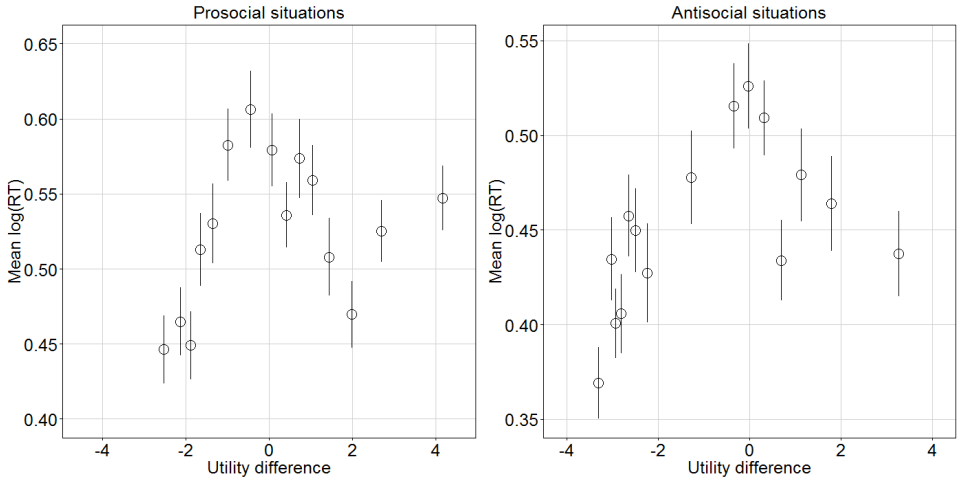


Figure 3.3 Mean log(RT) and utility difference (odd trials)

Result 1. *Both the cognitive process of prosocial and antisocial punishment complies with the sequential sampling process.*

The result above is based on the utility difference calculated at the aggregate level. Next, we corroborate this results but comparing the RTs of punishment and non-punishment decisions at the individual level. Subjects may have different utilities on punishment and non-punishment options. Subjects who place higher utility on punishment than non-punishment would be more likely to punish others and, on average, quicker in making punishment decisions. The reason is, on average, the drift rate in punishment decisions is higher than the drift rate in non-punishment decisions. Similarly, subjects who are less likely to punish others are slower in making punishment decisions. Therefore, if the cognitive process is consistent with the sequential sampling process, the RT difference between punishment and non-punishment should decreases with the frequency of punishment.

Table 3.2 OLS regressions of RTs (odd trials)

	Prosocial situations		Antisocial situations	
	(1)	(2)	(3)	(4)
Constant	0.758***	0.621***	0.740***	0.757***
	(0.027)	(0.061)	(0.029)	(0.031)
Punishment	0.087***	0.245***	0.009	0.172***
	(0.019)	(0.056)	(0.019)	(0.032)
Period	-0.009***	-0.009***	-0.008***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Decision Num.	-0.070***	-0.069***	-0.078***	-0.078***
	(0.006)	(0.006)	(0.006)	(0.006)
abs(Utility difference)	-0.025***	-0.026***	-0.031***	-0.025***
	(0.006)	(0.006)	(0.006)	(0.006)
Frequency Pun.		0.289***		-0.103*
		(0.092)		(0.056)
Punishment × Frequency Pun.		-0.329***		-0.338***
		(0.092)		(0.079)
R^2	0.134	0.164	0.117	0.154
Adj. R^2	0.132	0.161	0.115	0.152
Num. obs.	1679	1679	1813	1813
Num. groups	129	129	93	93

Notes. The dependent variable is log(RT). *Punishment* is a dummy which indicates whether the decision is a punishment decision or non-punishment decision. *Decision Num.* is the decision number within each period. The utility difference is calculated using coefficients from the regressions on the even trials.
***p < 0.01, **p < 0.05, *p < 0.1.

Similar to the utility difference between punishment and non-punishment, I calculate the frequency of punishment using the even trials in prosocial and antisocial situations, and then predict the RT difference between punishment and non-punishment decisions in the odd trials. Figure 3.4 plots the RT difference between punishment and non-punishment decisions as a function of the frequency of punishment. Each point represents one subject. The solid line is the regression line. In both prosocial and antisocial situations, the RT difference is negatively related with the frequency of punishment (Pearson correlation test, r = -0.101 and p = $4.338*10^{-6}$ for prosocial situations; r = -0.365 and p = $5.965*10^{-4}$ for antisocial situations). That is, punishment decisions are longer than non-punishment decisions for subjects who have low frequency of punishment, and punishment decisions are quicker than non-punishment decisions for subjects who have high frequency of punishment no matter whether they are prosocials or individualists. The OLS regressions in Table 3.2 also confirm this result. Regression (2) reveals that the frequency of punishment has negative effects on the RT of punishment decisions and positive effects on the RT of non-punishment decisions in prosocial situations. Regression (4) reveals that the frequency of punishment has negative effects on the RT of non-punishment decisions and stronger negative effects on the RT of punishment decisions in antisocial situations. Therefore, the RT difference decreases with the frequency of punishment in both prosocial and antisocial situations. This corroborates the first result.

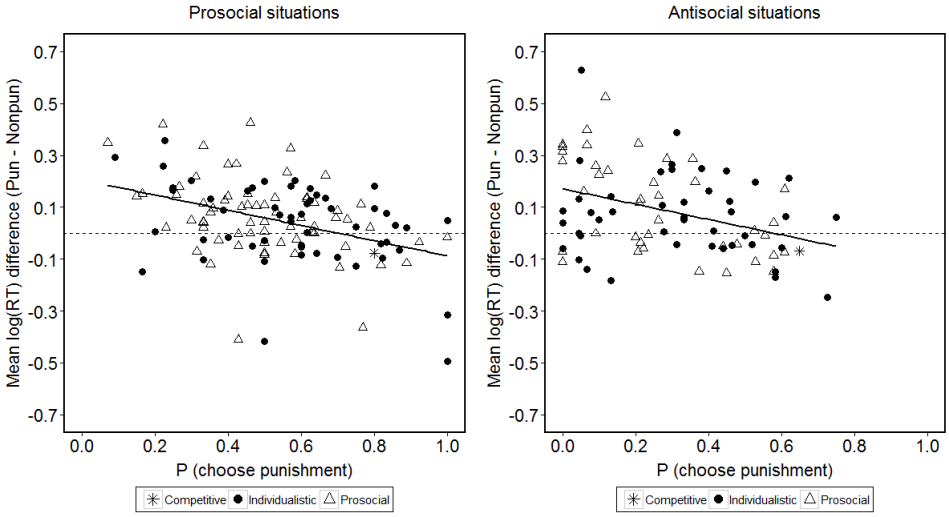


Figure 3.4 RT difference between punishment and non-punishment (odd trials)

3.5.2 Intuition and Deliberation of Punishment

In this part, we study whether prosocial punishment and antisocial punishment is more associated with the intuitive process or the deliberative process by investigating whether the relative starting point of the sequential sampling process is different from 0.5.

I estimate the intuitive DDM using the Kolmogorov-Smirnov approach of Fast DM (Voss and Voss 2007). In the estimation, the RT distribution is split into two parts, one for punishment decisions and the other for non-punishment decisions. Thus, the probability of punishment is implicitly contained by the RT distributions. To estimate the parameters of the model, a multidimensional search is conducted using a version of the simplex downhill method (Nelder and Mead 1965) to minimize the maximal vertical distance of the predicted and the empirical cumulative RT distributions. In the intuitive DDM, the drift rate depends on the stimuli of the decision situation. That is, different drift rates are assumed for different stimuli. In the estimation, we divide all decisions in prosocial situations into 40 conditions (4 punishment levels × 10 contribution differences) and all decisions in antisocial situations into 44 conditions (4 punishment levels × 11 contribution differences) based on punishment levels and contribution differences. Therefore, in the estimation we let the drift rate depend on conditions, and we assume other parameters are constant over conditions.

Table 3.3 shows the estimation results for the intuitive DDM. The relative staring point in prosocial situations (0.427) is less than 0.5. That is, the sequential sampling process in prosocial situations starts near the boundary of non-punishment, which

indicates that prosocial punishment is more deliberative. In antisocial situations, the relative starting point (0.542) is greater than 0.5, and the process in antisocial situations starts near the boundary of punishment, which indicates that antisocial punishment is more intuitive.

To check whether the deviation of the relative staring point is driven by individualistic subjects or prosocial subjects, we also estimated the intuitive DDM in which we let the relative starting point depend on subject's SVO type. Table 3.3 shows that the relative starting points for both prosocial and individualistic subjects are less than 0.5 in prosocial situations. In antisocial situations, the relative starting points for both prosocial and individualistic subjects are higher than 0.5. Interestingly, the sequential sampling process of prosocial subjects is different from that of individualistic subjects. The relative starting point for prosocial subjects are less than that of individualistic subjects in both prosocial and antisocial situations. That is, although prosocial punishment is more deliberative and antisocial punishment is more intuitive, prosocial punishment of prosocial subjects is even more deliberative than that of individualistic subjects, and antisocial punishment of individualistic subjects is even more intuitive than that of prosocial subjects.

Table 3.3 Estimation results of intuitive DDM

	Relative starting point (z)		Inter-trial variability of z (szr)	boundary (a)	non-decision time (t_0)
	Prosocial	Individualistic			
prosocial situation	0.4270 (0.0023)		0.2843 (0.0272)	3.7811 (0.0175)	1.0819 (0.0161)
	0.4011 (0.0043)	0.4491 (0.0051)	0.3467 (0.0340)	3.6877 (0.0254)	1.1911 (0.0315)
antisocial situations	0.5416 (0.0094)		0.2105 (0.0397)	3.5043 (0.0522)	0.8533 (0.0488)
	0.5243 (0.0087)	0.5443 (0.0056)	0.2831 (0.0882)	3.6459 (0.0497)	0.8129 (0.0559)

Notes. The standard errors of parameters are estimated using a jackknife procedure. The subjects who never punish others antisocially are not included in the estimation for antisocial situations. The competitive subject is also excluded for the estimation in which the relative starting point depends on the SVO type.

The deviation of the relative starting point can also be induced by two other factors. One is the frequency of punishment. The high frequency of punishment can induce a prior bias towards punishment and the low frequency of punishment can induce a prior bias towards non-punishment (Mulder et al. 2012). However, in our experiment, the mean frequency of punishment in prosocial situations is 0.518 ($sd = 0.197$), and the mean frequency of punishment in antisocial situations is 0.276 ($sd = 0.193$). The other is the response priming. This is relevant if one response is pre-activated or the default choice

(Voss et al. 2013). In the experiment, punishment and non-punishment options are randomly positioned on the right or left. Therefore, we can conclude that prosocial punishment is more deliberative and antisocial punishment is more intuitive.

Result 2. *Compared to non-punishment, prosocial punishment is more deliberative, and antisocial punishment is more intuitive.*

To confirm that the deliberation and intuition are indeed involved and matter in the sequential sampling process. I compare the within-sample fittings and the out-of-sample predictions of the intuitive DDM to those of the standard DDM. The results of within-sample fitting show that the intuitive DDM fits the data better than the standard DDM after adjusting for degrees of freedom. In prosocial situations, the BIC for the intuitive DDM is -375.294, which is less than the BIC of the standard DDM (-367.177). In antisocial situations, the BIC for the intuitive DDM is -413.346, which is also less than the BIC of the standard DDM (-404.935).[30]

In order to compare the out-of-sample predictions, I first calibrate the intuitive DDM and the standard DDM based on the even trials of prosocial and antisocial situations. Then I validate the models by using the derived estimates to predict decisions in the odd trials. This procedure uses enough data to estimate parameters reliably, but also forecasts out-of-sample to ensure models are not being overfit. More specifically, I use the estimated parameters to simulate 5000 times for each condition, and calculate the proportion of punishment as the predicted probability of punishment in that condition.

Since the number of trials are not the same in each condition of the estimation, I use Cramer's λ (Clithero 2016a, Cramer 1999, Webb et al. 2016b) to measure the predictive performance of the two models,

$$\lambda = \bar{P}^{+} - \bar{P}^{-}$$

in which $\bar{P}^{+}$ is the average predicted probability of punishment of all the trials in which subjects actually choose to punish in the experiment, and $\bar{P}^{-}$ is the average predicted probability of punishment of all the trials in which subjects actually choose not to punish.

[30] The BIC is calculated using the likelihood of the estimation using Maximum Likelihood approach of Fast-DM (Voss and Voss 2007). And the estimation results are shown in Table 3.6 in Appendix.

Thus, $\lambda \in [0,1]$ reflects the ability of the model to discriminate between choices, and measures the proportion of the total variation in choices that is explained.

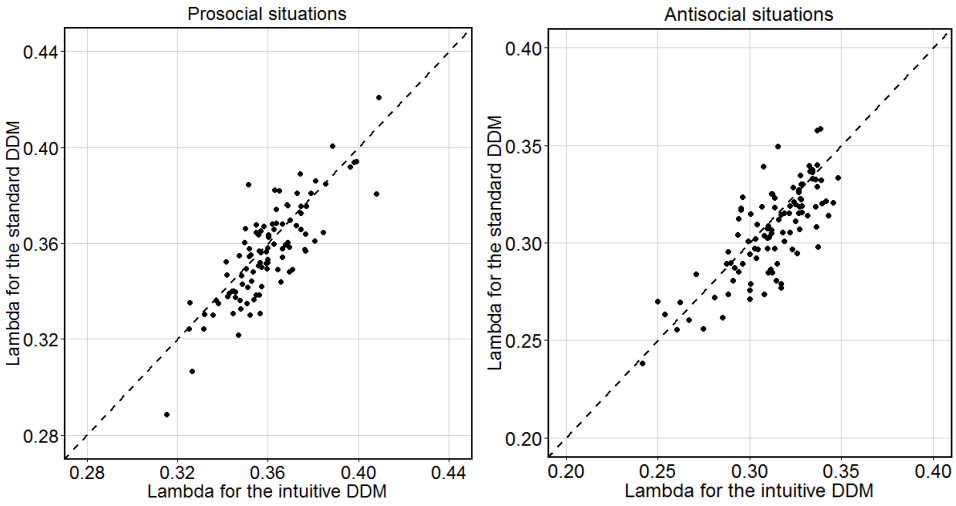


Figure 3.5 Cramer's λ for the intuitive DDM and the standard DDM

In prosocial situations, the Cramer's λ for the intuitive DDM is 0.362, which is higher than the Cramer's λ for the standard DDM (0.358). In antisocial situations, the Cramer's λ for the intuitive DDM is 0.352, which is also higher than the Cramer's λ for the standard DDM (0.341). In order to test whether the intuitive DDM significantly improves the out-of-sample predictions compared to the standard DDM, I randomly selected half of the data from prosocial situations and antisocial situations 100 times. Then I use the selected half of the data to estimate the model and predict decisions in the other half of the data. Figure 3.5 displays Cramer's λ in the 100 independent predictions. For prosocial situations, the intuitive DDM has higher λ than the standard DDM in 66 amongst 100 predictions, which is higher than the chance level (binomial test, $p = 8.95*10^{-4}$). For antisocial situations, the intuitive DDM has higher λ than the standard DDM in 68 of the 100 predictions, which is also higher than the chance level (binomial test, $p = 2.000*10^{-4}$). Therefore, we can conclude that the intuitive DDM has higher predictive power than the standard DDM, which corroborates that prosocial punishment is more associated with the deliberative process and antisocial punishment is more associated with the intuitive process.

3.6 Conclusion

This chapter studies whether the cognitive process of prosocial and antisocial punishment complies with sequential sampling models, and whether people have a

spontaneous tendency to punish others. First, the experimental evidence support the hypothesis that both the cognitive process of prosocial and antisocial punishment is in line with the sequential sampling process. The RT decreases with the strength of preference in both prosocial and antisocial situations, which is consistent with the prediction by sequential sampling models. This result can be corroborated by RT difference between punishment and non-punishment at the individual level. Second, the estimation of the intuitive DDM shows that the sequential sampling process starts near the boundary of non-punishment in prosocial situations, and the process starts near the boundary of punishment in antisocial situations. That is, prosocial punishment is more associated with the deliberative process and antisocial punishment is more associated with the intuitive process. This result can also be corroborated by the out-of-sample predictions.

This study gives a deeper understanding about the underlying process which guides prosocial and antisocial punishment behavior. Prosocial punishment and antisocial punishment share a common underlying process as perceptual and value-based decision making, but they are different on their association with the intuition process and the deliberative process. Previous studies usually treat the sequential sampling process and the intuition or deliberation of social behavior separately. This chapter shows that the intuition or deliberation is also involved in the sequential sampling process, which suggests that it is necessary to take the sequential sampling process and the intuition into account when investigating the cognitive process of social decision making. In this sense, the results of this chapter also helps to build more psychologically realistic models to describe and predict human behavior.

3.7 Appendix

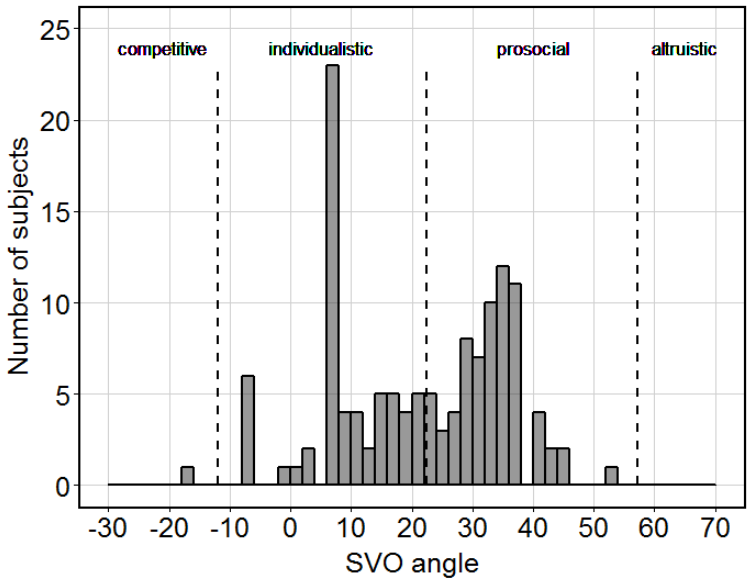


Figure 3.6 Distribution of SVO angle

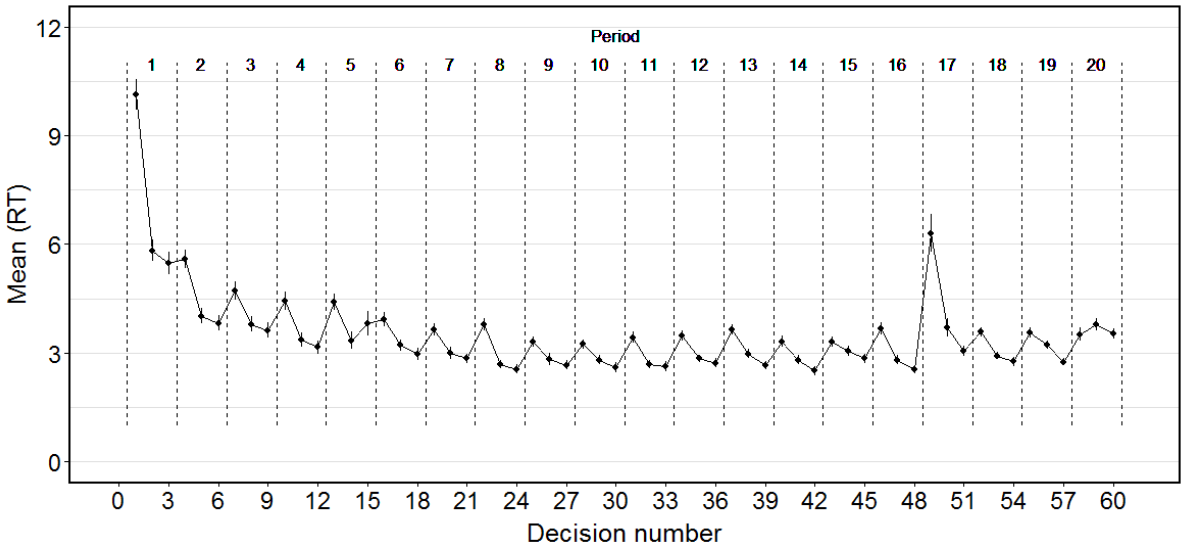


Figure 3.7 Mean RT of decisions across and within periods

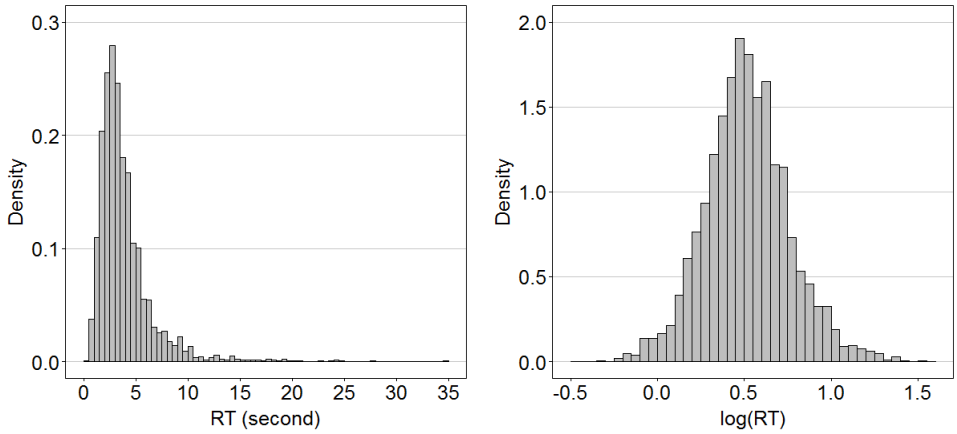


Figure 3.8 Distributions of RTs in prosocial situations

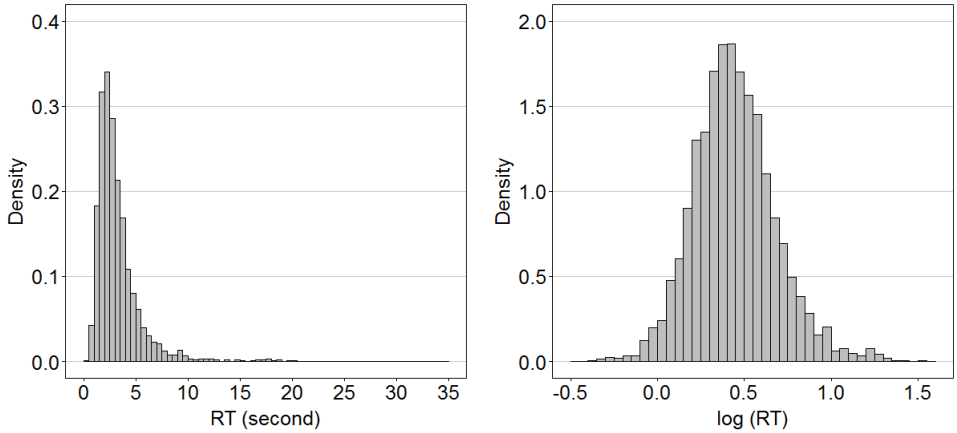


Figure 3.9 Distributions of RTs in antisocial situations

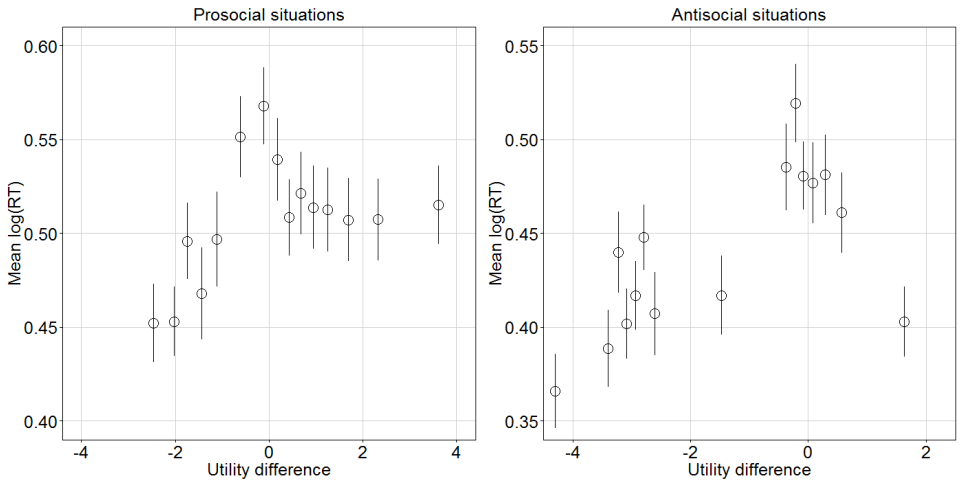


Figure 3.10 Mean log(RT) and utility difference (even trials)
The mean log(RT) is significantly correlated with the absolute value of the utility difference (Pearson correlation two-sided test, r = -0.079, p = 0.001 for prosocial situations, and r = -0.159, p = $1.467*10^{-11}$ for antisocial situations). Since these observation are not independent at the aggregate level, I also test the correlation at the individual level. In prosocial situations, the Pearson correlation between the mean log(RT) and the absolute value of the utility difference is positive for 76 of 125 subjects, which is significantly higher than the chance level of 50% (binomial test, p = 0.010). In antisocial situations, the Pearson correlation between the mean log(RT) and the absolute value of the utility difference is positive for 68 of 93 subjects, which is also significantly higher than the chance level of 50% (binomial test, p = $9.376*10^{-6}$).

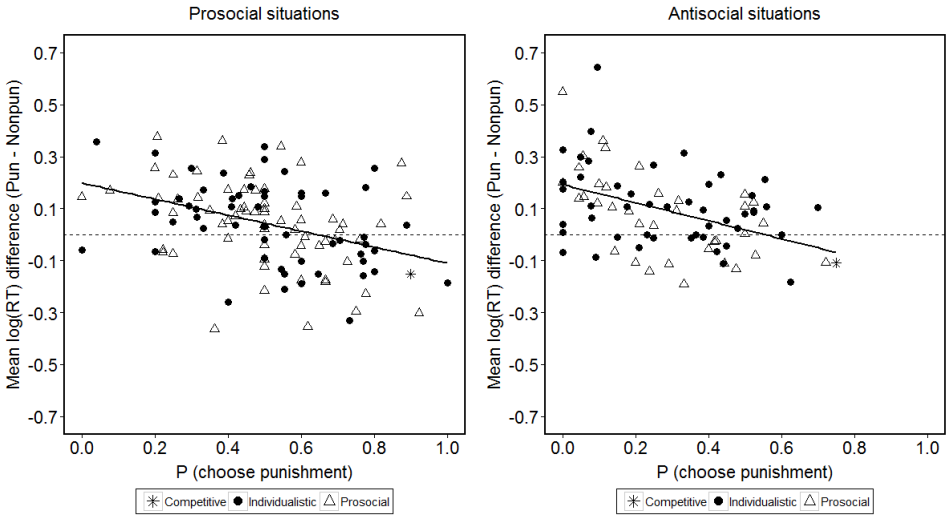


Figure 3.11 RT difference between punishment and non-punishment (even trials)
The RT difference between punishment and non-punishment decisions is negatively related with the frequency of punishment. r = -0.391, p = $7.455*10^{-6}$ for prosocial situations, and r = -0.446, p = $2.159*10^{-5}$ for antisocial situations (Pearson correlation two-sided test).

Table 3.4 Logit regressions of punishment for the even trials

	Prosocial situations	Antisocial situations
Constant	0.618	-0.403
	(0.421)	(0.317)
Contribution difference	0.107**	-0.120***
	(0.051)	(0.037)
Cost of punishment	-1.158***	-1.313***
	(0.139)	(0.190)
Impact of punishment	-0.030	-0.062**
	(0.039)	(0.029)
SVO angle	-0.028**	-0.012
	(0.013)	(0.012)
Contribution difference × Cost of punishment	-0.038**	0.068***
	(0.018)	(0.018)
Contribution difference × Impact of punishment	0.016***	-0.003
	(0.005)	(0.004)
Contribution difference × SVO angle	0.002	-0.002
	(0.001)	(0.001)
AIC	1636.151	1462.036
BIC	1679.677	1505.920
Log Likelihood	-810.076	-723.018
R^2	0.314	0.339
Num. obs.	1704	1782

Notes. The subjects who never punish others antisocially are not included in the regression of antisocial situations. The dependent variable is a dummy which indicates whether the decision is punishment (1) or non-punishment (0). The robust standard errors are clustered on participants and reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 3.5 OLS regressions of RTs (even trials)

	Prosocial situations		Antisocial situations	
	(1)	(2)	(3)	(4)
Constant	0.669***	0.560***	0.691***	0.696***
	(0.025)	(0.049)	(0.026)	(0.027)
Punishment	0.076***	0.198***	-0.020	0.167***
	(0.017)	(0.047)	(0.018)	(0.038)
Period	-0.005***	-0.004***	-0.005***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Decision Num.	-0.060***	-0.060***	-0.071***	-0.071***
	(0.006)	(0.006)	(0.006)	(0.006)
abs(Utility difference)	-0.020***	-0.022***	-0.027***	-0.025***
	(0.005)	(0.005)	(0.005)	(0.005)
Frequency Pun.		0.266***		-0.037
		(0.087)		(0.058)
Punishment × Frequency Pun.		-0.285***		-0.387***
		(0.087)		(0.092)
R^2	0.088	0.112	0.113	0.143
Adj. R^2	0.086	0.108	0.111	0.141
Num. obs.	1704	1704	1782	1782
Num. groups	130	130	93	93

Notes. The dependent variable is log(RT). *Punishment* is a dummy which indicates whether the decision is punishment decision or non-punishment decision. *Decision Num.* is the decision number within each period. The utility difference is calculated using coefficients from the regressions on the odd trials. Regressions (1) and (3) shows that the absolute value of the utility difference has negative effects on the RT in both prosocial and antisocial situations. Regression (2) and (4) reveals that the frequency of punishment has negative effects on the RT of punishment decisions, but positive effects on the RT of non-punishment decisions in prosocial situations and no significant effect on the RT of non-punishment decisions in antisocial situations. That is, the RT difference between punishment and non-punishment decreases with the frequency of punishment.
***p < 0.01, **p < 0.05, *p < 0.1.

Table 3.6 Model Estimation using MLE

		z		t_0	szr	b	BIC
		prosocial	individualistic				
Intuitive DDM	Prosocial situations	0.447		0.499	0.039	4.655	-375.294
		0.454	0.461	0.498	0.073	5.748	-383.321
	Antisocial situations	0.555		0.434	0.020	3.925	-413.346
		0.585	0.595	0.437	0.004	4.089	-421.701
Standard DDM	Prosocial situations	N/A		0.498	0.015	4.708	-367.177
	Antisocial situations	N/A		0.437	0.005	3.952	-404.935

Notes. z is the relative starting point. t_0 is the non-decision time. *szr* is the inter-trial variability of the relative starting point. *b* is the boundary. The estimation results show that the relative starting point is less than 0.5 for prosocial situations and the relative starting point is higher than 0.5 for antisocial situations, no matter whether the relative starting point depends on the subjects' SVO type or not. The BIC for the intuitive DDM is less than the BIC for the standard DDM. N/A, not applicable.

3.8 Instructions (Translated from German)

General Information

Welcome to this economic experiment.

Your decisions and possibly the decisions of other participants will affect your payoff in this experiment. Therefore, it is very important to read these instructions carefully. For the entire duration of the experiment, it is not allowed to communicate with other participants. We therefore ask you not to talk to each other. Besides, please turn off your mobile phones and tablets. If you do not understand something, please look into the experimental instructions again. If you still have questions, please give us a hand sign. We will come to you and answer your questions personally.

Today's experiment consists of two experiments, but these two experiments are independent. In the first experiment, there are 15 decision situations, and one of them will be randomly selected and paid. In the second experiment, there are 20 periods, and one of these periods will be randomly selected and paid.

During the whole experiment we do not talk about Euro but points. The income is calculated in points. The total number of points you get during the experiment is converted into Euro at the end with the exchange rate which is specified in each experiment. In addition to the points that you earn in the experiment, you will also receive 3 Euro for showing up. The exact course of the experiment is explained in detail on the following pages.

First you will get the instructions for the first experiment. When all participants have completed the first experiment, you will get the instructions for the second experiment.

Experiment A

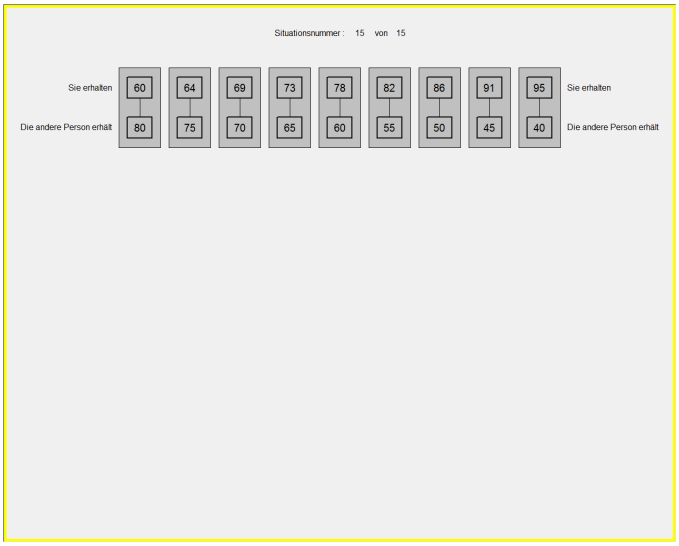
In this experiment, all participants in the lab are randomly divided into groups of two. That is, you are assigned to a group with another participant. Nobody will find out neither during or after the experiment with whom he or she is paired.

This experiment consists of 15 decision situations. In each situation, there are nine options which determine how to distribute points between yourself and the other participant in your group. You can choose one of nine options respectively. At the end of today's experiment, one of the situations will be randomly selected to be paid out according to your decision or the decision of the other participant in your group, which is randomly determined as well. The points in this experiment will be converted into Euro as follows:

1 Point = 0.05 Euro.

Decision on the Screen

The decision situation is displayed as shown in the following screenshot. You can see a possible option in each grey box: the upper number indicates how many points you receive, and the lower number indicates how many points the other participant in your group receives. For example, in the fourth distribution you get 73 points and the other participant in your group gets 65 points. You make your decision by clicking the left mouse button on one of the nine grey areas. After each decision, an intermediate screen appears. By clicking the "Next" button you will leave the screen and continue to the next decision situation.



Please answer the comprehension questions on the screen. Your answers will not affect your payment. If you have any questions, please raise your hand.

Experiment B

Overview

This experiment is repeated for 20 periods. One of the 20 periods will be randomly selected at the end and you will get the points in the corresponding period. The points you earn in this experiment will be converted into Euro as follows:

1 Point = 0.30 Euro.

In each period, all participants in the lab are randomly assigned to groups of four persons. And the group is randomly reassembled at the beginning of each period. Therefore, your group will consist of different participants in different periods. You will not know who is in your group and the other members in your group will not know who you are either.

Each period of this experiment consists of two stages. At the first stage you decide how many points you would like to contribute to a project. At the second stage you are informed about the contributions of the other group members, and you can assign deduction points to them.

The following pages describe the course of the experiment in detail.

Stage I

At the beginning of each period, each participant receives 20 points as an endowment. You decide how many of the 20 points you would like to contribute to a project. All the points that are contributed to the project by the members of your group are multiplied by 1.6 and equally divided among the four group members. Therefore, your income consists of two parts:

- The points you keep for yourself. This income is

  20 – Your contribution to the project.

- The income from the project. This income is

  1.6 * Sum of the contributions in your group / 4.

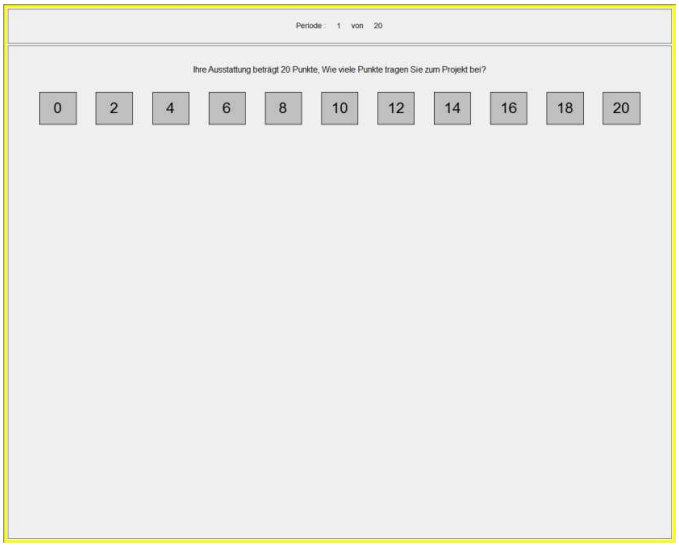
Therefore, your total income at the first stage is

(20 – Your contribution to the project) + (1.6 * Sum of the contributions in your group / 4).

The income of each group member is calculated in the same way. This means that the income *from the project* is equal for everyone in the group. For example, assume that you have contributed 12 points, and the other group members have contributed 16, 18 and 14 points. Thus, a total of 60 (=12+16+18+14) points have been contributed to the project. Then all the group members get 24 (=1.6*60/4) points from the project. Because you have contributed 12 points, you will also receive 8 (=20-12) retained points. Your total income is 32 (=8+24) points. The person who has contributed 16 points has 4 retained points and a total income of 28 (=4+24) points.

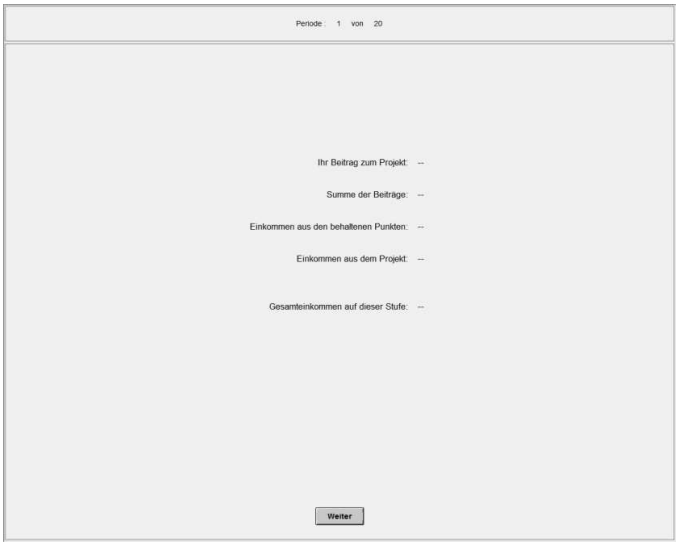
Basically, the more people contribute to the project, the higher income of the group, since the contributions are multiplied by 1.6 before being distributed. However, one may receive less than his/her own contribution because he/she can only get 0.4 times of the amount of the project.

Input on the Computer



You can contribute multiples of 2 points to the project, i.e., 0, 2, 4 to 20. These options are displayed in small grey boxes. You make your decisions by clicking the left mouse button on the corresponding boxes.

After all participants have made their decisions, you will be informed about the outcome of the period. As you can see in the following screenshot: your own contribution, the sum of all contributions, the income from the retained points, the income from the project and the total income are reported.



Stage II

At the second stage, you will learn how many points each of your group members has contributed and then decide whether to assign deduction points to them or not. If you assign deduction points, maybe it is costly for you. How many deduction points you can assign and what costs are associated with deduction is determined by chance. The following four possibilities occur with equal probabilities:

- 0:2 = You have a cost of 0 and the other group member is deducted by 2 points.
- 0:8 = You have a cost of 0 and the other group member is deducted by 8 points.
- 2:2 = You have a cost of 2 and the other group member is deducted by 2 points.
- 2:8 = You have a cost of 2 and the other group member is deducted by 8 points.

You have 120 seconds to decide for each of the three persons. If you do not decide within this time, 3 points will be deducted.

Conversely, the other group members can also assign deduction points to you.

Your income is thus composed of the following parts:

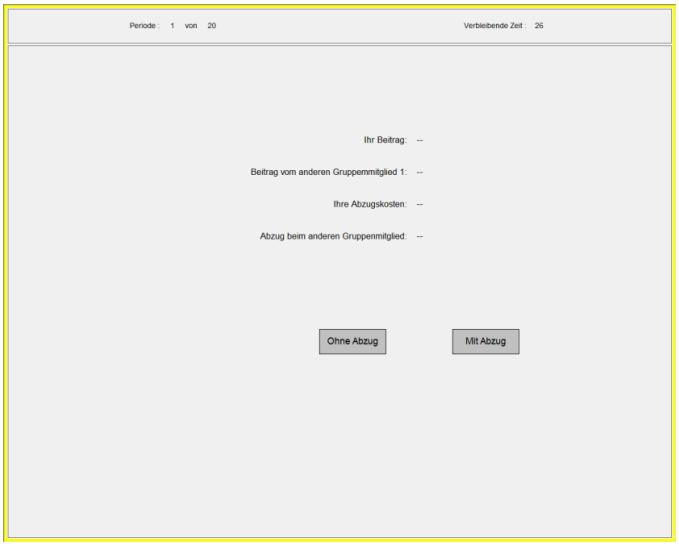
- The income from the first stage.
- The total cost of your decisions.
- The total deduction points that the others have assigned to you.
- Deduction of 3 points for every late decision.

Theoretically it would be possible that the income is negative. But in this case you will receive an income of 0.

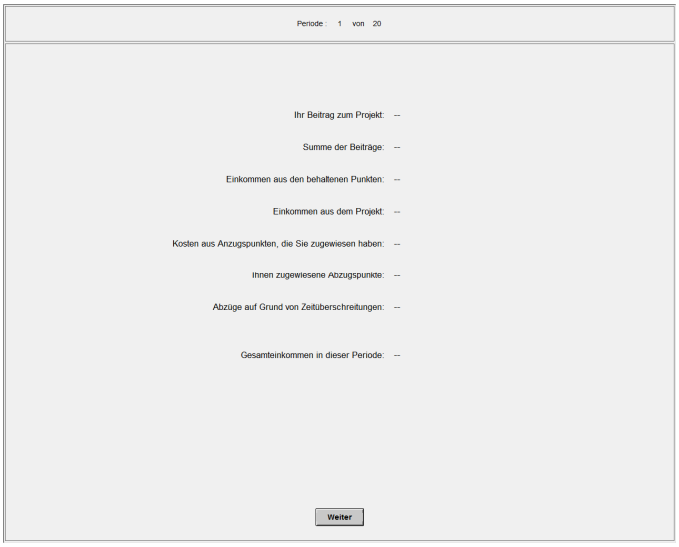
Display on the Screen

You make your decision successively on the other three group members. The order of the other three members is random.

Each decision is made on a screen similar to the following screenshot. On this screen you have the following information: your own contribution, the contribution of other group member, your cost when you assign the deduction and the number of points that are deducted from the group member. The decision is made similar to the contribution decisions. You click on the grey rectangle "Deduction" or "No Deduction" with the left mouse button to make decisions.



If all participants have made decisions on whether to assign deduction or not, you get the information listed on the following screen.



The experiment is repeated for 20 periods. After each period, all participants are assigned to a new groups randomly.

Payment

At the end of the experiment, the following information will be randomly determined:

In the first experiment:

- Which situation is paid.
- Which of the two participants' decision determines the payment in the corresponding situation.

In the second experiment:

- Which period is paid.

Now please solve the comprehensive questions on the screen. Your answers will not affect your payment.

On the bottom right of the screen, there is a button with which you can get the calculator of the computer.

If you have any questions, please contact us by raising hands.

3.9 References of Chapter 3

Alós-Ferrer, Carlos, and Fritz Strack. 2014. "From Dual Processes to Multiple Selves: Implications for Economic Behavior." *Journal of Economic Psychology* 41:1-11.

Artavia-Mora, Luis, Arjun S Bedi, and Matthias Rieger. 2016. "Intuitive Cooperation and Punishment in the Field." *IZA Working Paper*.

Bogacz, Rafal, Eric Brown, Jeff Moehlis, Philip Holmes, and Jonathan D Cohen. 2006. "The Physics of Optimal Decision Making: A Formal Analysis of Models of Performance in Two-Alternative Forced-Choice Tasks." *Psychological Review* 113 (4):700.

Bone, Jonathan E, and Nichola J Raihani. 2015. "Human Punishment Is Motivated by Both a Desire for Revenge and a Desire for Equality." *Evolution and Human Behavior* 36 (4):323-30.

Chen, Fadong, and Urs Fischbacher. 2015. "Cognitive Processes of Distributional Preferences: A Response Time Study." *TWI Working Paper*.

Chen, Fadong, and Ian Krajbich. 2016. "Sequential Sampling Process with Intuition." *Working Paper*.

Clithero, John A. 2016. "Improving out-of-Sample Predictions Using Response Times and a Model of the Decision Process." *Available at SSRN 2798459*.

Corgnet, Brice, Antonio M Espín, and Roberto Hernán-González. 2015. "The Cognitive Basis of Social Behavior: Cognitive Reflection Overrides Antisocial but Not Always Prosocial Motives." *Frontiers in Behavioral Neuroscience* 9.

Cramer, Jan Salomon. 1999. "Predictive Performance of the Binary Logit Model in Unbalanced Samples." *Journal of the Royal Statistical Society: Series D (The Statistician)* 48 (1):85-94.

Falk, Armin, Ernst Fehr, and Urs Fischbacher. 2005. "Driving Forces Behind Informal Sanctions." *Econometrica* 73 (6):2017-30.

Fehr, Ernst, and Simon Gächter. 2002. "Altruistic Punishment in Humans." *Nature* 415 (6868):137-40.

Fehr, Ernst, Karla Hoff, and Mayuresh Kshetramade. 2008. "Spite and Development." *American Economic Review* 98 (2):494-9.

Fehr, Ernst, and Antonio Rangel. 2011. "Neuroeconomic Foundations of Economic Choice—Recent Advances." *Journal of Economic Perspectives* 25 (4):3-30.

Fiedler, Susann, Andreas Glöckner, Andreas Nicklisch, and Stephan Dickert. 2013. "Social Value Orientation and Information Search in Social Dilemmas: An Eye-Tracking Analysis." *Organizational Behavior and Human Decision Processes* 120 (2):272-84.

Fischbacher, Urs. 2007. "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2):171-8.

Fowler, James H, Tim Johnson, and Oleg Smirnov. 2005. "Egalitarian Motive and Altruistic Punishment." *Nature* 433 (10.1038).

Frydman, Cary, and Gideon Nave. forthcoming. "Extrapolative Beliefs in Perceptual and Economic Decisions: Evidence of a Common Mechanism." *Management Science*.

Gächter, Simon, and Benedikt Herrmann. 2011. "The Limits of Self-Governance When Cooperators Get Punished: Experimental Evidence from Urban and Rural Russia." *European Economic Review* 55 (2):193-210.

Greiner, Ben. 2015. "Subject Pool Recruitment Procedures: Organizing Experiments with Orsee." *Journal of the Economic Science Association* 1 (1):114-25.

Gürerk, Özgür, Bernd Irlenbusch, and Bettina Rockenbach. 2006. "The Competitive Advantage of Sanctioning Institutions." *Science* 312 (5770):108-11.

Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, and Natalie Henrich. 2006. "Costly Punishment across Human Societies." *Science* 312 (5781):1767-70.

Herrmann, Benedikt, Christian Thöni, and Simon Gächter. 2008. "Antisocial Punishment across Societies." *Science* 319 (5868):1362-7.

Hutcherson, Cendri A., Benjamin Bushong, and Antonio Rangel. 2015. "A Neurocomputational Model of Altruistic Choice and Its Implications." *Neuron* 87 (2):451-62.

Kahneman, Daniel. 2003. "A Perspective on Judgment and Choice: Mapping Bounded Rationality." *American Psychologist* 58 (9):697-720.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*: Macmillan.

Krajbich, Ian, Carrie Armel, and Antonio Rangel. 2010. "Visual Fixations and the Computation and Comparison of Value in Simple Choice." *Nature Neuroscience* 13 (10):1292-8.

Krajbich, Ian, Bjorn Bartling, Todd Hare, and Ernst Fehr. 2015. "Rethinking Fast and Slow Based on a Critique of Reaction-Time Reverse Inference." *Nature Communication* 6:7455.

Krajbich, Ian, Todd Hare, Bjoern Bartling, Yosuke Morishima, and Ernst Fehr. 2015. "A Common Mechanism Underlying Food Choice and Social Decisions." *PLoS Comput Biol* 11 (10):e1004371.

Krajbich, Ian, Bastiaan Oud, and Ernst Fehr. 2014. "Benefits of Neuroeconomic Modeling: New Policy Interventions and Predictors of Preference." *American Economic Review* 104 (5):501-6.

Mulder, Martijn J, Eric-Jan Wagenmakers, Roger Ratcliff, Wouter Boekel, and Birte U Forstmann. 2012. "Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff." *Journal of Neuroscience* 32 (7):2335-43.

Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf. 2011. "Measuring Social Value Orientation." *Judgment & Decision Making* 6 (8):771-81.

Nelder, John A, and Roger Mead. 1965. "A Simplex Method for Function Minimization." *The Computer Journal* 7 (4):308-13.

Pfattheicher, Stefan, and Simon Schindler. 2015. "Understanding the Dark Side of Costly Punishment: The Impact of Individual Differences in Everyday Sadism and Existential Threat." *European Journal of Personality* 29 (4):498-505.

Piovesan, Marco, and Erik Wengström. 2009. "Fast or Fair? A Study of Response Times." *Economics Letters* 105 (2):193-6.

Rand, David G., Joshua D. Greene, and Martin A. Nowak. 2012. "Spontaneous Giving and Calculated Greed." *Nature* 489 (7416):427-30.

Ratcliff, Roger. 1978. "A Theory of Memory Retrieval." *Psychological Review* 85 (2):59-108.

Ratcliff, Roger, and Gail McKoon. 2008. "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks." *Neural Computation* 20 (4):873-922.

Ratcliff, Roger, and Philip L. Smith. 2004. "A Comparison of Sequential Sampling Models for Two-Choice Reaction Time." *Psychological Review* 111 (2):333-67.

Rubinstein, Ariel. 2007. "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal* 117 (523):1243-59.

Schulz, Jonathan F., Urs Fischbacher, Christian Thöni, and Verena Utikal. 2014. "Affect and Fairness: Dictator Games under Cognitive Load." *Journal of Economic Psychology* 41:77-87.

Strack, Fritz, and Roland Deutsch. 2004. "Reflective and Impulsive Determinants of Social Behavior." *Personality and Social Psychology Review* 8 (3):220-47.

Sylwester, Karolina, Benedikt Herrmann, and Joanna J Bryson. 2013. "Homo Homini Lupus? Explaining Antisocial Punishment." *Journal of Neuroscience, Psychology, and Economics* 6 (3):167.

Tinghog, Gustav, David Andersson, Caroline Bonn, Harald Bottiger, Camilla Josephson, Gustaf Lundgren, Daniel Vastfjall, Michael Kirchler, and Magnus Johannesson. 2013. "Intuition and Cooperation Reconsidered." *Nature* 498 (7452):E1-E2.

Voss, Andreas, Klaus Rothermund, Anne Gast, and Dirk Wentura. 2013. "Cognitive Processes in Associative and Categorical Priming: A Diffusion Model Analysis." *Journal of Experimental Psychology: General* 142 (2):536.

Voss, Andreas, and Jochen Voss. 2007. "Fast-Dm: A Free Program for Efficient Diffusion Model Analysis." *Behavior Research Methods* 39 (4):767-75.

Webb, Ryan, Paul W Glimcher, Ifat Levy, Stephanie C Lazzaro, and Robb B Rutledge. 2016. "Neural Random Utility: Relating Cardinal Neural Observables to Stochastic Choice Behaviour."

4 Response Time and Click Position: Cheap Indicators of Preferences

4.1 Introduction

There is an increasing interest in using process data to get a better understanding of economic decision making. Different methods such as response times (Rubinstein 2007), pupil dilations (Wang, Spezio, and Camerer 2010), eye movements (Reutskaja et al. 2011) and even neural activity (Smith et al. 2014) have been introduced and have shown significant correspondences between this kind of data and people's decisions. These methods vary significantly in their complexity and some of them are quite costly. We show that simple by-product data in lab or online economic experiments, such as response times and click positions, can provide information about people's preferences. With a proper design, this data can be collected at almost no cost. While response time analysis has been used frequently (for a review, see Spiliopoulos and Ortmann 2015), click positions have received attention mainly in computer science. There, click positions are analyzed routinely, in particular for the optimization of web pages (Guo, Li, and Faloutsos 2009, Liu, Dolan, and Pedersen 2010). However, these studies use the click position only to determine the choice. We study what attributes of a choice subjects click on and show that the personally attractive attributes of a choice have a higher probability to be clicked.

Specifically, we use a social value orientation (SVO) experiment in which the response times and the click positions are recorded. The SVO task consists of several decision situations, each consisting of a menu of allocations between the deciding subject and an anonymous partner. Thus, the SVO measures how much people care about their own as compared to the other player's welfare. The SVO task represents a prototypical situation in which people have to assess and trade off different attributes of an option – in the case of the SVO their own vs the other's payoff. Our results show that more prosocial subjects take longer to make their decisions in the SVO task. Concerning the click position, we find that individualistic subjects click more often on their own payoffs than on the others' payoffs, and they click more often on their own payoffs than prosocial subjects. Moreover, the response time information and the click position information are complementary in explaining subjects' SVO. Thus, response time and click position can be used as indicators of people's preferences.

Response time has increasingly been used in economic experiments (Hutcherson, Bushong, and Rangel 2015, Krajbich, Armel, and Rangel 2010, Krajbich, Oud, and Fehr 2014), in particular as an indicator of whether a decision is made intuitively or deliberatively (Achtziger and Alós-Ferrer 2014, Krajbich et al. forthcoming, Piovesan and

Wengström 2009, Rand, Greene, and Nowak 2012, Schulz et al. 2014). Most closely related to our study is Fiedler et al. (2013), who use eye-tracking technology to investigate the underlying processes of social decision making. They find that differences in SVO are accompanied by consistent differences in information search and response times. In another study by Liebrand & McClintock (1988) they find that the processing times vary systematically between different types of subjects. Specifically, cooperators and competitors are slower than altruists and individualists. There is a tradition of using the mouse to track the information search process or information processing. The most prominent example is probably MouseLab (Brocas et al. 2014, Payne, Bettman, and Johnson 1993).[31] In this environment, subjects access the information hidden behind boxes on the computer screen by moving the cursor over the boxes. Another method that uses the natural interaction is "response dynamics" (Kieslich and Hilbig 2014, Koop and Johnson 2011, Spivey, Grosjean, and Knoblich 2005). In this paradigm, the mouse response is tracked as subjects move from a central location to one of two disparately spaced options. The curvature of the mouse movement is taken as an indicator of cognitive conflicts and more curved response trajectories indicate stronger conflicts.

We confirm previous evidence that the response time correlates with subjects' preferences and find that the click position also correlates with subjects' preferences. The click position provides much less detailed information than MouseLab since it does not record the information search process, but it has the advantage that it records natural behavior while in MouseLab people need to move the mouse explicitly to acquire information. In this sense, the analysis of the click position is similar to the response dynamics paradigm, which also uses the computer interaction as an additional source of information. Our study highlights the potential benefits of recording response times and click positions in economic experiments involving multiple attribute decision making and using this data to infer people's preferences.

4.2 Experimental Design and Procedure

The experiment is built on the SVO Slider Measure by Murphy, Ackermann, and Handgraaf (2011). In this task, subjects make a series of allocation decisions. In each allocation decision, the subject chooses one out of nine allocations involving varying payoffs for herself and another anonymous participant. A screenshot of one item is shown in Figure 4.1. In this example, the subject can choose one of nine options from the most

[31] For details about MouseLab see http://www.mouselabweb.org

individualistic distribution on the right to the most altruistic distribution on the left. In the SVO task, there are 6 primary items that allow the assessment of the subject's concern for the payoff of the other player. The sum of the chosen allocations determines a vector in the plane of the own and the other player's payoff and provides a scale of SVO that indicates the concern for the welfare of others.[32] In addition, there are 9 secondary items, which makes it possible to distinguish the prosocial motives of inequality aversion and efficiency.[33] We do not refer to this distinction in our study, but we use the secondary items to check the robustness of our results.

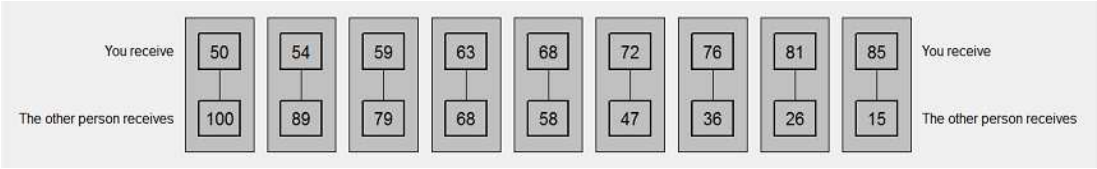


Figure 4.1 A screenshot of one item in the SVO task[34]

The subjects made their decisions by clicking on the allocation that they preferred most by clicking on one of the dark grey areas. For each decision, we recorded the subject's response time and the coordinate of the position that the subject clicked. This shows whether the subject clicked on the side of their own payoff or on the side of the payoff of the other player. Before the distributional decision, the subjects needed to click on a button that appeared in the lower region of the screen to get into the decision situation. Thus, the mouse was always positioned horizontally in the middle and vertically in the lower part of the screen immediately preceding each decision.

The experiment was computerized using z-Tree (Fischbacher 2007). A total of 132 students recruited via ORSEE (Greiner 2015) took part in the experiment between October and November 2014. The experiment took place in the Lakelab at the University of Konstanz. Upon entering the laboratory, subjects were randomly assigned into pairs. We randomized the order of the 15 items for each pair. At the end of the session, one of the 15 items was randomly selected to be paid out in each pair according to one person's choice, who was randomly determined as well. The SVO task was part of another experiment and conducted at the beginning of the session (see Instructions). The whole experiment lasted for about one and a half hours and the subjects earned 13.80 Euros on average.

[32] The definition of the scale can be found in the results section.
[33] Both the primary and the secondary items are shown in 4.5.1 in Appendix.
[34] Translated from German.

4.3 Results

4.3.1 Social Value Orientation

All choice items in the SVO measure correspond to nine equidistant points on a line in the self/other allocation plane. The six primary items correspond to the six lines that connect four points on a circle with center (50, 50) and radius 50. The points correspond to four SVO types: altruistic, prosocial, individualistic, and competitive types. The average of the chosen allocations in the primary items determines the SVO angle as follows:

$$SVO\ angle\ = \arctan\left(\frac{\overline{P_O} - 50}{\overline{P_S} - 50}\right)$$

where $\overline{P_S}$ is the average of the own payoff, and $\overline{P_O}$ is the average of the other's payoff in the chosen allocations. The SVO angle is a (practically) continuous measure of social value orientation and a higher angle indicates greater concern for the welfare of others. SVO angles of -16.26°, 0° and 61.39° indicate perfect competitiveness, perfect selfishness, and perfect altruism respectively. Given the computed SVO angles, subjects can also be categorized into four types: altruists (SVO° > 57.15°), prosocials (22.45° < SVO° < 57.15°), individualists (-12.04° < SVO° < 22.45°), and competitors (SVO° < -12.04°).[35]

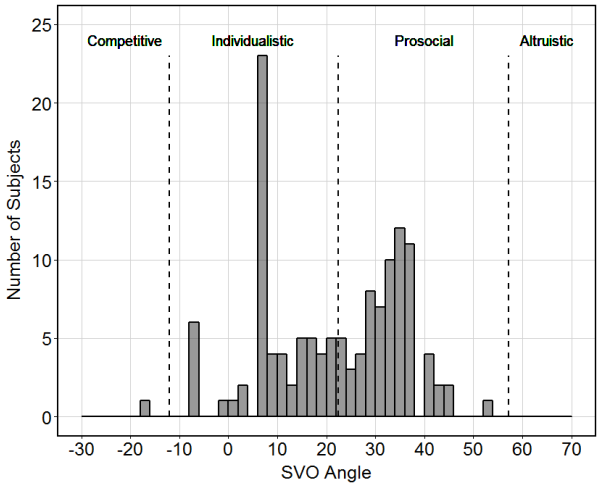


Figure 4.2 Distribution of the SVO angles

Subjects' SVO angles are reported in Figure 4.2. In line with the previous findings (Ackermann, Fleiß, and Murphy forthcoming, Murphy, Ackermann, and Handgraaf 2011), subjects vary considerably in their SVOs. In our study, the average SVO angle equals

[35] For further details, see Murphy, Ackermann, and Handgraaf (2011).

21.69° (*sd*. = 14.15°) with the lowest observation being an angle of -16.26° and the highest observation being an angle of 52.91°. Among the 132 subjects, 68 (51.52%) subjects are prosocial, 63 (47.73%) subjects are individualistic, and 1 (0.76%) subject is competitive.

4.3.2 Response Time and Preferences

Next we study the correlation between SVOs and response times. There are two reasons why we expect shorter response times for individualistic subjects in the SVO task. First, individualistic subjects have smaller information requirements because they do not care about the payoff of the other player. This has also been shown in eye tracking studies (Fiedler et al. 2013). Second, for the individualistic subjects the decision does not involve a conflict (or involves a weak conflict) between different motives while for prosocial subjects most of the decisions are associated with a stronger conflict between the own payoff and the payoff for the other player. It has been shown that the number of cognitive conflicts between motives and the degree of conflict lead to longer response times (Chen and Fischbacher 2015, Evans, Dillon, and Rand forthcoming). Therefore we expect that on average, the subjects with higher SVOs are slower in making decision in the SVO task.[36]

The distributions of response times and common logarithm of response times are shown in Appendix 4.5.2. The figures show that the distribution of response time is right skewed, but the common logarithmic transformation makes it symmetric and removes outliers.[37] The average of the logarithm of response times (ALRT) in the primary items is 0.898 with a standard deviation of 0.293, and the ALRT in the secondary items is 0.797 with a standard deviation of 0.310. The left panel of Figure 4.3 shows the relationship between the ALRT in the primary items and the SVO angle for each subject. The response times are positively correlated with the SVO angles (Spearman two-sided test, $\rho = 0.506$, $p < 10^{-9}$).[38] That is, the subjects with higher SVOs are slower in making decisions in the SVO task.

[36] If there are many subjects in each SVO type, the response time should increase from pure individualistic subjects to pure altruistic subjects, as well as from pure individualistic subjects to pure competitive subjects. That is, there should exist a U-relationship between response times and SVO angles. In the SVO slider measure, a perfectly consistent individualistic subject yields an angle between -7.82° and 7.82°. In our experiment, there is no altruistic subjects and only 1 subjects whose SVO angle is less than -7.82°. Therefore, we expect that, on average, the response time increases with the SVO angle in our study.

[37] After the logarithm transformation, all observations lie within (-2.792, 3.403) standard deviations.

[38] To test the robustness of our findings, we also conducted the analysis using the untransformed response times, which essentially led to the same results.

Result 1. *The subjects who are more prosocial are slower in making decisions in the SVO task.*

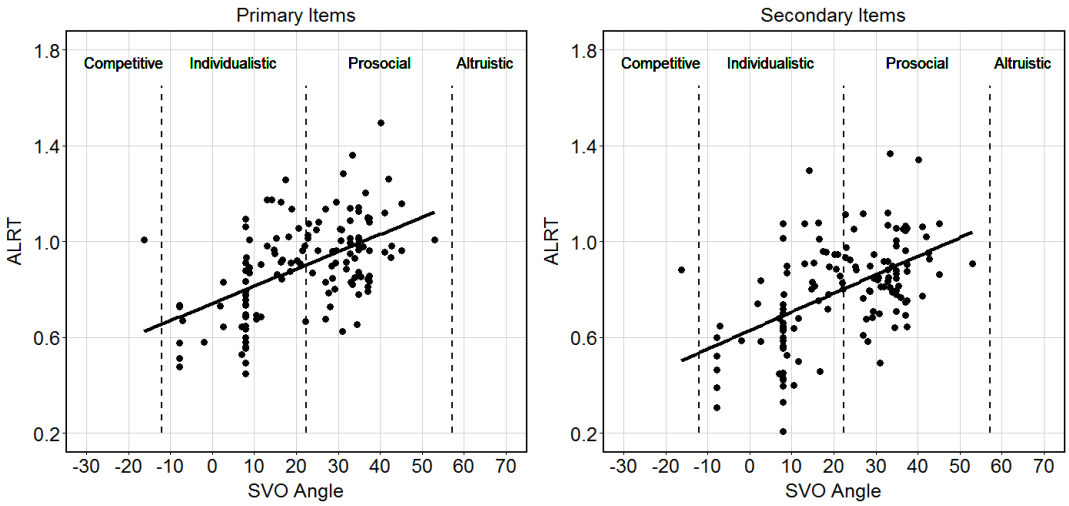


Figure 4.3 Relationship between ALRTs and SVO angles

This result is based on the data of the primary items, which we use to determine the subjects' SVOs. The secondary items are used to assess the degree of inequality aversion vs. the degree of joint gain maximization among prosocial subjects. Thus, the data of the secondary items allows us to check the robustness of our first result. The right panel of Figure 4.3 indicates that the response times in the secondary items are also positively correlated with the SVO angles (Spearman two-sided test, $\rho = 0.514$, $p < 10^{-9}$), which confirms our first finding. Moreover, regressions (1) and (2) in Table 4.1 show that the coefficients of *SVO angle* are significantly positive in both primary and secondary items for individualistic subjects, also if we control for the *decision number* and the *distance* that the cursor has to move from the button in the bottom-center to the chosen option. The coefficients of *SVO angle* are also positive for prosocial subjects (significant in the primary items, but not significant in the secondary items, $p = 0.109$). The sign of the *distance* that the cursor has to move is negative, which is counterintuitive at first glance. The reason is that the extreme decisions are easier to make. This is consistent with a view that at least some subjects were willing to take an even more extreme decision and, therefore, the decision is particularly easy. In regressions (3) and (4) of Table 4.1, we exclude the decisions in which the leftmost or the rightmost options were chosen. The results show that the distance has no significant effect on response time in the primary items and has significantly positive effect on response time in the secondary items. The horizontal extent of the decision area is larger than the vertical extent and only this

direction directly affects the decision. Nevertheless, the vertical click position could affect decision time. Because the own payoff was also on top and further away from the initial mouse position, we would expect that selfish decisions would be slower. We analyzed this question in a regression analysis and found that it is not the case. If anything, decisions in the upper part are quicker but this effect is far from significant.

Table 4.1 OLS regressions of RTs

	(1)	(2)	(3)	(4)
	Primary items	Secondary items	Primary items without extreme decisions	Secondary items without extreme decisions
SVO Angle	0.012*	0.012	0.012	0.008
	(0.007)	(0.007)	(0.012)	(0.009)
Individualist	-0.243	-0.261	-0.358	-0.099
	(0.241)	(0.256)	(0.396)	(0.337)
SVO Angle × Individualist	0.021**	0.020**	0.031*	0.016
	(0.009)	(0.009)	(0.018)	(0.013)
Decision Number	-0.023***	-0.037***	-0.024**	-0.034***
	(0.005)	(0.005)	(0.012)	(0.007)
Distance from the Bottom-Center	-0.014***	-0.012***	-0.006	0.020**
	(0.004)	(0.004)	(0.014)	(0.009)
Constant	4.239***	3.886***	3.074	-1.093
	(0.611)	(0.659)	(2.343)	(1.461)
R^2	0.192	0.220	0.142	0.081
Adj. R^2	0.187	0.217	0.123	0.073
Num. obs.	786	1179	239	549

Notes. The dependent variable is *ln(RT)*. *Individualist* is a dummy variable for SVO type (1 for individualistic type and 0 for prosocial type). *Decision Number* is the number of the decision situation which measures the subjects' experiences. *Distance from the Bottom-Center* is the linear distance that the cursor has to move from the continue button in the bottom-center to the center of the chosen option. The distance is measured using the same unit as the click position analysis. The robust standard errors are clustered on each subject and reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

4.3.3 Click Position and Preferences

Now, we turn to the correlation between the click positions and people's preferences. In the experiment, the subjects make decisions by clicking on the options they prefer most. The option contains the own as well as the other player's payoff. The click position reveals the focus of people's attention in the final stage of the decision. Our main interest is in whether the subjects click on their own or the others' payoffs. Thus, we focus on the *y*-coordinates of the click positions. We expect that individualistic subjects put most attention on their own payoffs since they care more about their own payoffs, while prosocial subjects distribute their attention among both payoffs since they care

about both their own payoffs and the others' payoffs (Fiedler et al. 2013, Funaki, Jiang, and Potters 2014).[39]

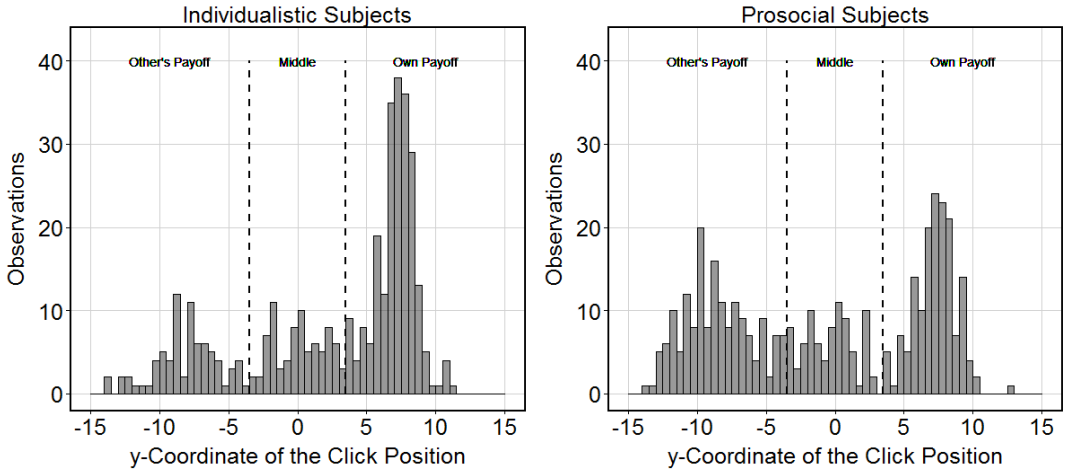


Figure 4.4 Distributions of y-coordinates of click positions in the primary items
The *y*-coordinate is normalized to the center of the clickable area.

Figure 4.4 shows histograms of the *y*-coordinates of the click positions of individual decisions for the primary items. Because the *y*-coordinate is normalized to the center of Figure 4.1, the *y*-coordinate of the own payoffs is positive and the *y*-coordinate of the other's payoffs is negative. Figure 4.4 shows that most of the click positions of individualistic subjects are located in the areas of their own payoffs, while a higher proportion of the click positions of prosocial subjects are located in the areas of the others' payoffs. For a statistical analysis, we count how many times each subject clicks on her own payoff, the other's payoff and the middle areas between her own payoff and the other's payoff. Figure 4.5 shows distributions of these frequency combinations for individualistic and prosocial subjects respectively. For example, 27 of the individualistic subjects always click in the regions of the own payoff, 4 always click in the regions of the other player's payoff and 2 always click in the middle areas. Of the 63 individualistic subjects, 41 subjects click more often on the own payoff and only 18 click more often on the other's payoff. These frequencies are significantly different (two-sided binomial test, $p = 0.004$). This is different for prosocial subjects. Of the 68 prosocial subjects, 27 subjects click more often on the own payoff and 31 click more often on the other's payoff. So prosocial subjects even click more frequently on the payoff of the other player but the frequencies are not significantly different (two-sided binomial test, $p = 0.694$). We can also directly compare the distribution of the individualistic subjects with the distribution of the prosocial

[39] There is only 1 competitive subject in our experiment. We omit this subject in the click position analysis.

subjects. Individualistic subjects click significantly more on their own payoffs than prosocial subjects, they click significantly less often on the others' payoffs than prosocial subjects, and the number of times that subjects click on the middle areas are not significantly different between the two groups (two-sided bootstrap Kolmogorov-Smirnov tests, p = 0.014, p = 0.007 and p = 0.994 respectively). That is, individualistic subjects click more often on their own payoffs than on the others' payoffs; and they click more often on their own payoffs than prosocial subjects.

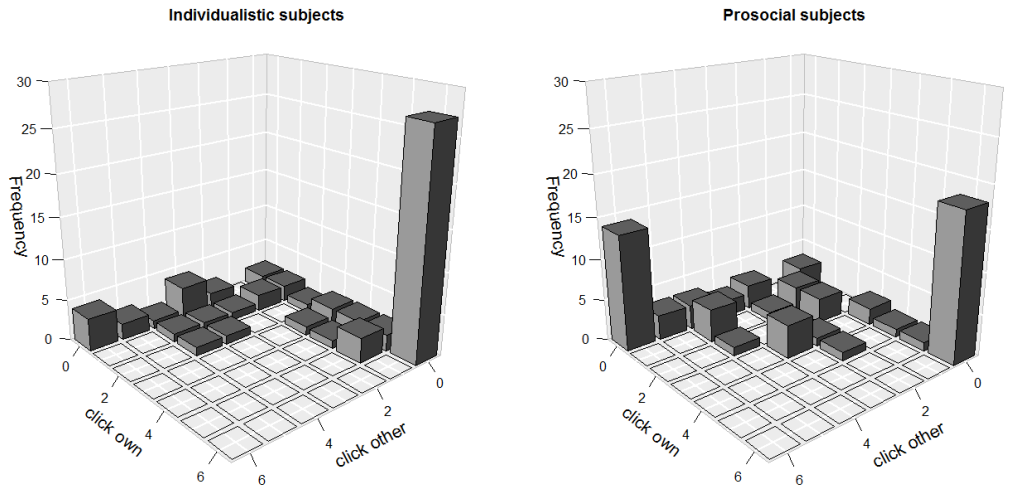


Figure 4.5 Frequency combinations of clicking own vs. the other's payoff in the primary items

Result 2. *Individualistic subjects click more often on their own payoffs than on the others' payoffs; and they click more often on their own payoffs than prosocial subjects.*

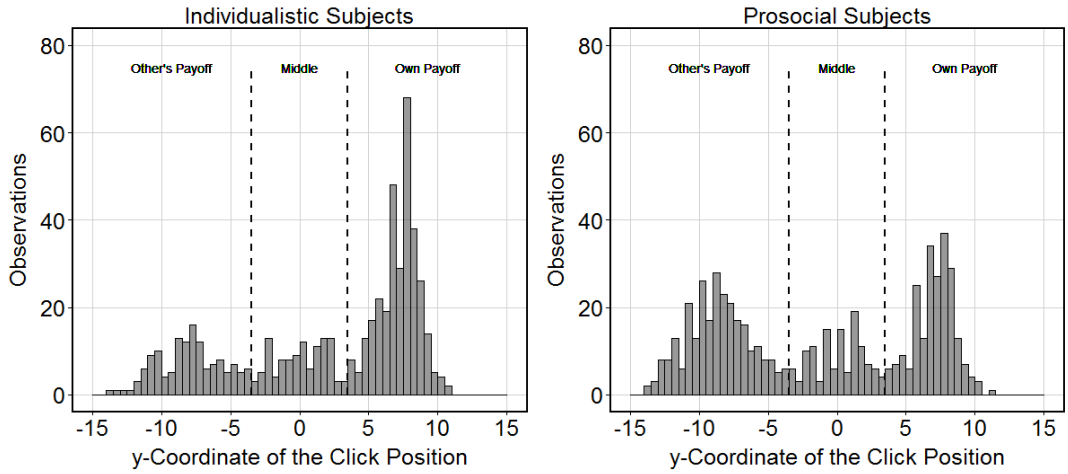


Figure 4.6 Distributions of y-coordinates of click positions in the secondary items
The y-coordinate is normalized to the center of the clickable area.

Similar to the response time analysis, we can also use the data of the secondary items to check the robustness of our result regarding the click positions. Figure 4.6 displays histograms of the *y*-coordinates of the click positions in the secondary items. As in the primary items, most of the click positions of individualistic subjects are located in the regions of their own payoffs, while a higher proportion of the click positions of prosocial subjects are located in the regions of the others' payoffs. Figure 4.7 shows distributions of the frequency combinations of the click positions in the secondary items, separately for individualistic and prosocial subjects. Also for the secondary items, the majority of individualistic subjects click more frequently on the own payoff, 39 subjects in comparison to 23 who click more frequently on the other player's payoff. These frequencies are marginally significantly different (two-sided binomial test, $p = 0.056$). This is also different for prosocial subjects. Of the 68 prosocial subjects, 24 subjects clicked more often on the own payoff and 38 clicked more often on the other's payoff. Prosocial subjects click more frequently on the payoff of the other player but the frequencies are only marginally statistically different (two-sided binomial test, $p = 0.098$). Nevertheless, the two distributions are significantly different. Individualistic subjects click significantly more often on their own payoffs than prosocial subjects, they click significantly less often on the others' payoffs than prosocial subjects, and the number of times that subjects click on the middle area are not significantly different between the two groups (two-sided bootstrap Kolmogorov-Smirnov tests, $p = 0.022$, $p = 0.003$ and $p = 0.777$ respectively). That is, individualistic subjects click more often on their own payoffs than on the others' payoffs; and they click more often on their own payoffs than prosocial subjects. Thus, the data of the secondary items confirm the robustness of our result regarding the click positions.

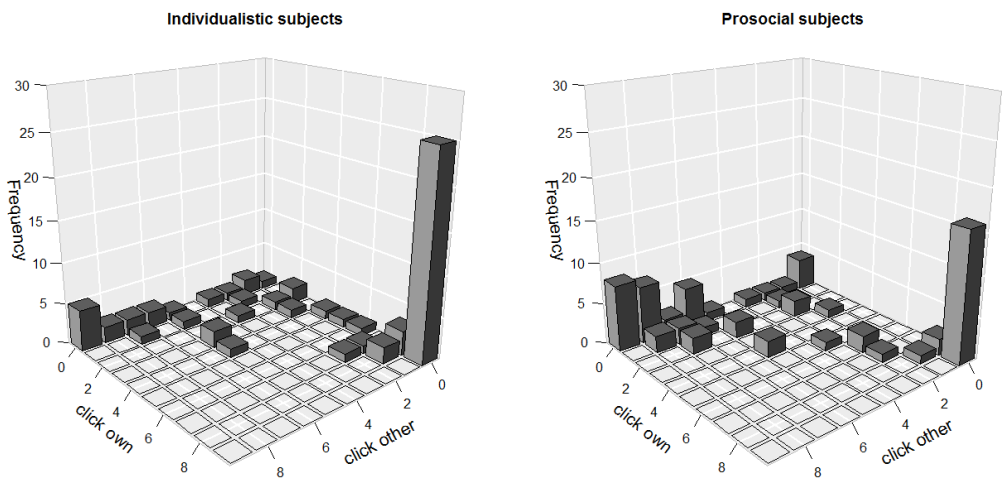


Figure 4.7 Frequency combinations of clicking own vs. the other's payoff in the secondary items

We conclude this subsection with an analysis of whether the correlation between click position pattern and SVO type is driven by the type of the subject or by the individual decision. In all items that are used to calculate the SVO angle, the other player's payoff varies linearly. This makes it possible to define an altruism measure for every decision. It equals one for the most altruistic choice, zero for the most selfish choice and is linearly interpolated for the values in between. In Table 4.2, we show how these measures affect whether the own payoff is clicked or not. The regressions show that both the type of the subject measured with the SVO angle as well as the individual decision are important but the type of the subject is more important. It has a higher R^2 as a single explanatory variable and it is more robust if we include the other variable.

Table 4.2 OLS regressions on click position in the primary items

	(1)	(2)	(3)
Constant	0.625***	1.080***	1.127***
	(0.171)	(0.296)	(0.299)
Degree of altruism of decision	-0.750***		-0.349**
	(0.188)		(0.153)
SVO angle		-0.034***	-0.030***
		(0.011)	(0.011)
Degree of altruism × SVO angle			
AIC	881.328	863.523	862.833
BIC	890.297	872.492	876.287
Log Likelihood	-438.664	-429.761	-428.417
R^2	0.018	0.037	0.041
Num. obs.	655	655	655

Notes. The dependent variable is a dummy variable of clicking on the upper part (1: own payoff) or the lower part (0: other's payoff). Item 2 is excluded (selfishness and altruism aligned). The robust standard error are clustered on each subject and reported in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

4.3.4 Comparison of Response Time and Click Position

In this part, we investigate how much of the variation in preferences can be explained by response time and click position, and whether one is redundant if the other is available. First, we calculate the classification rate when assigning a subject to a prosocial or an individualistic type using only the response time measure (ALRT) or the click position measure (number of clicks on the other's payoff). As there are 63 individualistic and 68 prosocial subjects according to the classification standard of the SVO task, we use the 0.48 (63/(63+68)) quantile of the response time measure or the click position measure as the classification boundary. The classification results in Table 4.3 show that the click position works almost equally as well as the response time. For instance, in the primary items, 40 of 63 subjects who are below or equal to the response time boundary are individualistic, and 45 of 68 subjects who are above or equal to the response time boundary are prosocial. With respect to the click position, 38 of 63 subjects who are below or equal to the click position boundary are individualistic, and 43 of 68 subjects who are above or equal to the click position boundary are prosocial.

Table 4.3 Classification based on RT or click position

		Primary items		Secondary items	
		Individualistic	Prosocial	Individualistic	Prosocial
Response time	Individualistic (63)	40 (0.63)	23 (0.37)	42 (0.67)	21 (0.33)
	Prosocial (68)	23 (0.34)	45 (0.66)	21 (0.31)	47 (0.69)
Click position	Individualistic (63)	38 (0.60)	25 (0.40)	41 (0.62)	25 (0.38)
	Prosocial (68)	25 (0.37)	43 (0.63)	29 (0.39)	46 (0.61)

Next we compare the explanatory power of response time and click position on subjects' preferences. Specifically, we use response time (ALRT) and click position information (number of clicks on the other's payoff) to predict subjects' SVO type.[40] Table 4.4 shows the regression results. The dependent variable is a dummy variable for SVO type (1 for prosocial and 0 for individualistic type). In both primary and secondary items, the response time has stronger explanatory power than the click position (R^2: 0.135 vs 0.057 in the primary items, and 0.144 vs 0.049 in the secondary items). However, if we put both the response time and the click position in the regressions simultaneously, the

[40] We can get similar results if we use response time and click position to predict the SVO angle, results are shown in Table 4.5 in Appendix.

click position adds 36.30% and 26.39% explanatory power in the primary and secondary items respectively. This is not surprising because the two measures are only weakly and insignificantly correlated (Spearman correlation test, $\rho = 0.1094$, p = 0.2134 for primary items, and $\rho = 0.1095$, p = 0.2133 for the secondary items).

Table 4.4 Regressions of RT and click position on SVO type

	Primary items			Secondary items		
	(1)	(2)	(3)	(4)	(5)	(6)
ALRT	5.111***		5.180***	4.983***		4.942***
	(1.173)		(1.222)	(1.127)		(1.155)
Num. of clicks on the other's payoff		0.265***	0.273***		0.157***	0.152**
		(0.086)	(0.096)		(0.054)	(0.060)
Constant	-4.496***	-0.398*	-5.045***	-3.889***	-0.401*	-4.310***
	(1.067)	(0.233)	(1.144)	(0.920)	(0.240)	(0.970)
AIC	160.890	175.059	153.992	159.248	176.461	154.480
BIC	166.641	180.809	162.618	164.998	182.211	163.106
Log Likelihood	-78.445	-85.529	-73.996	-77.624	-86.230	-74.240
R^2	0.135	0.057	0.184	0.144	0.049	0.182
Num. obs.	131	131	131	131	131	131

Notes. This analysis does not include the competitive subject. The dependent variable is a dummy variable for SVO type (1 for prosocial and 0 for individualistic type). *ALRT* is the average logarithm of response times for the subject. *Number of clicks on the other's payoff* is how many times the subject clicked on the other's payoff. ***p < 0.01, **p < 0.05, *p < 0.1.

Result 3. *Response time information and click position information are complementary in explaining subjects' social value orientation.*

4.4 Conclusion

This chapter explores the usefulness of response time and click position in economic experiments. The results of the SVO experiment confirm earlier results that more prosocial people have longer response times. Our novel results concern click position. We find that the click position provides information about people's social preferences, i.e. it reveals information what features (own or other's payoff) people care about. Individualistic subjects click more often on their own payoffs than on the others' payoffs and they are more likely to click on their own payoffs than prosocial subjects. Response time information and click position information are complementary in explaining subjects' preferences. Furthermore, data of the secondary items of the SVO task confirms the robustness of our results.

Our results are relevant since data about response times and click positions can be collected at almost no cost in lab or online experiments and it is wasteful to not make use

of them. Our setting investigates the relevance of different attributes for social preferences. However, the general idea can be applied to any type of choice between multi-attribute options such as consumer decisions, where options differ in price and brand attributes. Also, in these kinds of decisions response time data and click positions provide additional information about people's preferences and can be used to improve the predictive power. Process data can enhance the analysis of the decision by providing additional evidence that can be economically relevant. For example, response times can provide information about the intuitiveness or difficulty of a decision, and the click position can deliver evidence of potential motives for a decision in addition to the decision itself.

4.5 Appendix

4.5.1 The Items in the SVO Task

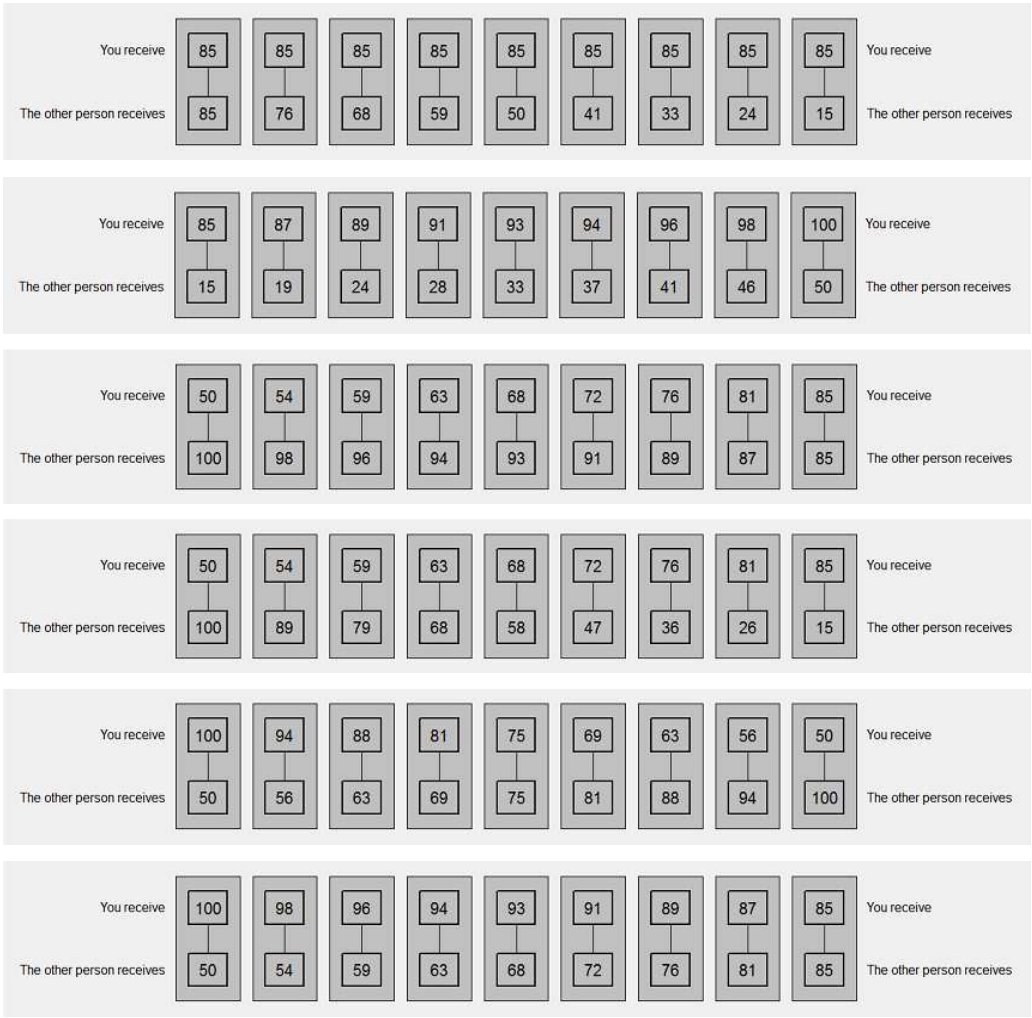


Figure 4.8 Primary items

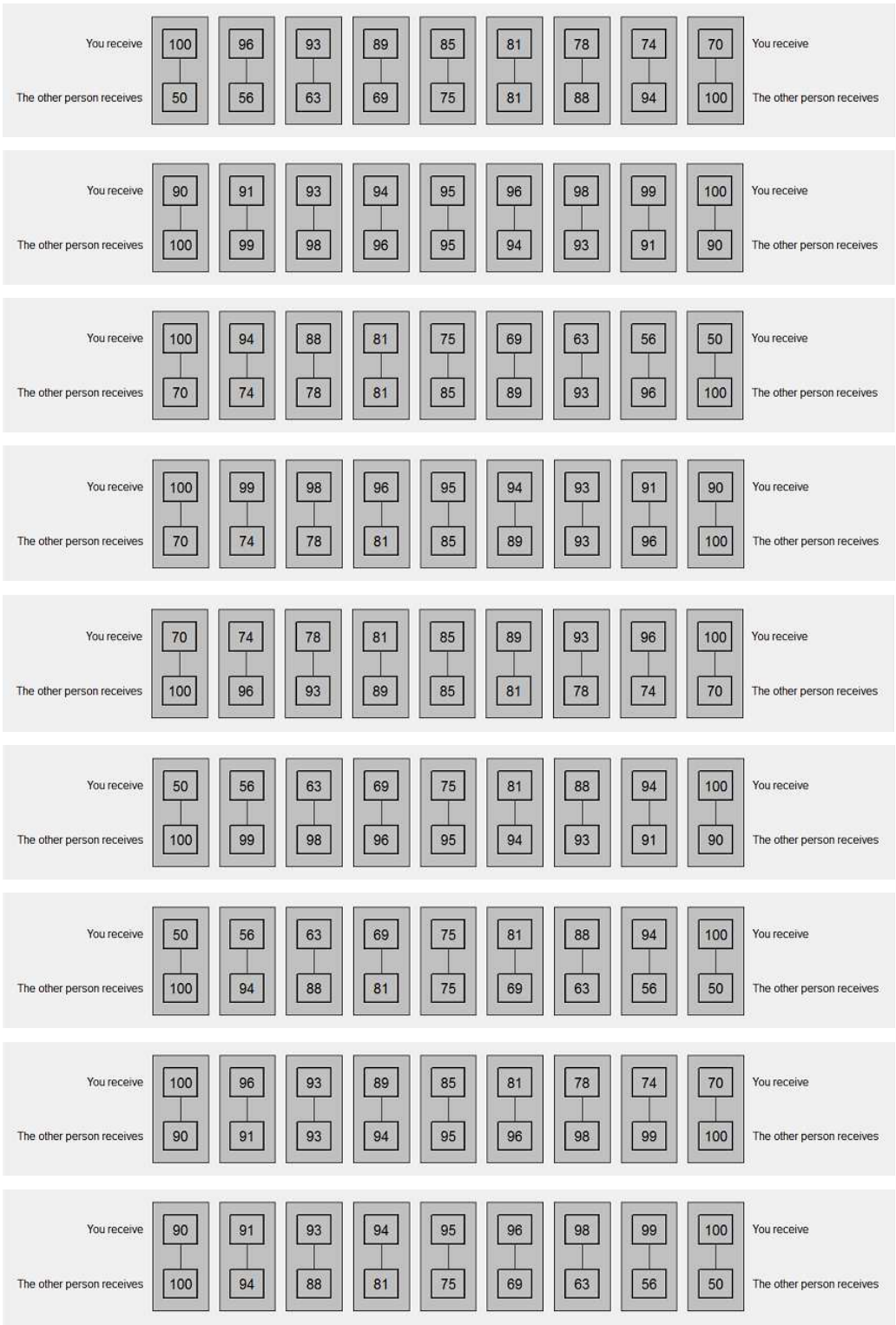


Figure 4.9 Secondary items

4.5.2 Distributions of Response Times and the Common Logarithm of Response Times

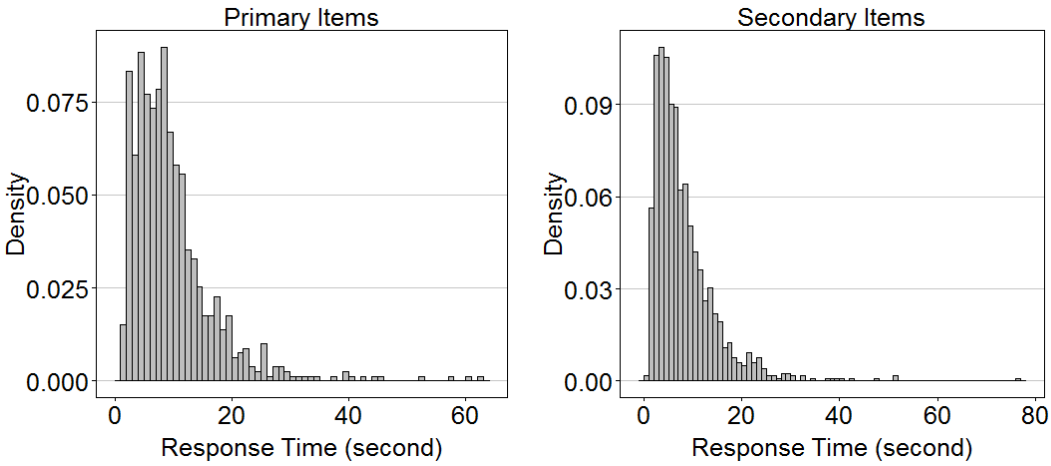


Figure 4.10 Distributions of RTs

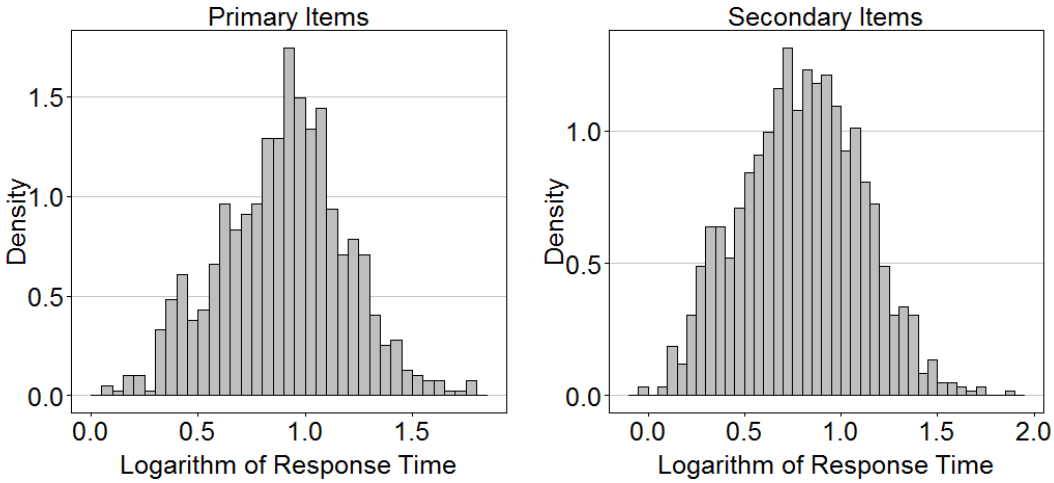


Figure 4.11 Distributions of the common logarithm of RTs

4.5.3 Compare the Explanatory Power of Response Time and Click Position on SVO Angle

Table 4.5 Regressions of RT and click position on SVO angle

	Primary items			Secondary items		
	(1)	(2)	(3)	(4)	(5)	(6)
ALRT	38.792***		37.040***	35.431***		33.771***
	(5.188)		(5.058)	(4.819)		(4.768)
Num. of clicks on the other's payoff		1.725***	1.360***		1.023***	0.738**
		(0.521)	(0.442)		(0.340)	(0.292)
Constant	-12.821***	18.763***	-13.783***	-6.242	18.789***	-7.218*
	(4.762)	(1.514)	(4.624)	(3.971)	(1.578)	(3.909)
R^2	0.302	0.078	0.350	0.295	0.066	0.329
Adj. R^2	0.297	0.071	0.340	0.290	0.058	0.318
Num. obs.	131	131	131	131	131	131

Notes. This analysis does not include the competitive subject. The dependent variable is the SVO angle. *ALRT* is the average logarithm of response time for the subject. *Number of clicks on the other's payoff* is how many times the subject clicked on the other's payoff. ***p < 0.01, **p < 0.05, *p < 0.1.

4.6 Instructions (Translated from German)

The instructions of this experiment are the same as the instructions in 3.8.

4.7 References of Chapter 4

Achtziger, Anja, and Carlos Alós-Ferrer. 2014. "Fast or Rational? A Response-Times Study of Bayesian Updating." *Management Science* 60 (4):923-38.

Ackermann, Kurt A., Jürgen Fleiß, and Ryan O. Murphy. forthcoming. "Reciprocity as an Individual Difference." *Journal of Conflict Resolution*.

Brocas, Isabelle, Juan D. Carrillo, Stephanie W. Wang, and Colin F. Camerer. 2014. "Imperfect Choice or Imperfect Attention? Understanding Strategic Thinking in Private Information Games." *Review of Economic Studies* 81 (3):944-70.

Chen, Fadong, and Urs Fischbacher. 2015. "Cognitive Processes of Distributional Preferences: A Response Time Study." *Working Papaer*.

Evans, Anthony M., Kyle D. Dillon, and David G. Rand. forthcoming. "Fast but Not Intuitive, Slow but Not Reflective: Decision Conflict Drives Reaction Times in Social Dilemmas." *Journal of Experimental Psychology: General*.

Fiedler, Susann, Andreas Glöckner, Andreas Nicklisch, and Stephan Dickert. 2013. "Social Value Orientation and Information Search in Social Dilemmas: An Eye-Tracking Analysis." *Organizational Behavior and Human Decision Processes* 120 (2):272-84.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2):171-8.

Funaki, Yukihiko, Ting Jiang, and Jan Potters. 2014. "Eye Tracking Social Preferences." *Working Paper*.

Greiner, Ben. 2015. "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." *Journal of the Economic Science Association* 1 (1):114-25.

Guo, Fan, Lei Li, and Christos Faloutsos. 2009. Tailoring Click Models to User Goals. Paper read at Proceedings of the 2009 workshop on Web Search Click Data.

Hutcherson, Cendri A., Benjamin Bushong, and Antonio Rangel. 2015. "A Neurocomputational Model of Altruistic Choice and Its Implications." *Neuron* 87 (2):451-62.

Kieslich, Pascal J., and Benjamin E. Hilbig. 2014. "Cognitive Conflict in Social Dilemmas: An Analysis of Response Dynamics." *Judgment & Decision Making* 9 (6):510-22.

Koop, Gregory J., and Joseph G. Johnson. 2011. "Response Dynamics: A New Window on the Decision Process." *Judgment & Decision Making* 6 (8):750-8.

Krajbich, Ian, Carrie Armel, and Antonio Rangel. 2010. "Visual Fixations and the Computation and Comparison of Value in Simple Choice." *Nature Neuroscience* 13 (10):1292-8.

Krajbich, Ian, Björn Bartling, Todd Hare, and Ernst Fehr. forthcoming. "Rethinking Fast and Slow Based on a Critique of Reaction-Time Reverse Inference." *Nature Communications*.

Krajbich, Ian, Bastiaan Oud, and Ernst Fehr. 2014. "Benefits of Neuroeconomic Modeling: New Policy Interventions and Predictors of Preference." *American Economic Review* 104 (5):501-6.

Liebrand, Wim BG, and Charles G McClintock. 1988. "The Ring Measure of Social Values: A Computerized Procedure for Assessing Individual Differences in Information Processing and Social Value Orientation." *European journal of personality* 2 (3):217-30.

Liu, Jiahui, Peter Dolan, and Elin Rønby Pedersen. 2010. Personalized News Recommendation Based on Click Behavior. Paper read at Proceedings of the 15th international conference on Intelligent user interfaces.

Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf. 2011. "Measuring Social Value Orientation." *Judgment & Decision Making* 6 (8):771-81.

Payne, John W., James R. Bettman, and Eric J. Johnson. 1993. *The Adaptive Decision Maker*: Cambridge University Press.

Piovesan, Marco, and Erik Wengström. 2009. "Fast or Fair? A Study of Response Times." *Economics Letters* 105 (2):193-6.

Rand, David G., Joshua D. Greene, and Martin A. Nowak. 2012. "Spontaneous Giving and Calculated Greed." *Nature* 489 (7416):427-30.

Reutskaja, Elena, Rosemarie Nagel, Colin F. Camerer, and Antonio Rangel. 2011. "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study." *American Economic Review* 101 (2):900-26.

Rubinstein, Ariel. 2007. "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal* 117 (523):1243-59.

Schulz, Jonathan F., Urs Fischbacher, Christian Thöni, and Verena Utikal. 2014. "Affect and Fairness: Dictator Games under Cognitive Load." *Journal of Economic Psychology* 41:77-87.

Smith, Alec, B. Douglas Bernheim, Colin F. Camerer, and Antonio Rangel. 2014. "Neural Activity Reveals Preferences without Choices." *American Economic Journal: Microeconomics* 6 (2):1-36.

Spiliopoulos, Leonidas, and Andreas Ortmann. 2015. "The Bcd of Response Time Analysis in Experimental Economics." *Working Paper*.

Spivey, Michael J., Marc Grosjean, and Günther Knoblich. 2005. "Continuous Attraction toward Phonological Competitors." *Proceedings of the National Academy of Sciences of the United States of America* 102 (29):10393-8.

Wang, Joseph Tao-Yi, Michael Spezio, and Colin F. Camerer. 2010. "Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games." *American Economic Review* 100 (3):984-1007.

Abgrenzung

Ich versichere hiermit, dass ich die Kapitel 1 bis 4 der vorliegenden Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

Kapitel 1: *Cognitive Processes of Distributional Preferences: A Response Time Study* entstammt einer gemeinsamen Arbeit mit Herrn Prof. Dr. Urs Fischbacher (Universität Konstanz). Es basiert auf einer Idee von Herr Fischbacher. Das experimentelle Design wurde gemeinsam erarbeitet. Programmiert und durchgeführt wurde das Experiment durch mich. Auch die Analyse stammt zum überwiegenden Teil von mir. Die erste Version des Artikels entstammt meiner Hand und wurde gemeinsam überarbeitet.

Kapitel 2: *Sequential Sampling Process with Intuition* entstammt einer gemeinsamen Arbeit mit Herrn Prof. Dr. Ian Krajbich (The Ohio State University). Es basiert auf meiner Idee. Das experimentelle Design wurde gemeinsam erarbeitet. Programmiert, durchgeführt und analysiert wurde das Experiment durch mich. Die erste Version des Artikels entstammt meiner Hand und wurde gemeinsam überarbeitet.

Kapitel 3: *Cognitive Processes of Prosocial and Antisocial Punishment* habe ich selbstständig ohne Hilfe Dritter erarbeitet.

Kapitel 4: *Response Time and Click Position: Cheap Indicators of Preferences* entstammt einer gemeinsamen Arbeit mit Herrn Prof. Dr. Urs Fischbacher. Es basiert sich auf einer Idee von Urs Fischbacher und das experimentelle Design wurde gemeinsam erarbeitet. die Programmierung, Durchführung des Experiments und die Analyse habe ich selbstständig durchgeführt. Die erste Version des Artikels entstammt meiner Hand und wurde gemeinsam überarbeitet.

Konstanz, den 20. Mar. 2017 _____

(Fadong Chen)