

## Cognitive processes underlying distributional preferences: a response time study

Fadong Chen<sup>1,2</sup> · Urs Fischbacher<sup>3,4</sup> 

### Abstract

There is ample evidence that people differ considerably in their preferences. We identify individual heterogeneity in type and strength of social preferences in a series of binary three-person dictator games. Based on this identification, we analyze response times in another series of games to investigate the cognitive processes of distributional preferences. We find that response time increases with the number of conflicts between individually relevant motives and decreases with the utility difference between choice options. The selfish motive is more intuitive for subjects who are more selfish. Our findings indicate that the sequential sampling process and the intuition of selfishness jointly produce distribution decisions, and provide an explanation for the mixed results on the correlations between response time and prosociality. Our results also show that it is important to take heterogeneity of preferences into account when investigating the cognitive processes of social decision making.

**Keywords** Distributional preferences · Cognitive process · Response time · Heterogeneity

**JEL Classification** A12 · C72 · C91 · D30 · D87

✉ Urs Fischbacher  
Urs.Fischbacher@uni-konstanz.de

Fadong Chen  
fadongchen@zju.edu.cn

<sup>1</sup> School of Management, Zhejiang University, Hangzhou 310058, China

<sup>2</sup> Neuromanagement Lab, Zhejiang University, Hangzhou 310058, China

<sup>3</sup> Department of Economics, University of Konstanz, Universitätsstr. 10, 78457 Konstanz, Germany

<sup>4</sup> Thurgau Institute of Economics, Hauptstr. 90, 8280 Kreuzlingen, Switzerland

## 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 1993). 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. Identifying the cognitive processes underlying social decision making has major implications for understanding human nature, and from a practical point of view for modeling cognitive processes in order to predict human behavior or design institutions to promote prosocial behavior (Bear and Rand 2016; Cone and Rand 2014; Krajbich et al. 2014, 2015b).

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 intuitive and deliberative processes, since intuitive processes are executed more quickly than deliberative processes (Krajbich et al. 2015a). In the domain of social preferences, this raises the question of whether some 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 processes is mixed so far (Chen and Krajbich 2018). Some studies find that social motives are more intuitive since prosocial decisions are quicker than selfish decisions and people tend to be more prosocial under time pressure or cognitive load (Cappelen et al. 2016; Cappelletti et al. 2011; Cornelissen et al. 2011; Lotito et al. 2013; Nielsen et al. 2014; Peysakhovich and Rand 2016; Rand et al. 2012; Rubinstein 2007; Schulz et al. 2014), while other studies find that the selfish motive is more intuitive (Duffy and Smith 2014; Lohse et al. 2017; Merkel and Lohse 2019; Piovesan and Wengström 2009; Tinghög et al. 2013; Verkoeijen and Bouwmeester 2014).

Unlike dual-process theories, which assume that decisions are the interaction results of intuitive and deliberative processes, sequential sampling models assume that decisions are made by a single comparison process. Specifically, sequential sampling models 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. 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 (Chen and Krajbich 2018; Dickhaut et al. 2013; Frydman and Nave 2016; Hutcherson et al. 2015; Krajbich et al. 2010, 2014, 2015a). These studies indicate that sequential sampling models can capture key patterns of choice, response time, attention, and neural response in the brain for social decision making. Intuitively, we can think of the accumulation process as an accumulation of arguments for the decision. While dual-process theories claim that decisions associated with intuitive processes are quicker than decisions associated with deliberative processes, sequential sampling models argue that the strength of preference which is based on the utility difference between choice options determines RTs, and stronger preference for one option results in shorter RTs. Particularly, the RT of selfish decisions is not significantly different from that of social decisions after controlling for the strength of preference (Krajbich et al. 2015a).

On account of the current debate and the conflicting results, we conjecture that both the sequential sampling process and the intuition of selfishness or prosociality may be part of the cognitive processes underlying social decision making. In addition, it is crucial to take heterogeneity into account, both with respect to the preferences as well as in explaining the cognitive process of social decision making. This paper investigates these topics by experimentally 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 2004; Erlei 2008; Fisman et al. 2007; Kerschbamer 2015). For instance, some people care more about efficiency, while others care more about fairness (Murphy et al. 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 selfishness.

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, selfish or social, can be considered more intuitive when controlling for the utility difference between choice options.

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-subjects analysis on the SP games shows that 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-subjects 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. Our findings demonstrate that both the sequential sampling process and the intuition of selfishness or prosociality are involved in the cognitive

processes underlying social decision making. Our findings also show that it is important to take into account the heterogeneity of the cognitive processes.

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 intuitive processes. This provides a way to reconcile the mixed results on the correlations between RT and prosociality, and also provides an explanation to resolve the debate between the dual-process and a single comparison process underlying social decision making. We argue that conflicts between individually relevant motives, strength of preference, and the intuitiveness or deliberativeness of selfishness all contribute to variations in RT. Thus, it is crucial to take these factors and the heterogeneity of preferences into account when investigating the cognitive processes of social decision making.

## 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.

### 2.1 Experimental design

The experiment consists of a series of 64 binary three-person dictator games.<sup>1</sup> 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. Thirty-two 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- $\alpha$ ) and advantageous inequality aversion (FS- $\beta$ ). 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 results in minimizing the difference between one's payoff and the highest payoff of others (Engelmann and Strobel 2004). The minimization of disadvantageous inequality corresponds to a high value of the  $\alpha$ -term in the inequality aversion model developed by Fehr and Schmidt (1999), and the minimization of advantageous inequality corresponds to a high value of the  $\beta$ -term in

---

<sup>1</sup> Supplementary Material A lists the 64 games.

**Table 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	$A_3 = 19$	$19 = B_3$	$A_3 = 13$	$20 = B_3$

the model. 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 gives examples for the two types of the binary three-person dictator games.

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 numeric as well as graphical form, in order to make them quickly accessible. The screen layout is shown in Fig. E2 in Supplementary Material E.

## 2.2 Procedural details

The experiment was computerized and conducted with the “z-Tree” software (Fischbacher 2007). We conducted four sessions in October and November 2013 at the Lakelab of the University of Konstanz. There were 105 subjects in the experiment. Subjects were recruited using the “ORSEE” online recruitment system (Greiner 2015). Each session lasted about 50 min, 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 shown the instructions (Supplementary Material E). Subjects were asked to respond to a set of control questions to ensure that they understood 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 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 to choose the left or right option on the screen. We recorded RTs on the server. The RT measure was the time between when the allocation was sent to the client and when the server received the message that the key was pressed. 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.

### 3 Behavior analyses

In this section, we first set up the finite mixture model and introduce the estimation procedure of the finite mixture model, then we present the estimation results which characterize the heterogeneity of social preferences.

#### 3.1 Model setup

In contrast to standard regression techniques, estimating behavior only at the aggregate level, finite mixture models identify types of behavior and produce type-specific sets of parameter estimates (Breitmoser 2013; Bruhin et al. 2010; Houser et al. 2004; McLachlan and Basford 1988; McLachlan and Jones 1988).<sup>2</sup> 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  $\theta_c$ —by the proportions  $\pi_c$  of the  $C$  different types in the sample. Maximizing  $\ln L(\Theta; X)$  yields the maximum likelihood estimates for the preference type parameters  $\hat{\theta}_c$  and the corresponding relative type sizes  $\hat{\pi}_c$ . Once we obtain the type-specific parameters, we can calculate the posterior probability that an individual  $i$  is of type  $c$  using Bayes' rule,  $\tau_{ic} = \frac{\hat{\pi}_c f(\hat{\theta}_c; x_i)}{\sum_{m=1}^C \hat{\pi}_m f(\hat{\theta}_m; x_i)}$ . then we classify each individual into the preference type with the highest posterior probability.

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

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

<sup>2</sup> Finite mixture models have a long history and a comprehensive description of the models can be found in McLachlan et al. (2019).

The entropy is low if all  $\tau_{ic}$  are either close to 1 or close to 0. And the entropy is high if many  $\tau_{ic}$  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  $C$ .

The TP games in the experiment allow us to deal with the heterogeneity of distributional preferences, even for selfish subjects. In each game, the subject made a binary decision according to her personal norm. We use a logistic model to capture the importance of potential social motives in the estimation of the finite mixture model. The dependent variable is the dummy variable *Decision* which indicates whether the subject chose Option A (1) or Option B (0). The independent variables refer to the non-selfish motives efficiency seeking, envy, caring about the poorest (maximin), and inequality aversion. Specifically, the independent variables are the differences between the strength of different motives and the signs of these differences. The differences of the strength of these motives between Option A and Option B are calculated as follows:

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

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

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

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

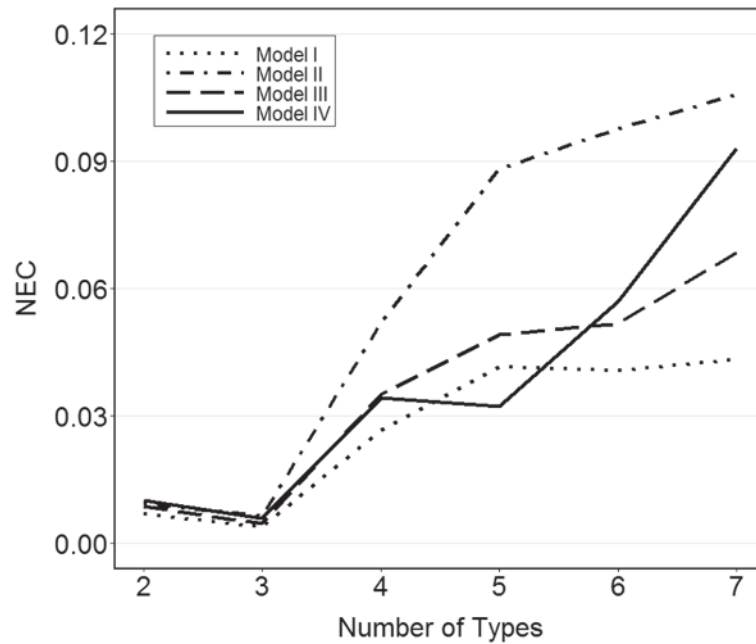
$$DiffFS-\alpha = 1/2 * \{[\max(B_1 - B_2, 0) + \max(B_3 - B_2, 0)] \\ - [\max(A_1 - A_2, 0) + \max(A_3 - A_2, 0)]\}.$$

$A_i$  and  $B_i$  are the payoffs for Player  $i$  in Option A and Option B. We interpret the signs of these differences 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. In the first model (Model I), we include all the variables of signs and differences on social motives. We have to leave out the variable  $DiffFS-\alpha$  in order to avoid collinearity.<sup>3</sup> The second model (Model II) includes the signs of all the social motives. The independent variables of the third model (Model III) are the subset of all variables which fit the data best according to the Akaike Information Criterion (AIC) at the aggregate level. And the independent variables of the fourth model (Model IV) are the subset of all variables which fit the data best according to the Bayesian Information Criterion (BIC) at the aggregate level.

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

<sup>3</sup> Due to the definitions of differences, there exists a linear relationship between  $DiffFS-\alpha$ ,  $DiffFS-\beta$  and  $DiffEfficiency$ , that is,  $DiffEfficiency = 2(DiffFS-\beta - DiffFS-\alpha)$ .



**Fig. 1** NEC for different number of types in the four models. The optimal number of types is 3 in all the four models

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

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

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

### 3.2 Behavior results

Within each of the four models, we first determine the optimal number of types, and then we classify each subject into their preference types based on the posterior probability. The results of the finite mixture analysis are robust.<sup>4</sup> All four models show that the optimal number of types equals three according to *NEC* (Fig. 1), and all four models create the same classification. In addition, all subjects can be assigned to one distinct type with high posterior probability.<sup>5</sup> 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’.

<sup>4</sup> We conducted the finite mixture analysis using the R package flexmix (Grün and Leisch 2008).

<sup>5</sup> In Model I, all subjects can be classified into their types with probabilities of greater than 0.93, and 93.3% of all subjects are classified into their types with probabilities of greater than 0.99. In Model II, only one subject is classified into her type with the probability of less than 0.90 (0.895), and 91.4% of all subjects are classified into their types with 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 probabilities of greater than 0.99. In Model IV, two subjects are classified into their types with probabilities of less than 0.90 (0.79 and 0.86), and 95.2% of all subjects are classified into their types with probabilities of greater than 0.99.



**Table 2** Logistic regression results of finite mixture model IV

	10	16	79
	Norm Type I	Norm Type II	Norm Type III
Num. subjects	10	16	79
SignEfficiency	0.613*** (0.237)	-4.612 (28.857)	2.395*** (0.176)
DiffEnvy	-0.010 (0.040)	0.286*** (0.050)	-0.048 (0.033)
DiffMaximin	0.073* (0.041)	0.383*** (0.055)	0.552*** (0.050)
SignFS- $\beta$	-0.089 (0.222)	5.024 (29.857)	0.539*** (0.119)
Constant	0.136 (0.127)	-0.297** (0.135)	-0.058 (0.108)
Num. obs	320	512	2528
AIC	1801.709		
BIC	1905.743		
Log Likelihood	-883.854		

The dependent variable is a dummy *Decision*

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

**Table 3** Pure social motives of each norm type

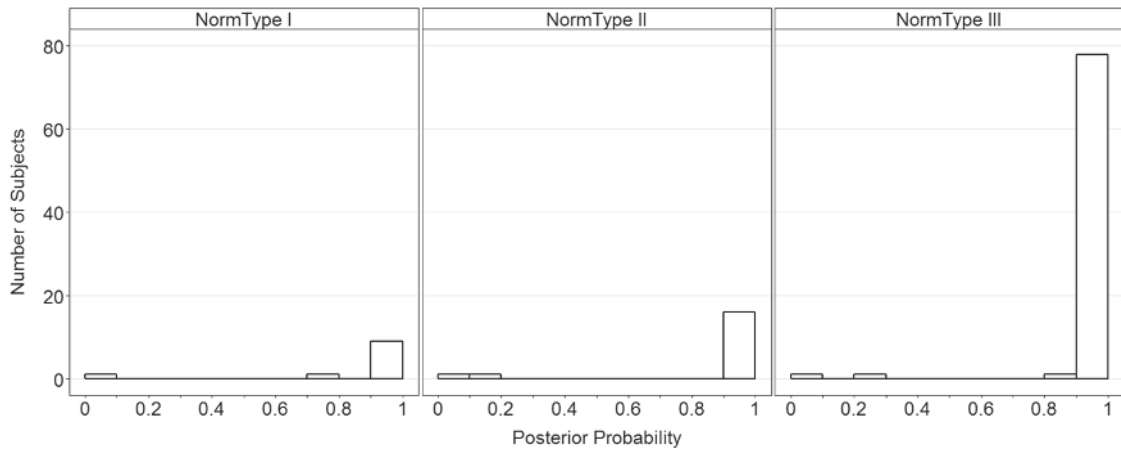
Norm type	Social motives	Number of subjects
I	Efficiency, Maximin	10
II	Envy, Maximin	16
III	Efficiency, Maximin, FS- $\beta$	79

The social motives of each type are identified according to the significance of the coefficients in Table 2

In the following, 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 model. Table 2 shows the regression results of the finite mixture model IV.<sup>6</sup> We identify subjects' social motives according to whether the coefficient of the motive is significant or not in Table 2. The relevant motives for each norm type in Model IV are summarized in Table 3.<sup>7</sup> Subjects of Norm Type I care mainly about efficiency and somewhat about maximin. This type corresponds to the model of

<sup>6</sup> The regression results of Model I, Model II, and Model III are shown in Table A2.

<sup>7</sup> One exception is the coefficient of *SignEnvy* for subjects of the first norm type in Model I. Envy is inequality aversion toward the person with the highest income. If people care about efficiency, they might like situations that envious people dislike. The relevant motives for each norm type in Model I, Model II, and Model III are shown in Table A3. Apart from some minor differences, all four models identify almost identical norm types.



**Fig. 2** The posterior probabilities that the number of subjects can be assigned to a norm type. This figure displays the posterior probabilities which are larger than 0.01. A peak at probability 1 indicates that a norm type is well separated from the other types, and no significant mass in the middle of the unit interval indicates clean classification

Charness and Rabin (2002). Subjects of Norm Type II care about envy and maximin, which corresponds to inequality aversion. Subjects of Norm Type III care about efficiency, maximin and advantageous inequality. As Type I, they are compatible with Charness and Rabin (2002), but in comparison to them, they care more about subjects who are behind. The posterior probabilities of Model IV are shown in Fig. 2.<sup>8</sup>

## 4 Process analyses

Based on the identification of heterogeneity in preferences above, we investigate the cognitive processes of distribution decisions by analyzing the RTs of SP decisions.<sup>9</sup> In this section, we first derive the hypotheses concerning the cognitive processes on the basis of sequential sampling models and dual-process theories, and then we report the results, which support the hypotheses.

### 4.1 Process hypotheses

In this part, we first focus on how the difficulty of the decision affects RT. We use the utility difference and the number of cognitive conflicts between the two options as measures of difficulty. Sequential sampling models assume that the speed of the evidence accumulation process (drift rate) represents the average of the strength of preference. The stronger the preference, the higher the drift

<sup>8</sup> The posterior probabilities for Model I, Model II, and Model III are shown in Fig. A1 in Supplementary Material A.

<sup>9</sup> The evolution and the distributions of RTs in the experiment are shown in “The evolution of mean response time” and “The distribution of response times in the second-party decisions” in Supplementary Material B.

rate. Thus, sequential sampling models predict that the probability of choosing an option is a sigmoidal function of the value (utility) closeness between the two options, and RT increases with the difficulty (measured by the value closeness between the two options) of the decision. Therefore, we expect RT to decrease with the utility difference between the two options.

Turning to the cognitive conflicts, we use the relevant motives identified before and define decision situations as conflicting (consistent) if there is a (no) conflict between these motives. We expected 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. Apart from the binary measure of conflict, we also consider *the number of conflicts*. Similar to the argument above, we also expect that RT increases with the number of conflicts. Therefore, we have hypothesized the following:

**Hypothesis 1** If the cognitive processes of distribution decisions comply with sequential sampling models, the response time will increase with the number of conflicts between individually relevant motives and decrease with the utility difference between choice options.

Dual-process theories postulate that decisions are produced by two interacting processes: the intuitive process and the deliberative process. Thus, people's reasoning does not consistently conform to a rationality norm. 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, whether selfish or social motives are more related to the intuitive process is still disputed. Our next hypothesis serves to investigate whether we can identify the selfish or social motive as generally intuitive for the subjects, i.e. that all subject have the same intuition. Evans et al. (2015) and Krajbich et al. (2015a) have shown that 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 need to control the cognitive conflict and the utility difference when investigating the relevance of intuition. Our hypothesis about the intuitiveness (deliberativeness) of selfishness is as follows:

**Hypothesis 2.1** While controlling for the cognitive conflict and the utility difference between choice options, if the selfish motive is more intuitive (deliberative) than the social motives, selfish decisions should be quicker (slower) than social decisions.

Hypothesis 2.1 is based on the assumption that subjects are homogeneous on whether their intuition favors selfishness or prosociality. Another reason for the conflicting evidence regarding whether selfish or social motives are more associated with the intuitive process is that people may differ with respect to what is the intuitive response to them. That is, it could be that the selfish motive is more

intuitive for some people and deliberative for others. Our experimental design allows us to take into account both the heterogeneity in social motives and the heterogeneity in selfishness. Taking the heterogeneity of selfishness into account, we conjecture that the strength of selfishness may influence subjects' intuition towards the selfish or social motives. Specifically, we expect that the selfish motive is more intuitive for subjects who are more selfish, and the selfish motive is more deliberative for subjects who are more prosocial. The reason is that the selfish (social) option may consistently yields better outcome for selfish (prosocial) subjects in daily life (Rand et al. 2014). Thus, we have the second hypothesis about the intuitiveness (deliberativeness) of selfishness:

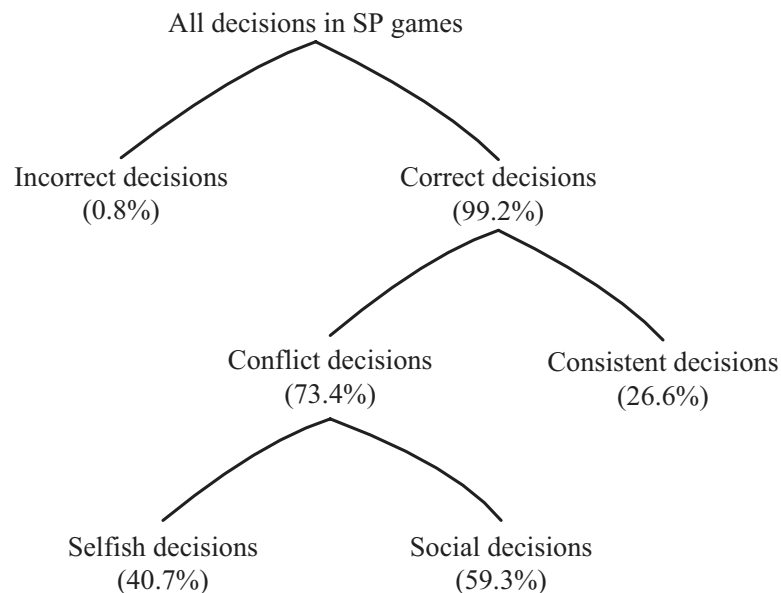
**Hypothesis 2.2** People are heterogeneous on whether their intuition favors selfishness or prosociality, the selfish motive is more intuitive (deliberative) for subjects who are more selfish (prosocial).

## 4.2 Process results

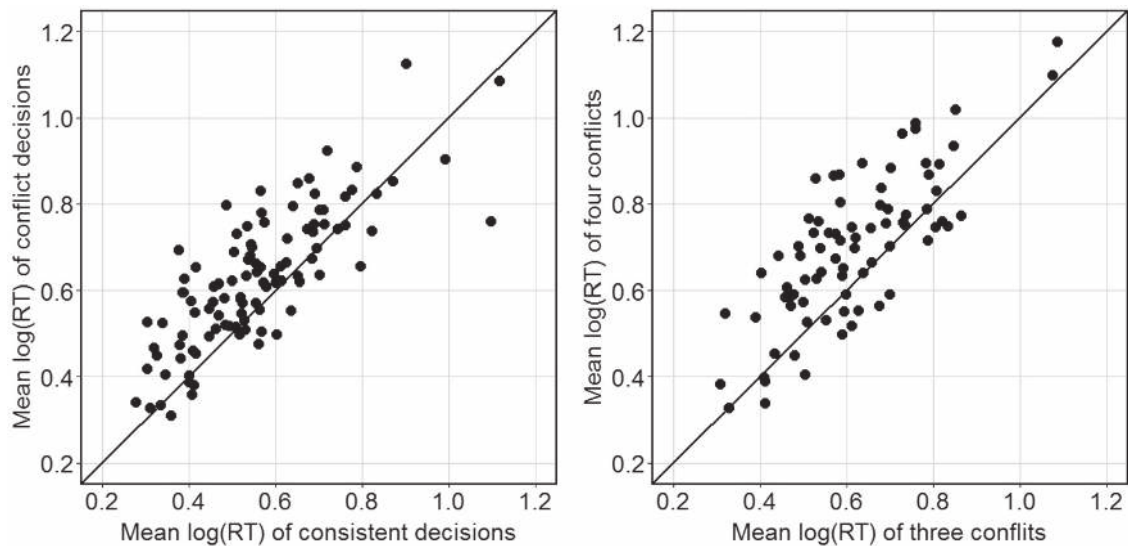
### 4.2.1 Evidence for sequential sampling models

In this part, we test Hypothesis 1 and provide evidence that the cognitive processes of distributional preferences comply with sequential sampling models.

First, we compare the RT for what we call *conflict decisions* and what we call *consistent decisions* in SP games. We define *conflict decisions* as decisions made



**Fig. 3** Classification of all decisions in the second-party dictator 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 in that decision situation, it is considered as consistent; otherwise it is considered as a conflict decision. A conflict decision is a selfish or social decision depending on which motive people are driven by in that decision. The numbers in brackets are the proportions of decisions. More details about how the decisions are classified are shown in Table 4



**Fig. 4** The RT of decisions with different number of conflicts in the second-party dictator games. Each point represents one subject. The left panel displays the mean log(RT) of conflict and consistent decisions for each subject, and the right panel displays the mean log(RT) of decisions with three conflicts and four conflicts for each subject

under circumstances in which the subject's behavior is consistent with some of her motives but in conflict with the other motives. *Consistent decisions* are decisions in circumstances where the subject's behavior is consistent with all her motives (Fig. 3). The left panel of Fig. 4 displays the mean log(RT) of conflict and consistent decisions in SP games.<sup>10</sup> 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 signed-rank test,  $p < 10^{-9}$ ).<sup>11</sup>

We get *the number of conflicts* by considering the pairwise comparisons of the individually relevant motives. The number of conflicts for each type of subject in SP decisions is shown in Table 4. 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 Fig. 4 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 signed-rank test,  $p < 10^{-7}$ ).

<sup>10</sup> To test the robustness of our results, we also conducted the analysis using untransformed RT, which essentially leads to the same results.

<sup>11</sup> Theoretically, it is difficult to put these incorrect decisions into conflict decisions or consistent decisions. On one hand, in these incorrect decisions, all the motives point to the same option, and in this sense, they can be considered to be consistent situations (decisions). On the other hand, the incorrect decisions can also be considered as conflict decisions since choices for these situations are against all the motives. Regardless of whether we consider these decisions to be conflict or consistent decisions, it does not affect the main results.

**Table 4** Classification of SP decisions

Norm type	Situation number	In-line-with					Decision type			Num. conflicts	Num. decisions
		Selfishness	Maximin	Envy	Efficiency	FS- $\beta$	Correct	Incorrect			
								Conflict	Consistent		
I	1	1	0	0	0	0	1			2	91
	2	1	1	0	0	0	1			2	30
	3	1	0	0	1	1	1			2	49
	4	1	1	1	1	1		1		0	88
	5	0	0	0	0	0			1		9
	6	0	1	1	0	0		1		2	14
	7	0	0	0	1	1		1		2	10
	8	0	1	1	1	1		1		2	29
II	1	1	0	0	0		1			2	53
	2	1	1	0	0		1			2	61
	3	1	0	1	1		1			2	24
	4	1	1	1	1			1		0	96
	5	0	0	0	0				1		15
	6	0	1	0	0			1		2	46
	7	0	0	1	1			1		2	25
	8	0	1	1	1			1		2	192

Table 4 (continued)

Norm type	Situation number	In-line-with				Decision type				Num. conflicts	Num. decisions
		Selfishness	Maximin	Envy	Efficiency	FS-β	Correct		Incorrect		
							Conflict				
							Selfish	Social			
III	1	1	0	0	0	0	1		3	220	
	2	1	0	1	0	0	1		4	287	
	3	1	1	0	1	1	1		3	0	
	4	1	1	0	0	0	1		4	101	
	5	1	0	1	1	1	1		3	0	
	6	1	0	0	0	1	1		4	0	
	7	1	1	1	1	0	1		3	79	
	8	1	1	1	1	1	1	1	0	702	
	9	0	0	0	0	0		1	4	4	
	10	0	0	1	1	0	1		3	0	
	11	0	1	0	0	1	1		4	191	
	12	0	1	0	0	0	1		3	1	
	13	0	0	1	1	1	1		4	215	
	14	0	0	0	0	1	1		3	0	
	15	0	1	1	1	0	1		4	0	
	16	0	1	1	1	1	1		3	728	

In *In-line-with* 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 and selfish motives. 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)

**Table 5** Logistic regression for the even trials of the second-party decisions

	Norm type I	Norm type II	Norm type III
Constant	−0.101 (0.190)	0.119 (0.146)	0.232*** (0.063)
Latent FMM based on TP decisions	0.577** (0.285)	0.129*** (0.042)	0.359*** (0.026)
DiffSelfish	0.298 (0.186)	0.338 (0.211)	0.620*** (0.091)
SignSelfish	1.106* (0.591)	−0.229 (0.381)	0.180 (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

The dependent variable is *Decision*. The variables *DiffSelfish* and *SignSelfish* which reflect the selfish motive are defined similar to those of social motives. The robust standard errors are clustered on subjects and reported in parentheses

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

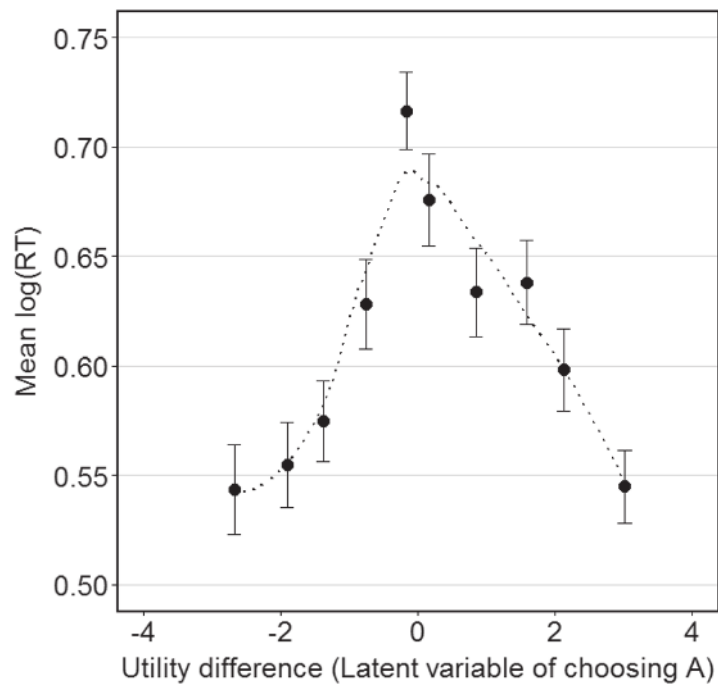
Turning to the utility difference between choice options, we construct a measure of utility difference by calculating the latent variable of choosing Option A in logistic regressions. To do an out-of-sample analysis, we conducted logistic regressions for the *even* trials of SP decisions with the dummy *Decision* as the dependent variable explained by variables reflecting the selfish incentive and the latent variables derived from TP decisions (*Latent FMM based on TP decisions* in Table 5). Table 5 shows the regression results. The latent variable *Latent FMM based on TP decisions* in Table 5 are calculated using the coefficients of the logistic regressions in Table 2.<sup>12</sup> Then we use the coefficients of the regressions in Table 5 to calculate the latent variable of choosing Option A for the *odd* trials of SP decisions and take this 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 Supplementary Material C.

Figure 5 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 are divided into 10 bins of equal size according to the utility difference. Each dot represents one bin and the solid line is the standard error of the mean log(RT). Figure 5 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

<sup>12</sup> Here we separate selfishness from the other motives when calculating the utility. We also run additional analyses in which we take both the selfish motive and social motives into account, and the results are similar. In particular, utility difference is a significant explanatory variable for response time.



**Fig. 5** The mean  $\log(\text{RT})$  and the utility difference in the odd trials of the second-party dictator games. All the data are divided into 10 bins of equal size according to the utility difference. Each dot represents one bin and the solid line is the standard error of that bin. The dotted line is the smoothing line using method “loess”



difference increases in either direction. The correlation is significant (Pearson’s correlation test,  $r = -0.205$ ,  $p < 10^{-15}$ ). Since the observations are not independent, we also test the correlation at the individual level. The Pearson correlation coefficient between the mean  $\log(\text{RT})$  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,  $p < 10^{-5}$ ).<sup>13</sup>

The dependence of RTs on the number of conflicts and the utility difference is also assessed with mixed-effects regressions. The regression results are shown in Table 6. The dependent variable is  $\log(\text{RT})$ . 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 significant 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.

**Result 1** *The cognitive processes of distributional preferences comply with the sequential sampling models. Specifically, the response time increases with the*

<sup>13</sup> We also use random samples to test the correlation between RT and the utility difference at the individual level. Specifically, we randomly select half of the data for 100 times, and then we use the selected data to predict the correlation in the other half of the data in each of the 100 times. The results show that there are 73 subjects who always have a positive correlation in all the 100 times, which is different from the chance level of 50% (two-sided binomial test,  $p < 10^{-4}$ ).

**Table 6** Mixed-effects regressions of response times on the conflict and the utility difference

	All SP decisions		The odd trials of SP decisions	
	All subjects	Type III	All subjects	Type III
	(1)	(2)	(3)	(4)
Constant	0.659*** (0.017)	0.504*** (0.038)	0.772*** (0.023)	0.709*** (0.059)
Conflict decision	0.076*** (0.008)		0.040*** (0.012)	
Decision number	-0.003*** (0.000)	-0.004*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)
Number of conflicts		0.076*** (0.010)		0.036** (0.014)
abs(Utility difference)			-0.052*** (0.006)	-0.047*** (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

The dependent variable is  $\log(RT)$

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

*number of conflicts between individually relevant motives and decreases with the utility difference between the choice options.*

#### 4.2.2 Evidence for the intuition/deliberation of selfishness

In this part, we examine the relevance of dual-process theories based on tests of whether selfishness or prosociality can be associated with the intuitive process. We first compare the RTs of selfish and social decisions to test Hypothesis 2.1, that is, whether subjects' intuition favors selfishness or prosociality based on the assumption of homogeneity. We exclude all consistent decisions and focus only on conflict 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. 3). We conduct a mixed-effects regression with  $\log(RT)$  as the dependent variable, and the regression result is shown in regression (1) of Table 7.<sup>14</sup> In this regression, the variable *Decision number* controls for learning, the variable *Utility difference* controls for the strength of preference, and the variable *Conflict within norms* (a dummy which indicates

<sup>14</sup> Regression (1) in Table 7 is the regression results for the odd trials in conflict decisions. The regression results for the even trials in conflict decisions are shown in regression (1) in Table C3.

**Table 7** Regressions of response times on the decision type and the strength-of-selfishness (the odd trials in conflict decisions)

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

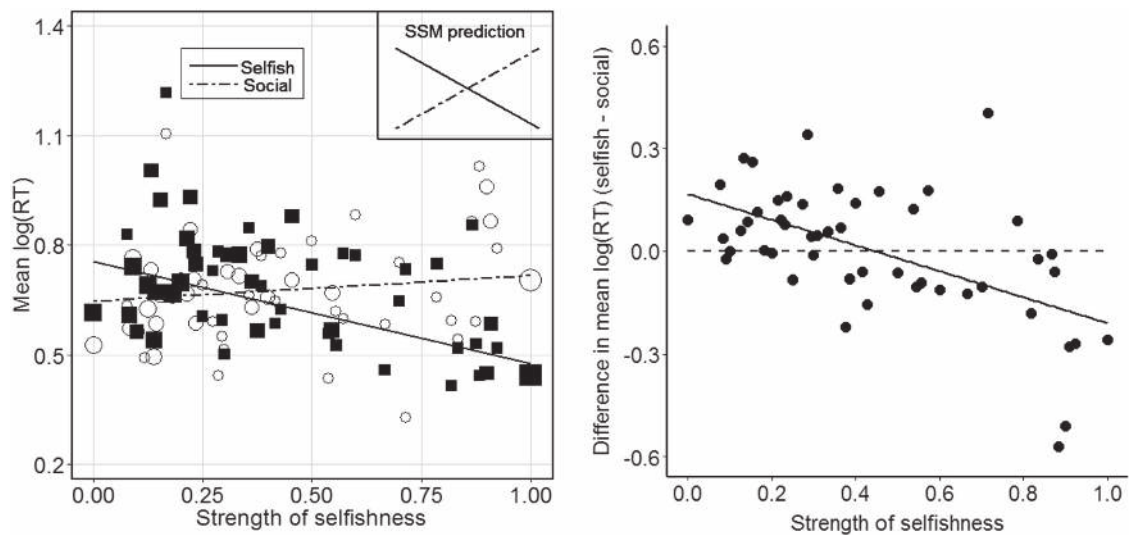
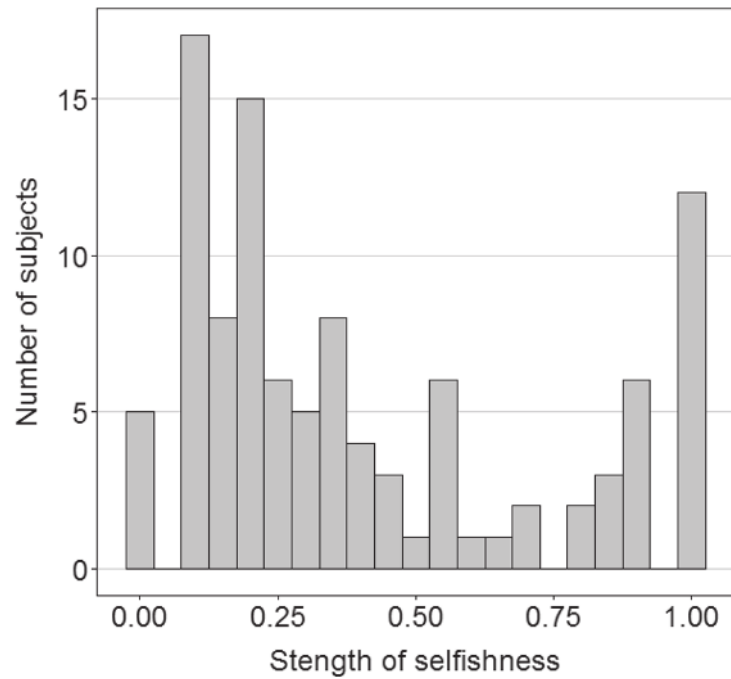
The dependent variable is  $\log(RT)$ . *Selfish Decision* is a dummy variable which indicates the decision is a selfish decision or social decision. *Conflict within norms* is a dummy which indicates whether there is a conflict within the social motives or not. *Male* is a dummy variable which indicates gender. The robust standard errors for regressions (2) and (3) are clustered on subjects and reported in parentheses

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

whether there is a conflict within the social motives or not) controls for the cognitive conflict of the decision. The coefficient of *selfish decision* is not significantly different from 0. That is, the RT does not differ between selfish and social decisions across all subjects while controlling for the cognitive conflict and the utility difference between choice options. The evidence does not support Hypothesis 2.1.

The analyses above show that there is no evidence that the selfish or the social decision is more intuitive if we ignore the heterogeneity in selfishness. Next, to test Hypothesis 2.2, we take the heterogeneity of selfishness into account and examine the influence of the strength of selfishness on subjects' intuition towards the selfish or the social motive. Specifically, we calculate the proportion of selfish decisions in the *even* trials of conflict decisions for each subject, and use this proportion as

**Fig. 6** The distribution of the strength-of-selfishness in the even trials of conflict decisions



**Fig. 7** Response times of selfish and social decisions in the odd trials of conflict decisions. The left panel plots the relationship between the strength of selfishness and the RTs of selfish/social decisions. The circle represents social decisions and the square represents selfish decisions. The size of the circle or square indicates the number of subjects in that circle or square. The dotted and solid lines are regression lines for social and selfish decisions separately. The “SSM prediction” in the upper right corner represents the correlation between the strength of selfishness and the RTs of selfish decisions (solid line)/social decisions (dashed line) predicted by the standard sequential sampling models (SSM), which should be symmetric. The right panel plots the strength of selfishness versus the RT difference between selfish and social decisions

our measure of the strength of selfishness.<sup>15</sup> Figure 6 shows the distribution of the strength of selfishness in the even trials of conflict decisions.

The left panel of Fig. 7 displays the relationship between the strength of selfishness and the RTs of selfish (social) decisions in the *odd* trials of conflict decisions.<sup>16</sup> The circle represents social decisions and the square represents selfish decisions. The size of the circle or the square indicates the number of subjects in that circle or square. The dashed line is the regression line for social decisions and the solid line is the regression line for selfish decisions. The left panel of Fig. 7 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,  $r = 0.059$ ,  $p = 0.117$ ) although the correlation has the expected sign. The right panel of Fig. 7 displays the RT 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.

The econometric analysis in Table 7 corroborates these results. Regression (3) in Table 7 shows that the RTs of selfish decisions significantly decrease with the strength of selfishness. However, the strength of selfishness has no significant effects on the RTs of social decisions although it has the expected sign. In a recent study, Rand et al. (2016) showed that intuition favors prosociality for women but not for men. We also control for the gender in our analysis. As shown in regressions (2) and (3) in Table 7, the RTs of selfish or social decisions for males are not significantly different from those for females. The econometric analysis controls for the strength of conflict and the strength of preference, which deals with the argument brought up by Krajbich et al. (2015a) and Evans et al. (2015). Therefore, our results show that the strength of selfishness interacts with the speed of selfish versus social decisions.

On the other hand, the relationship in the right panel of Fig. 7 can also be produced by sequential sampling models, since sequential sampling models predict that selfish decisions are quicker (slower) than social decisions for subjects who are more selfish (prosocial) because more selfish (prosocial) people put higher weight on the selfish (social) dimension (Chen and Krajbich 2018; Dickhaut et al. 2013; Frydman and Nave 2016; Hutcherson et al. 2015; Krajbich et al. 2010; Krajbich et al. 2015a; Krajbich et al. 2014). Based on this argument, in addition to the relationship in the right panel of Fig. 7, sequential sampling models also predict that 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

<sup>15</sup> We also checked whether subjects' selfish behavior is consistent over the rounds. The results show that the strength of selfishness (i.e., the frequency of choosing the selfish option in conflict decisions) in the first half rounds is not different from that in the second half rounds (Wilcoxon signed-rank test,  $p = 0.303$ ).

<sup>16</sup> We also calculate the strength of selfishness using the odd trials and study the RT in the even trials of conflict decisions. The results are similar and shown in Supplementary Material C.

corner of the left panel in Fig. 7 (SSM prediction). Obviously, the symmetric relationship is not observed in our data (the left panels of Fig. 7 and Fig. C3 in Supplementary Material C). That is, only sequential sampling models cannot explain all the patterns observed in our data.

Of course, the control is incomplete because the utility difference and the strength of 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. Subjects for whom selfishness is intuitive should look at their own payoff at the early stage of the decision process. We did not collect 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. D1 in Supplementary Material D).<sup>17</sup> To look at each fixation individually, we also plot the proportion of subjects who fixated on their own payoffs in each fixation. It shows that 78.6% of selfish subject fixated on their own payoffs in the first fixation, while 58.5% of pro-social subjects fixated 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. D2 in Supplementary Material D). 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 more prosocial.*

## 5 Conclusion

This paper studies the cognitive mechanisms underlying 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 and the heterogeneity in the strength of selfishness. We find evidence for both sequential sampling models and dual-process theories. First, 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 results indicate that the sequential sampling process and intuition coexist in the cognitive process underlying social decision making. And our results provide a possible explanation for the conflicting results concerning the automaticity of the selfish motive observed in the previous literature, as well as the debate

---

<sup>17</sup> The results are shown in Supplementary Material D.

between the dual-process explanation and the sequential sampling explanation of the cognitive processes underlying social decision making. Future work can investigate whether the sequential sampling process and intuition act on distinct levels. For instance, the option preferred by the intuitive process could have a head start and sampling from the deliberative system starts later. Or both the intuitive and the deliberative process could work via sequential sampling. Future work can also try to develop formal mathematical models that describe how they are combined, and in what order, to produce decisions.

Individual heterogeneity in social preferences is not at all controversial. But data limitations often force us to assume homogeneity, and explicitly taking heterogeneity into account is still rare. Our study shows that heterogeneity in preferences is reflected in the differences in 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 dual-process theories 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 us to draw causal inferences. However, the sequential sampling models make clear predictions that conflicts between motives increase RT and less utility difference between choice options needs 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 that the motive that is more relevant behaviorally is also more intuitive in the sense of a dual-process model. 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 such as time pressure or cognitive load.

**Acknowledgements** We would like to thank Carlos Alós-Ferrer, Dominik Bauer, Colin Camerer, David Dohmen, Fabian Dvorak, Gerald Eisenkopf, Ernst Fehr, David Grammling, Jan Hausfeld, Konstantin Hesler, Ian Krajbich, Daniel Martin, Gideon Nave, David Rand, Katrin Schmelz, Simeon Schudy, Roberto Weber and Irenaeus Wolf as well as participants of 2014 ESA European Meeting in Prague, 2014 Zurich Workshop on Experimental and Behavioral Economic Research, the Second International Meeting on Experimental and Behavioral Social Sciences in Toulouse for helpful comments and discussions. Fadong Chen gratefully acknowledges support from the National Natural Science Foundation of China (Grant No. 71803174), the Qiantang River Talents Program, and the Fundamental Research Funds for the Central Universities in China. Urs Fischbacher gratefully acknowledges support from the German Research Foundation (DFG) through research unit FOR 1882 “Psychoeconomics”.

## References

- Achtziger, A., & Alós-Ferrer, C. (2014). Fast or rational? A response-times study of Bayesian updating. *Management Science*, *60*(4), 923–938.
- Alós-Ferrer, C., & Strack, F. (2014). From dual processes to multiple selves: Implications for economic behavior. *Journal of Economic Psychology*, *41*, 1–11.
- Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737–753.

- Bear, A., & Rand, D. G. (2016). Intuition, deliberation, and the evolution of cooperation. *Proceedings of the National Academy of Sciences*, *113*(4), 936–941.
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, *90*(1), 166–193.
- Breitmoser, Y. (2013). Estimation of social preferences in generalized dictator games. *Economics Letters*, *121*(2), 192–197.
- Brocas, I., & Carrillo, J. D. (2014). Dual-process theories of decision-making: A selective survey. *Journal of Economic Psychology*, *41*, 45–54.
- Bruhlin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, *78*(4), 1375–1412.
- Cappelen, A. W., Nielsen, U. H., Tungodden, B., Tyran, J.-R., & Wengström, E. (2016). Fairness is intuitive. *Experimental Economics*, *19*(4), 727–740.
- Cappelletti, D., Güth, W., & Ploner, M. (2011). Being of two minds: Ultimatum offers under cognitive constraints. *Journal of Economic Psychology*, *32*(6), 940–950.
- Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, *13*(2), 195–212.
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.
- Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, *117*(3), 817–869.
- Chen, F., & Krajbich, I. (2018). Biased sequential sampling underlies the effects of time pressure and delay in social decision making. *Nature Communications*, *9*(1), 3557.
- Cone, J., & Rand, D. G. (2014). Time pressure increases cooperation in competitively framed social dilemmas. *PLoS ONE*, *9*(12), e115756.
- Cornelissen, G., Dewitte, S., & Warlop, L. (2011). Are social value orientations expressed automatically? Decision making in the dictator game. *Personality and Social Psychology Bulletin*, *37*(8), 1080–1090.
- Dickhaut, J., Smith, V., Xin, B., & Rustichini, A. (2013). Human economic choice as costly information processing. *Journal of Economic Behavior & Organization*, *94*, 206–221.
- Duffy, S., & Smith, J. (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, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, *47*(2), 268–298.
- Engelmann, D., & Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review*, *94*(4), 857–869.
- Erlei, M. (2008). Heterogeneous social preferences. *Journal of Economic Behavior & Organization*, *65*(3–4), 436–457.
- Evans, A. M., Dillon, K. D., & Rand, D. G. (2015). Fast but not intuitive, slow but not reflective: Decision conflict drives reaction times in social dilemmas. *Journal of Experimental Psychology: General*, *144*(5), 951–966.
- Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, *54*(2), 293–315.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 817–868.
- Fiedler, S., Glöckner, A., Nicklisch, A., & Dickert, S. (2013). Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational Behavior and Human Decision Processes*, *120*(2), 272–284.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.
- Fisman, R., Kariv, S., & Markovits, D. (2007). Individual preferences for giving. *American Economic Review*, *97*(5), 1858–1876.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.
- Frydman, C., & Nave, G. (2016). Extrapolative beliefs in perceptual and economic decisions: Evidence of a common mechanism. *Management Science*, *63*(7), 2340–2352.
- Fudenberg, D., & Levine, D. K. (2006). A dual-self model of impulse control. *American Economic Review*, *96*(5), 1449–1476.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125.



- Grün, B., & Leisch, F. (2008). FlexMix Version 2: Finite mixtures with concomitant variables and varying and constant parameters. *Journal of Statistical Software*, 28(4), 1–35.
- Hauge, K. E., Brekke, K. A., Johansson, L.-O., Johansson-Stenman, O., & Svedsäter, H. (2016). Keeping others in our mind or in our heart? Distribution games under cognitive load. *Experimental Economics*, 19(3), 562–576.
- Houser, D., Keane, M., & McCabe, K. (2004). Behavior in a dynamic decision problem: An analysis of experimental evidence using a Bayesian type classification algorithm. *Econometrica*, 72(3), 781–822.
- Hutcherson, C. A., Bushong, B., & Rangel, A. (2015). A neurocomputational model of altruistic choice and its implications. *Neuron*, 87(2), 451–462.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9), 697–720.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.
- Kerschbamer, R. (2015). The geometry of distributional preferences and a non-parametric identification approach: The equality equivalence test. *European Economic Review*, 76, 85–103.
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 1292–1298.
- Krajbich, I., Bartling, B., Hare, T., & Fehr, E. (2015a). Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature Communications*, 6, 7455.
- Krajbich, I., Hare, T., Bartling, B., Morishima, Y., & Fehr, E. (2015b). A common mechanism underlying food choice and social decisions. *PLoS Computational Biology*, 11(10), e1004371.
- Krajbich, I., Oud, B., & Fehr, E. (2014). Benefits of neuroeconomic modeling: New policy interventions and predictors of preference. *American Economic Review*, 104(5), 501–506.
- Lohse, J., Goeschl, T., & Diederich, J. H. (2017). Giving is a question of time: Response times and contributions to an environmental public good. *Environmental & Resource Economics*, 67(3), 455–477.
- Lotito, G., Migheli, M., & Ortona, G. (2013). Is cooperation instinctive? Evidence from the response times in a public goods game. *Journal of Bioeconomics*, 15(2), 123–133.
- McLachlan, G. J., & Basford, K. E. (1988). *Mixture models: Inference and applications to clustering* (Vol. 84). New York: M. Dekker.
- McLachlan, G. J., & Jones, P. N. (1988). Fitting mixture models to grouped and truncated data via the EM algorithm. *Biometrics*, 571–578.
- McLachlan, G. J., Lee, S. X., & Rathnayake, S. I. (2019). Finite mixture models. *Annual Review of Statistics and Its Application*, 6, 355–378.
- Merkel, A. L., & Lohse, J. (2019). Is fairness intuitive? An experiment accounting for subjective utility differences under time pressure. *Experimental Economics*, 22(1), 24–50.
- Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. (2011). Measuring social value orientation. *Judgment and Decision Making*, 6(8), 771–781.
- Nielsen, U. H., Tyran, J.-R., & Wengström, E. (2014). Second thoughts on free riding. *Economics Letters*, 122(2), 136–139.
- Peysakhovich, A., & Rand, D. G. (2016). Habits of virtue: Creating norms of cooperation and defection in the laboratory. *Management Science*, 62(3), 631–647.
- Piovesan, M., & Wengström, E. (2009). Fast or fair? A study of response times. *Economics Letters*, 105(2), 193–196.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83(5), 1281–1302.
- Rand, D. G., Brescoll, V., Everett, J. A., Capraro, V., & Barcelo, H. (2016). Social heuristics and social roles: Intuition favors altruism for women but not for men. *Journal of Experimental Psychology: General*, 145(4), 389–396.
- Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, 489(7416), 427–430.
- Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., et al. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, 5, 3677.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111(2), 333–367.
- Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *The Economic Journal*, 117(523), 1243–1259.

- Schulz, J. F., Fischbacher, U., Thöni, C., & Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology, 41*, 77–87.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3–22.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*(3), 220–247.
- Tinghög, G., Andersson, D., Bonn, C., Böttiger, H., Josephson, C., Lundgren, G., et al. (2013). Intuition and cooperation reconsidered. *Nature, 498*(7452), E1–E2.
- Verkoeijen, P. P., & Bouwmeester, S. (2014). Does intuition cause cooperation? *PLoS ONE, 9*(5), e96654.