Quantitative Economics 4 (2013), 515–547 1759-7331/20130515

Elicited beliefs and social information in modified dictator games: What do dictators believe other dictators do?

Nagore Iriberri
University of the Basque Country UPV/EHU, IKERBASQUE, Basque Foundation for Research and Universitat Pompeu Fabra

Pedro Rey-Biel
Universitat Autònoma de Barcelona and Barcelona GSE

Using data from modified dictator games and a mixture-of-types estimation technique, we find a clear relationship between a classification of subjects into four different types of interdependent preferences (selfish, social welfare maximizers, inequity averse, and competitive) and the beliefs subjects hold about others' distributive choices in a nonstrategic environment. In particular, selfish individuals fall into false-consensus bias more than other types, as they can hardly conceive that other individuals incur costs so as to change the distribution of payoffs. We also find that selfish individuals are the most robust preference type when repeating play, both when they learn about others' previous choices (*social information*) and when they do not, while other preference types are more unstable.

Keywords. Interdependent preferences, social welfare maximizing, inequity aversion, belief elicitation, social information, experiments, mixture-of-types models.

JEL classification. C72, C91, D81.

1. Introduction

In the last 20 years, the experimental literature has challenged the classic assumption that individuals are only motivated by their own individual material payoff, as they may in fact take into account the outcome of their decisions on others' payoffs.[1] This evi-

Nagore Iriberri: nagore.iriberri@gmail.com
Pedro Rey-Biel: pedro.rey@uab.es

We thank Jose Apesteguia, Giuseppe Attanasi, Ghazala Azmat, Miguel A. Ballester, Maite Cabeza, Gary Charness, Vincent P. Crawford, Steffen Huck, Manuel Mosquera, Rosemarie Nagel, Joel Sobel, Karl Schlag, Carmit Segal, anonymous referees, and seminar audiences at Games World Congress, Universitat Autònoma de Barcelona, University College London, University of Copenhagen, Universidade Nova de Lisboa, Universitat Pompeu Fabra, and University of Michigan, University of California at Berkeley and Maastricht University for their comments. We are grateful to Aniol Llorente-Saguer and Natalia Montinari for their help in running the experiments. Nagore Iriberri acknowledges financial support from Ministerio de Ciencia e Innovación (ECO2011-25295) and Ministerio de Economía y Competitividad (ECO2012-31626). Pedro Rey-Biel acknowledges financial support from Ministerio de Ciencia y Tecnología (Grant ECO2009-07616), Ministerio de Educación (Grant ECO2012-31962) and Generalitat de Catalunya (Grant 2009SGR-169).

[1] See Fehr and Schmidt (2000) and Sobel (2005) for comprehensive and excellent surveys.

Copyright © 2013 Nagore Iriberri and Pedro Rey-Biel. Licensed under the Creative Commons Attribution-NonCommercial License 3.0. Available at http://www.qeconomics.org.
DOI: 10.3982/QE135

dence has given rise to extensive work on interdependent (or "social") preferences.[2] Different types of preferences have been proposed. Standard "selfish" preferences assume individuals only care about their own material payoff. "Social welfare maximizing" preferences correspond to individuals who care positively about others' payoffs. "Inequity averse" preferences imply that individuals care positively about others' payoffs when they are better off than others, but care negatively when they are worse off than others (Fehr and Schmidt (1999), and Bolton and Ockenfels (2000)). Finally, "competitive" preferences assume that individuals care negatively about others' payoffs. Charness and Rabin (2002) (CR from here on) encompassed these four types of interdependent preferences in a simple piecewise linear utility model with two parameters that capture the weight on others' payoffs.[3]

Several studies have aimed to find the interdependent utility function that explains best the distributional choices made by subjects in experiments performed in the laboratory.[4] More recently, researchers have aimed to identify and quantify different types of interdependent preferences in experiments where subjects make distributive decisions (Andreoni and Miller (2002), Fisman, Kariv, and Markovits (2007), and Blanco, Engelmann, and Normann (2011)). An important finding of these studies is that the existence of heterogeneity in interdependent preferences cannot be ignored. In particular, around half of the subjects in these experiments behave as selfish, while a minority behave as competitive. Moreover, there exists a significant portion of subjects whose behavior is consistent with social welfare maximizing preferences and with inequity aversion.

This paper goes one step further and studies the role *beliefs* about other individuals' choices play in one's own choice. We identify individuals' beliefs about others' actions and analyze how these beliefs relate to different interdependent preferences types. Furthermore, we study the robustness of interdependent preferences types to repeated play after eliciting beliefs about others' actions, that is, *repeated play*, and after eliciting beliefs about others' actions and actually knowing about others' choices, that is, *social information*.

With respect to beliefs, any application that assumes heterogeneity of interdependent preferences requires assumptions about individuals' beliefs about others' actions and thus preferences. In this paper, we elicit and measure beliefs about others' actions to test how much of this heterogeneity in preferences is actually known by the subjects. In particular, this paper empirically tests whether a relationship exists between subjects' distributive choices and their beliefs about others' distributive choices. For instance, do

[2] "Social preferences" and "other-regarding preferences" have been used to refer to distributional preferences as well as reciprocity concerns. Since our setting is nonstrategic, we focus on purely distributional preferences and thus use the term "interdependent preferences" to refer to purely distributional concerns. See also Fisman et al. (2007) for a discussion on the difference between preferences for giving and social preferences.

[3] We refer to the model presented on page 822 in Charness and Rabin (2002), where parameter q is set to be equal to zero (no reciprocity issues considered). Thus, there are only two relevant parameters in the model, ρ and σ, the weights for others' payoffs when the subject is ahead and behind, respectively (as reproduced in our equation (1) in Section 4).

[4] See, for example, Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Engelmann and Strobel (2004).

selfish individuals think others are also selfish? What about the beliefs of other interdependent preferences types?

Furthermore, regarding repeated play and social information, if interdependent preferences only depend on the distribution of payoffs, then repeating behavior after eliciting beliefs about others' actions and after letting them know about others' actions should not affect individual behavior. To test this robustness, we study how individuals react to the information on others' choices when confronted with the same choice problems after having elicited beliefs about them. For instance, do selfish individuals modify their behavior when they are provided with information on others' actions compared to when they are simply asked about others' actions?

We depart from the current experimental literature on belief elicitation by using a purely decision making and, therefore, nonstrategic setting.[5] In our experiment, subjects express their beliefs about actions taken by other subjects with whom they never interact and whose actions can never affect their own payoffs. We proceed in this way for two reasons. First, this offers a clean test for the role beliefs and knowledge of heterogeneity in others' actions might play, if any, in purely interdependent preferences. Second, nonstrategic environments control for issues such as intention-based utilities, perceptions of kindness and unkindness, and/or reciprocal behavior.[6] Thus, nonstrategic settings are an ideal test bed for whether beliefs and social information are relevant in modelling other-regarding behavior.

With regard to the effect of social information on choices, there exists work both in the laboratory and in the field.[7] Cason and Mui (1998) are closest to our work. They studied the effect of social information on behavior using a regular dictator game design and tested the "social influence hypothesis" (see Akerlof (1982), Jones (1984), and Bernheim (1994)), that is, whether an individual's perception of what constitutes socially appropriate behavior depends on her estimate of others' beliefs regarding what constitutes socially appropriate behavior. Additionally, Cason and Mui (1998) studied the effect of social information on different types of individuals, but they only distinguished between selfish and nonselfish types. We add to this literature by analyzing how four different interdependent preferences types react to social information given that our design allows for a connection between four different interdependent preferences types and social information.

Our experimental design is a modified dictator game inspired by the designs of both Andreoni and Miller (2002) and Fisman et al. (2007), appropriately adapted to fit our research question dealing with beliefs and social information.[8] Deciders in our experi-

[5]See, for example, Croson (2000), Nyarko and Schotter (2002), Costa-Gomes and Weizsäcker (2008), Rey-Biel (2009), and Ivanov (2011).

[6]Dufwenberg and Gneezy (2000), Cabrales et al. (2010), Gächter and Renner (2010), and Bellemare, Kröger, and van Soest (2008) elicited beliefs about others' actions in strategic settings where actions taken by other subjects affect own payoffs.

[7]Duffy and Kornienko (2010) studied how competitive or generous environments affect giving in sequential dictator games. In the field, Frey and Meier (2004) and Croson and Shang (2008) found that it is possible to use information on how much others have contributed to increase charitable contributions.

[8]Our main difference with respect to their designs resides in having only three available actions. Making the choice set discrete allows us to elicit beliefs and provide information on other subjects' actions in a simple and meaningful manner.

ment have to choose in 16 different decision tables among three different options that yield different payoff distributions for a decider and a receiver. The three options consist of a self-payoff maximizing choice, a surplus creating choice, in which deciders give up one payoff unit to allow the receiver to obtain $s > 1$ more units, and a surplus destroying choice, in which deciders give up one payoff unit to destroy the receiver's payoff in the same $s > 1$ units. The 16 decision tables differ on whether the decider is better off or worse off than the receiver, as well as in s.

This simple design allows us to identify four of the most prominent types of interdependent preferences included in the CR model: selfish (SF), social welfare maximizing (SW), inequity averse (IA), and competitive (CP). Regarding choices allowed for each type, notice that a SF decider always chooses the self-payoff maximizing choice. A SW decider either chooses the selfish or surplus-creating action, but never a surplus-destroying action. An IA decider, on the other hand, either chooses the selfish or surplus-creating action when he is better off than the receiver, but either the selfish or surplus-destroying action when he is worse off than the receiver. Finally, a CP decider either chooses the selfish or surplus-destroying action, but never a surplus-creating action. Furthermore, for SW, IA, and CP, choices between the selfish and the surplus-creating and -destroying actions depend on the *price* for such actions, given by $1/s$, in such a way that the types show consistency in their behavior. Fisman et al. (2007) distinguished an additional type thet call lexicographic preferences for self over others, which are found to be very prevalent. These preferences reflect a concern for own payoffs first and, provided own payoffs do not change, a concern for others' payoffs. Given that in our design nonselfish actions are costly in terms of own payoffs, we cannot distinguish the lexicographic from the selfish type.

The experiment consists of three parts. First, subjects take actions over the 16 decision tables. Given their decision profile, we are able to perform a within subject classification of subjects into the four different preferences types. This first part, although similar to previous work, is crucial to later relate different types of interdependent preferences to beliefs and social information. Second, we elicit deciders' beliefs about other deciders' actions in exactly the same 16 decision tables. This allows us to identify different beliefs among the subject population and to classify each individual into a belief types. We compared the preferences-type classification with the belief-type classification to measure how much of the existent heterogeneity in actions is known to the subjects and also to test for correlation between their actions and beliefs. Finally, deciders take actions over the 16 decision tables as in the first part of the experiment, again after belief elicitation. We have two treatments. In one of them, we provide subjects with information about the distribution of choices other deciders previously made. In the other, we do not provide them with any information. Comparing the treatment with social information with a treatment without it allows us to study the robustness of type classification to repeated play and social information.

We find a preferences-type distribution very similar to those found by Andreoni and Miller (2002) and Fisman et al. (2007). Selfish preferences type is the most frequent (47% of the subjects), followed by inequity averse individuals (21%) and social welfare maximizers (19%). A nontrivial fraction of subjects are classified as being competitive (13%).

More importantly, we find that individuals who have different interdependent preferences indeed have different beliefs about others' actions and, furthermore, beliefs are correlated with their type. Selfish individuals systematically state that they believe other individuals take selfish actions, while other preferences types are more aware of the existent heterogeneity in actions. Finally, and contrary to the purely *interdependent* preferences assumption, although the preferences-type classification is relatively robust to social information and repeated play, it affects different types very differently. While almost no selfish subject changes his type, around half of the subjects classified as having other-regarding preferences (SW, IA, and CP) change their type, tending to behave overall more selfishly both with and without social information. This last result shows that other-regarding preferences' non-stability is mainly driven by repeated play and not by social information.

Our results suggest that it may be problematic to assume that beliefs about others' choices are independent of preferences types, as well as to assume that all types are as robust to repeated play and social information. These findings have important implications for modelling interdependent preferences, as well as for the application of interdependent preferences to both nonstrategic and strategic settings. Thus, this paper contributes to the discussion of whether purely interdependent preferences, which take into account only payoff differences, capture the essence of other-regarding preferences or, on the contrary, whether extended models, which include others' *expected* behavior, are required. We find that, in the simplest nonstrategic setting in which other-regarding preferences affect behavior, other-regarding preferences are a more complicated object than the reduced form modelled by purely interdependent preferences.

The rest of the paper is organized as follows. Section 2 explains the experimental design and procedures. Section 3 shows the main descriptive statistics in the three parts of the experiment. Section 4 describes the classification of subjects into four interdependent preferences types according to their choices in the first part of the experiment. Section 5 explains the belief-type identification and classification, and studies the correlation between the actions-based and beliefs-based classifications. Section 6 shows the new classification of subjects according to their actions in Part 3 of the experiment, both when they have been exposed to social information and when they have not. Section 7 concludes. Experimental instructions are included in the Appendix.

2. Experimental design and procedures

Six experimental sessions were conducted in the Laboratori d'Economia Experimental (LEEX) at Universitat Pompeu Fabra using z-tree experimental software (Fischbacher (2007)) in February 2008 and January 2011. A total of 240 subjects, 40 per session, were recruited using the Online Recruiting System for Economic Experiments (ORSEE) recruiting system (Greiner (2004)), ensuring that subjects had not participated in similar experiments in our laboratory in the past. After arrival, subjects extracted a piece of paper from a bag that randomly determined whether they would stay in the lab or they

would go to a different classroom. We refer to the 120 subjects in the lab as "deciders" and to the 120 subjects in the classroom as "receivers."[9] Further, the 20 deciders in each session were divided into two groups of 10 subjects each, which will be relevant for Parts 2 and 3 of the experiment. Instructions for each of the subject roles were also read aloud in each room before tasks were performed. Once the experiment had concluded, subjects filled in a voluntary questionnaire while they waited to be paid.

Each experimental session lasted one and a half hours (including assignment of subjects to rooms and payment). Throughout the experiment, we ensured anonymity and effective separation between deciders and receivers, locating them physically in different rooms, to minimize any interpersonal influences that could stimulate other-regarding behavior. Subjects were paid individually and in private, using a closed envelope and starting with deciders first.

Deciders performed three tasks that determined the payoffs for both player roles. Receivers were also read the deciders' instructions for Parts 1 and 3 such that they would know how their own payoffs were determined and waited in a separate classroom, filling in a voluntary questionnaire that had no influence on their payoffs. Tasks were presented in three different parts. For all three parts, deciders were shown the same 16 decision tables that describe the allocation of experimental units among two subjects.[10] The order in which tables were shown to subjects was changed randomly from one task to the other, aiming to control for possible order effects and to keep subjects engaged.

Each decision table contained three options, which show different allocations of experimental units between deciders and receivers, as illustrated in Figure 1. One of the options contained the highest number of experimental units for the decider; we refer to such an option as the *selfish* action (S). Another option was constructed such that the decider would lose one experimental unit so as to *increase* the receiver's allocation by $s > 1$ units. We refer to this option as the *surplus-creating* action (C). The third option was constructed such that the decider would lose again one experimental unit, but this time so as to *decrease* the receiver's allocation by $s > 1$ units. We refer to this option as the *surplus-destroying* action (D).[11] We refer to $1/s$ alternatively as the price of creating or destroying surplus. Tables differed on (i) the difference between the decider's and the receiver's allocations ($|x - y|$), (ii) the decider's relative position with respect to the receiver ($x \gtrless y$) and whether this would change depending on the chosen action, for example, if $x > y$, whether $x - 1 \gtrless y + s$,[12] and (iii) the number of created and destroyed experimen-

[9]Subjects know their role in the experimental task before they take any action (role certainty). In Iriberri and Rey-Biel (2011), we found that selfish types were much less prevalent when using role uncertainty than when using role certainty.

[10]An experimental unit was equal to 0.25 euro.

[11]Labels for options used neutral language and their order was randomized.

[12]In 6 out of 16 tables, Tables 2, 3, 5, 7, 11, and 12, deciders' payoffs were higher than receivers' for all three available choices. In another 6 tables, Tables 1, 6, 8, 10, 13, and 14, deciders' payoffs were lower than receivers' for all available choices. Finally, in 4 out of 16 tables, deciders' relative positions changed depending on the chosen action. In Tables 9 and 15, the deciders' relative positions changed from ahead to behind only when the surplus-creating actionwas chosen. In Tables 4 and 16, the decider's relative positions changed from behind to ahead only when the surplu-destroying action was chosen. When referring to subjects' relative positions in a table, we refer to their positions when taking the selfish action.

Table 1 ($s = 7$)	Option 1	Option 2	Option 3
Decider	7	7	8
Receiver	10	24	17

Table 2 ($s = 5$)	Option 1	Option 2	Option 3
Decider	16	17	16
Receiver	3	8	13

Table 3 ($s = 2$)	Option 1	Option 2	Option 3
Decider	20	19	19
Receiver	5	7	3

Table 4 ($s = 7$)	Option 1	Option 2	Option 3
Decider	10	10	11
Receiver	21	7	14

Table 5 ($s = 4$)	Option 1	Option 2	Option 3
Decider	17	16	16
Receiver	8	12	4

Table 6 ($s = 3$)	Option 1	Option 2	Option 3
Decider	8	7	7
Receiver	17	14	20

Table 7 ($s = 3$)	Option 1	Option 2	Option 3
Decider	17	16	16
Receiver	8	11	5

Table 8 ($s = 5$)	Option 1	Option 2	Option 3
Decider	8	7	7
Receiver	17	12	22

Table 9 ($s = 6$)	Option 1	Option 2	Option 3
Decider	13	14	13
Receiver	5	11	17

Table 10 ($s = 4$)	Option 1	Option 2	Option 3
Decider	4	5	4
Receiver	24	20	16

Table 11 ($s = 7$)	Option 1	Option 2	Option 3
Decider	16	16	17
Receiver	1	15	8

Table 12 ($s = 4$)	Option 1	Option 2	Option 3
Decider	20	19	19
Receiver	5	1	9

Table 13 ($s = 2$)	Option 1	Option 2	Option 3
Decider	4	4	5
Receiver	22	18	20

Table 14 ($s = 6$)	Option 1	Option 2	Option 3
Decider	7	7	8
Receiver	23	11	17

Table 15 ($s = 3$)	Option 1	Option 2	Option 3
Decider	13	13	14
Receiver	8	14	11

Table 16 ($s = 5$)	Option 1	Option 2	Option 3
Decider	10	10	11
Receiver	19	9	14

FIGURE 1. The 16 decision tables.

	S	C	D
Decider	x	$x-1$	$x-1$
Receiver	y	$y+s$	$y-s$

FIGURE 2. Decision table.

tal units, that is, on s, which varied between 2 and 7. The structure of the tables is shown in Figure 2.

Deciders' tasks were as follows. In Part 1, they had to choose one of the three options in each of the 16 tables, knowing that they were randomly and anonymously matched with a different participant in each table and that their payoffs corresponded to that of decider, while the receiver's payoffs corresponded to a matched receiver in another classroom.

In Part 2, we elicited deciders' beliefs about other deciders' actions. The deciders' task was to guess how many of 10 deciders in the same session had chosen each of the three options in each of the 16 tables.

Finally, in Part 3, deciders had to choose again among the three options in each of the 16 tables. In a control treatment, half the number of subjects, 60, chose without having any additional information. In the *social information treatment*, the other 60 subjects were informed about the exact distribution of choices previously made by 10 deciders in their session in each of the 16 tables in Part 1. Deciders were again matched randomly and anonymously to a receiver in another classroom who was different from the one in Part 1 to avoid possible compensations between amounts allocated in Part 1 and Part 3.

At the end of the experiment, three tables were randomly chosen to determine payments. Deciders received the sum of a 3 euro participation fee plus the allocation they had chosen for decider in the randomly chosen tables in Parts 1 and 3 plus the amount earned according to a quadratic scoring rule that rewards accuracy of their elicited beliefs in the randomly chosen table in Part 2.[13] Receivers earned the 3 euro participation fee plus the allocation for the receiver chosen by their randomly matched decider in the randomly chosen tables in Parts 1 and 3. Average total payments were 13.94 euros for deciders and 9.25 euros for receivers.

3. Descriptive statistics

Table 1 reports the number of times each of the available actions—selfish, surplus-creating, and surplus-destroying actions—was chosen in Part 1 of the experiment, separating for decider's initial relative position (see footnote 12). The selfish action was chosen with highest frequency on average (72%) and also in each of the 16 tables. It was chosen slightly less frequently when deciders were ahead (68%) than when they were behind (76%). The surplus-creating action was chosen with second highest frequency overall

[13]The particular quadratic scoring rule (QSR) used in the experiment is shown in the Instructions. There is no consensus yet among experimentalists about the optimal incentive mechanism for eliciting beliefs. For other applications, see Huck and Weizsäcker (2001) and Charness and Dufwenberg (2006). For a discussion on QSRs, see Offerman and Sonnemans (2001) and Andersen et al. (2007).

TABLE 1. Actions in Part 1 of the experiment.

	Overall			Decider's Position: Ahead			Decider's Position: Behind			
	S	*C*	*D*	*S*	*C*	*D*	*S*	*C*	*D*	Total
Number of actions	1383	376	161	655	246	59	728	130	102	1920
Average by subject	11.53	3.13	1.34	5.46	2.05	0.49	6.07	1.08	0.85	16
Stand. dev.	(2.59)	(2.31)	(1.58)	(2.76)	(2.59)	(1.35)	(2.38)	(1.87)	(1.77)	
Frequency of play	0.72	0.20	0.08	0.68	0.26	0.06	0.76	0.14	0.11	

TABLE 2. Elicited beliefs.

	Overall			Decider's Position: Ahead			Decider's Position: Behind		
	S	*C*	*D*	*S*	*C*	*D*	*S*	*C*	*D*
Average	0.73	0.15	0.11	0.72	0.17	0.1	0.74	0.13	0.12
Stand. dev.	(0.21)	(0.15)	(0.14)	(0.22)	(0.18)	(0.13)	(0.22)	(0.15)	(0.17)

(19%), although it was more frequently chosen when the decider was ahead (26%) than when behind (14%). Finally, the surplus-destroying action was the least chosen (8%). Deciders chose to destroy surplus more frequently when behind (11%) than when ahead (6%). Although average behavior did not change much across tables, standard deviations indicate that there exists variability across subjects, which can be explained with the existence of different preferences types.

Table 2 reports the average frequency subjects assigned to each of the actions, taken by a group of 10 deciders in the same session. We observe that subjects expected the selfish action to be chosen, on average, with highest frequency (73%), which as we have seen, was correct. Furthermore, on average, subjects consistently believed that the selfish action was chosen with highest frequency in all 16 tables, no matter what the decider's relative position. Surplus-creating and -destroying actions were expected to be chosen with lower frequencies (15% and 11%, respectively). The surplus-creating action was believed to be chosen with slightly higher frequency when deciders were ahead (17%) than behind (13%). Finally, the surplus-destroying action was expected to be chosen slightly more frequently when deciders were behind (12%) than ahead (10%). Standard deviations also indicate that heterogeneity in beliefs exists.

Finally, Table 3 reports the frequency with which each of the available actions—selfish, surplus-creating, and -destroying actions—was chosen in the third part of the experiment in both the control and the social information treatments. We again observe the familiar pattern that the selfish action was chosen with highest frequency in both

Table 3. Actions in Part 3 of the experiment.

	Overall			Decider's Position: Ahead			Decider's Position: Behind			
	S	*C*	*D*	*S*	*C*	*D*	*S*	*C*	*D*	Total
				With Social Information						
Number of actions	716	183	61	342	114	24	374	69	37	960
Average	11.93	3.05	1.02	5.7	1.9	0.4	6.23	1.15	0.62	16
Stand. dev.	(5.12)	(4.75)	(2.79)	(2.8)	(2.7)	(1.29)	(2.59)	(2.33)	(1.63)	
Frequency of Play	0.75	0.19	0.06	0.71	0.24	0.05	0.78	0.14	0.08	
				Without Social Information						
Number of actions	740	137	83	369	89	22	371	48	61	960
Average	12.33	2.28	1.38	6.15	1.48	0.37	6.18	0.8	1.02	16
Stand. dev.	(4.36)	(3.62)	(2.88)	(2.41)	(2.35)	(1.02)	(2.52)	(1.58)	(2.12)	
Frequency of play	0.77	0.14	0.09	0.77	0.19	0.05	0.77	0.1	0.13	

treatments and irrespective of the decider's relative position. Comparing average frequency of play in Parts 1 and 3 of the experiment, in all decision tables but three (decision tables 1, 8, and 13 in Figure 1), the selfish action was chosen with higher frequency in Part 3 than in Part 1. The surplus-creating action was more frequently chosen when ahead (24% in the control, 19% in the social information treatment) than when behind (14% in the control and 10% in the social information treatment). Finally, the surplus-destroying action was chosen with the lowest frequency, although more frequently when deciders were behind (8% in the control, 1% in the social information treatment) than when ahead (5% in both treatments).

Standard deviations in the three tables indicate that heterogeneity exists in chosen actions and beliefs across subjects. The analysis in the following sections studies the sources of heterogeneity behind the average behavior reported here by looking at individual decisions and beliefs.

4. Results in Part 1 of the experiment: Estimation of the distribution of interdependent preferences types

This section describes the identification strategy of different interdependent preferences types in the first part of the experiment and presents the estimated type distribution.

Our econometric specification follows a structural estimation using the mixture-of-types models of Stahl and Wilson (1994, 1995), Harless and Camerer (1994), El-Gamal and Grether (1996), Costa-Gomes, Crawford, and Broseta (2001), Camerer, Ho, and Chong (2004), Costa-Gomes and Crawford (2006), and Crawford and Iriberri (2007a,

2007b).[14] Readers who are familiar with the application of mixture-of-type models can skip ahead to the results.

The structural model is based on CR's piecewise linear preferences utility function shown in equation (1). Deciders' utility (u_D) depends on both the decider's own payoff (π_D) and the receiver's payoff (π_R). The two key parameters refer to the weight on the receiver's payoff, ρ, when the decider is ahead of the receiver ($\pi_D > \pi_R$), and the weight, σ, when the decider is behind the receiver ($\pi_R > \pi_D$). Each decider i at decision table t has three available actions $a = \{S, C, D\}$, which refer to selfish (S), surplus-creating (C), and surplus-destroying (D) actions, respectively. The utility of a given decider at decision table t taking action a is thus given by

$$\begin{aligned} u_D(\pi_{Rta}, \pi_{Dta}) &= (\rho r + \sigma v)\pi_{Rta} + (1 - \rho r - \sigma v)\pi_{Dta} \\ &\quad \text{for } t = 1, \ldots, T \text{ and } a = \{S, C, D\}, \end{aligned} \tag{1}$$

where $r = 1$ if $\pi_D > \pi_R$ and $r = 0$ otherwise, and $v = 1$ if $\pi_D < \pi_R$ and $v = 0$ otherwise.

Four different interdependent preferences types are considered: selfish (SF), social welfare maximizers (SW), inequity averse (IA), and competitive (CP). A preferences type k is defined by the sign of parameters ρ and σ, such that each type imposes specific constraints on the signs these parameters can have. For the SF type, both parameters are zero such that a SF individual always chooses the selfish action. For the SW type, both parameters are strictly positive in a way that a SW individual chooses either the surplus-creating action or the selfish action, regardless of the decider's relative position. For the IA type, ρ is strictly positive and σ is nonpositive, and the IA individual chooses either the surplus-creating action or the selfish action when ahead, while choosing either the selfish action or the surplus-destroying action when behind. Finally, for the CP type, both parameters are nonpositive and at least one parameter is strictly negative, and the CP individuals choose either the surplus-destroying action or the selfish action, regardless of their relative position.[15] Therefore, preferences types determine a specific region for each of the parameters (ρ_k, σ_k). We refer to the utility of the decider, who belongs to preferences type k, as $u_{Dk}(\cdot)$.

More importantly, for all types, except for the SF type, the choice between the surplus-creating (-destroying) action and the selfish action must be *consistent* with respect to the *price* of the creating (destroying) action, given by $(1/s)$, where s is the number of created (destroyed) units in the decision tables. In other words, preferences types also determine choices to be consistent across different decision tables. For example, if a SW individual decides to create surplus at a certain price, that is, when price $= 1/6$

[14] Our specification—individual by individual estimation and uniform errors—is closest to El-Gamal and Grether (1996) and Costa-Gomes et al. (2001). The main difference from El-Gamal and Grether's (1996) and Costa-Gomes et al.'s (2001) applications is that types are determined by the values (ρ_k, σ_k) can take.

[15] Subjects who take surplus-creating actions when ahead but who behave as purely selfish when behind are considered to be inequity averse since their behavior would always yield a more equalitarian distribution of payoffs. Subjects who choose either the selfish action or the surplus-destroying action are considered to be competitive. For an extended model that considers these types as separate types, see Kerschbamer (2012).

($s = 6$), she also should choose to create surplus when it is cheaper to do so, that is, when price $= 1/7$ ($s = 7$). The same applies for IA and CP types.

We consider a probabilistic model, where all preferences-types' actions are subject to error. We use a uniform independent and identically distributed (iid) error across different decision tables. With some probability, given by ε, to be estimated, each preferences type chooses any of the available three actions with equal probability. Hence, according to CR's utility function and the iid error, the predicted choice at decision table t for a decider who belongs to preferences type k is

$$\begin{aligned} &\text{PredictedChoice}(a|\rho_k, \sigma_k, \varepsilon)_{Dkt} = (1-\varepsilon)1_{a=\arg(\max_a u_{Dk}(\pi_{Rta}, \pi_{Dta}))} + \frac{\varepsilon}{3} \\ &\quad \text{for } a = \{S, C, D\}, k = \{S, SW, IA, CP\} \text{ and } t = 1, \ldots, T. \end{aligned} \tag{2}$$

The indicator function 1_a takes value 1 if action a yields the highest utility and 0 otherwise. The uniform error is easily interpreted. With no error, $\varepsilon = 0$, the action that yields highest utility is chosen with probability 1. With positive error, $1 > \varepsilon > 0$, the action that yields the highest utility is chosen with higher probability than other actions, although it is chosen with probability less than 1. Finally, if $\varepsilon = 1$, the individual is purely random and chooses any of the available actions with equal probability.[16]

Notice that CR's utility function is restrictive in its specific *linear* form. Therefore, the error term is capturing two types of errors. One type of error is, for example, taking both surplus-creating and -destroying actions in tables in which the subject's relative position is kept constant. No preferences type k can explain this type of error, which is not implied by the linearity restriction, but by the basic consistency restriction that indifference curves should not cross. The other type of error is creating or destroying surplus for a certain price, but not doing so for a lower price. For example, a SW individual who chose to create surplus in all the tables except in decision table 11 in Figure 1, where she chose the selfish action, will be considered to be making a mistake in this last decision, yielding a positive ε. In decision table 11 in Figure 1, it is cheaper to create surplus ($s = 7$) compared to many other tables where she did choose to create surplus ($s < 7$), such that this individual did not make a consistent choice according to SW. These second types of mistakes partly come from the linearity restriction. Using a more flexible utility function such as constant elasticity of substitution could accommodate some of this second type of errors. For simplicity, we will stick to CR's linear utility function.[17]

The decision data collected in Part 1 of the experiment consisted of T decisions over S, C, and D actions for each of the N deciders, called, in general, *Choice*. The typical observation, called $\text{Choice}(a)_{Dit}$, takes value 1 if individual i took action a at decision table t and 0 otherwise. Having described the predicted choice in equation (2) and decider's actions data, we can now construct the likelihood function.

[16]Note that the error term ε captures the uncertainty involved in the estimation. The lower (higher) is the ε, the lower (higher) is the uncertainty involved in the estimation. In the extreme case of $\varepsilon = 0$, the uncertainty is lowest and, thus, individual behavior is captured by a specific preferences type with certainty. In the extreme case of $\varepsilon = 1$, uncertainty is highest, since no preferences type captures individual behavior.

[17]A constant elasticity of substitution utility function includes an extra parameter that determines the curvature of the indifference curve, allowing for linear but also Cobb–Douglas or Leontief functional forms.

The specification we use is an individual by individual estimation, which yields a set of estimated parameters (p_i, ε_i) for each individual i, where p_i is estimated to be equal to 1 for the preferences type that explains best decider i and 0 for other types, and ε_i is the individual-specific error rate.[18] The overall preferences-type distribution is obtained by counting the number of subjects classified in each type. We do not report on the values of the parameters ρ_i and σ_i because they are not identified at the individual level.[19] The likelihood function to be maximized is

$$\begin{aligned} &L_{Di}(p_{ki}, \rho_i, \sigma_i, \varepsilon_i | \text{Choice}_{Di}) \\ &\quad = \sum_{k=S,SW,IA,CP} p_{ki} \prod_{t=1}^{T} \prod_{a=\{S,C,D\}} \text{PredictedChoice}(a|\rho_i, \sigma_i, \varepsilon_i)_{Dikt}^{\text{Choice}(a)_{Dit}}. \end{aligned} \tag{3}$$

Table A.1, available in a supplementary file on the journal website http://qeconomics.org/supp/135/supplement.pdf, shows the individual by individual estimation. The first six columns summarize the number of decision tables in which each decider takes the selfish, surplus-creating, and surplus- destroying actions, separating for deciders' relative position. The last three columns present estimated ε, the log likelihood, and the estimated preferences type. On the one hand, 38% of the subjects (46 out of 120) are estimated without any error and their preferences types are readable directly from their actions. There is at least one subject that can be classified into each preferences type without any error. Subject 4, among many others, is classified as SF because he consistently chose the selfish action in all decision tables. Subject 37 is classified as SW because she consistently chose the surplus-creating action in all decision tables. Subject 52 is classified as IA because she chose the surplus-creating action once when ahead but never when behind. Finally, subject 44 is classified as CP since she consistently chose the surplus-destroying action in all decision tables. Finally, there are only 4 subjects out of 120 (subjects numbers 60, 76, 98, and 106) whose type is not identified, since their behavior is consistent with various preferences types.

Table 4 summarizes the distribution of preferences types. SF is the most frequent type since almost half of the subjects are classified as having selfish preferences. It is followed by SW and IA types, with frequencies around 20% for each of the types. Finally, the least frequent type is CP, with a frequency of 13%. This preferences-type distribution

[18]If we considered a different specification for the error term, such as logit, ρ and σ would be identified. However, we are not interested in the actual estimation of these parameters, but the preferences type identification, and both the logit and the uniform error yield the same type classification. Furthermore, the uniform error is easier to interpret. With uniform error, we have also considered population level specifications, where the error rate is allowed to depend on the preferences type and where it is restricted to be equal for all types instead of allowing it to differ from individual to individual. Overall, the estimated type distribution is fairly robust across the three specifications.

[19]For example, an individual who chooses to create surplus when $s = 7$ but decide to take the selfish action when $s \leq 6$ will have an estimated value of $p_i = 1$ for SW type with an estimated $\varepsilon = 0$. However, about ρ and σ, we can only say that they will be $\geq 1/8$ (given by $1/(1+s)$, where $s = 7$) and $\leq 1/7$ (given by $1/(1+s)$, where $s = 6$) such that the likelihood function will be flat for any value of ρ and σ between 1/8 and 1/7. Note that this interval would be finer and more precise if we added more decision tables with different values for s between 6 and 7.

TABLE 4. Summary of individual by individual estimation (Part 1).

	Number of Individuals	Frequency	$\bar{\varepsilon}_k$
SF	55	0.47	0.04
SW	22	0.19	0.28
IA	24	0.21	0.25
CP	15	0.13	0.27
Total	116	1	

TABLE 5. Frequency of play separately for preferences types (Part 1) ($N = 116$ subjects).

				Decider's Position					
	Overall			Ahead			Behind		
	S	C	D	S	C	D	S	C	D
SF	0.97	0.02	0.01	0.98	0.01	0.01	0.97	0.02	0.01
SW	0.37	0.57	0.05	0.32	0.64	0.04	0.43	0.51	0.06
IA	0.59	0.32	0.08	0.42	0.55	0.03	0.77	0.09	0.13
CP	0.52	0.08	0.40	0.57	0.10	0.32	0.46	0.06	0.47
Total	0.72	0.19	0.08	0.69	0.25	0.06	0.76	0.14	0.10

is fairly similar to those previously found in the literature by Andreoni and Miller (2002) and Fisman et al. (2007).[20]

Table 5 reports the average frequency of play of the three available actions, $a = \{S, C, D\}$, separately for subjects classified in each of the preferences types. This table clearly shows the idea behind our identification strategy for different preferences types. Subjects classified as SF almost always chose the selfish action (97% of the time). Subjects classified as SW chose the surplus-creating action with the highest frequency (57% of the time) and very rarely decided to destroy surplus (5% of the time). Subjects classified as IA chose the selfish action with the highest frequency, but also took the surplus-creating and -destroying actions with nontrivial frequency (32% and 8% of the time, respectively). As expected, subjects classified as IA created surplus much more when ahead (55%) than when behind (0.9%) and they destroyed surplus more when behind (13%) than when ahead (3%). Subjects classified as CP were, as expected, the ones who took the surplus-destroying action with the highest frequency (40% of the time).

[20]Andreoni and Miller's (2002) design cannot distinguish between selfish and competitive preferences. They found the distribution for selfish, social welfare, and inequity averse to be, respectively, 44%, 21%, and 35%. Fisman et al.'s (2007) design can further identify what they call lexicographic for self over other while we cannot, so if we add up their lexicographic and selfish frequencies, their type distribution for selfish, social welfare, inequity averse, and competitive is 62%, 13%, 19%, and 5%, respectively. Both distributions are quite similar to ours.

5. Results in Part 2 of the experiment: Belief-type identification and correlation between the interdependent preferences types and belief types

This section describes the belief-type identification strategy and presents the estimated belief types with different econometric specifications. We also look for correlations between the identified belief types and the interdependent preferences types already estimated in the previous section.

We start by commenting on overall belief accuracy. We calculate the average squared error (ASE) between subjects' beliefs and the real distribution of actions, averaging across all subjects. The ASE over all 16 tables was 20.1, that is, around 10% of the maximum error subjects could have made.[21] Although this ASE seems to indicate that subjects were reasonably accurate, averaging across subjects gives a misleading idea of the knowledge subjects had about the heterogeneity in actions. As will become clear later, there exist significant differences in beliefs across subjects.

Our objective is to identify the belief types present in the subject population and measure the level of heterogeneity in their beliefs. Only after we have identified them will we be able to classify each subject into different classes of beliefs so as to relate the two classifications, one based on their actions, and thus on preferences, and the other based on their beliefs.

One simple way to identify beliefs and look for differences among different preferences types consists of averaging elicited beliefs across individuals who were classified into the same preferences type according to their choices in Part 1 of the experiment. Table 6 shows the average frequency of play expected by those subjects classified as belonging to each of the preferences types. All types of subjects assign highest frequency

Table 6. Expected frequency (and standard deviations) of play separately for preferences types.

	S	C	D
SF	0.89	0.07	0.05
	(0.13)	(0.09)	(0.08)
SW	0.55	0.31	0.14
	(0.18)	(0.17)	(0.12)
IA	0.64	0.22	0.14
	(0.16)	(0.13)	(0.14)
CP	0.61	0.13	0.26
	(0.20)	(0.10)	(0.15)
Average	0.74	0.15	0.11
	(0.21)	(0.15)	(0.14)

[21]The minimum possible ASE is obviously 0, while the maximum possible ASE is 200, corresponding, for example, to stating beliefs $(10, 0, 0)$ while the frequency of actions taken were $(0, 10, 0)$. Looking at the average squared error in each of the 16 tables, we do not observe clear differences. Subjects on average were most accurate in decision table 6 in Figure 1 (ASE = 13.93), while they were most imprecise in decision table 15 in Figure 1 (ASE = 26.45).

to others choosing the selfish action, which, as Table 1 showed, is right. However, Table 6 suggests that different preferences types have different beliefs. In particular, SF subjects assign much higher frequency (89%) to other subjects taking the selfish action than do other types (55%, 64%, and 61% for SW, IA, and CP types, respectively). Additionally, SW subjects are the ones who assign the highest frequency (31%) to the surplus-creating action, while CP subjects are the ones who assign the highest frequency (26%) to the surplus-destroying action. This result also points in the direction of the existence of "false-consensus bias." This is a regularity found in psychology literature as well as in economics, which describes the fact that individuals tend to believe that others are more likely to be like themselves, that is, in our experiment, they would assign high frequency to other subjects taking the same actions as they themselves took.[22] On average, subjects assigned a frequency of 0.68 to their own actions being taken. SF subjects assign the highest frequency (0.87) to others taking their own actions, while SW subjects assign the lowest frequency (0.46) to their own actions being taken. IA and CP subjects assign frequencies 0.57 and 0.55, respectively, to their own actions being taken.

However, averaging beliefs of those individuals classified into a preferences type can be misleading since it imposes the assumption that all individuals who belong to a type according to their actions should have similar beliefs. Since this is, in fact, our research question, we opted for a different strategy. We take a purely empirical strategy to identify experimental subjects' beliefs about the actions of other individuals. Furthermore, we empirically test whether individuals classified as belonging to different preferences types actually have different beliefs about others' actions.

We follow a mixture-of-types model, using the elicited belief data, to identify belief types, as well as the frequencies associated with each of the belief types. The elicited belief data consist of a probability distribution over the three available actions, $a = \{S, C, D\}$, for each of the T decision tables and each of the N individuals. The typical elicited belief observation is given by (Seb_{it}, Ceb_{it}), where Seb_{it} and Ceb_{it} represent the frequencies decider i associates to observing the selfish and surplus-creating actions at decision table t. Notice that the belief about the surplus-destroying action is given by 1 minus the beliefs about selfish and surplus-creating actions. For example, if decider i states that half of the 10 participants in the other group of deciders chose the selfish action and the other half chose the surplus-creating action, then the elicited belief observation will take the values $(0.5, 0.5)$.

When applying a mixture-of-types model to the analysis of beliefs, we have to make some specification decisions. First, we need to address what the specification of a belief type is. We consider two different belief-type specifications, depending on whether the relative position of a subject matters (or not) for belief statements. Our unrestricted specification defines a belief type as two different probability distributions over selfish, surplus-creating, and surplus-destroying actions: one when the decider is better off than the receiver and another when the decider is worse off. This specification thus

[22]False consensus bias was first mentioned by psychologists (Ross, Greene, and House (1977), and Mullen et al. (1985)). Economists have also found evidence of it; see Selten and Ockenfels (1993) and Charness and Grosskopf (2001). Engelmann and Strobel (2000) defined a *real* false consensus effect as weighting own decisions more heavily than those of a randomly selected person from the same population.

separates the elicited beliefs about others' actions into two different sets, depending on the decider's relative position (rp), $rp = \{A, B\}$, which we name A and B, *which* refer to ahead and behind, respectively. The typical belief type k will then be given by $(SbA_k, CbA_k, SbB_k, CbB_k)$. The restricted specification defines a belief type as a distribution over selfish, surplus-creating, and surplus-destroying actions, without differentiating for the decider's relative position. The typical belief type is then given by (Sb_k, Cb_k). The decision about whether the data fit one specification better than the other will be taken using a likelihood ratio test.[23]

The second question we need to address is how many belief types we should consider. We took a conservative position and started by allowing for only one belief type, which yields exactly the average beliefs in the subject population. We then added types one by one until the explanatory power of adding one more type was offset by the increased number of parameters to be estimated. For the decision over the number of belief types, we again used likelihood ratio tests. The restricted model refers to the specification with $(k-1)$ belief types and the unrestricted model refers to the specification with (k) belief types.

The likelihood function for the k different belief types in the specification where the decider's relative position matters is

$$\begin{aligned} &L(p_k, SbA_k, CbA_k, SbB_k, CbB_k | Seb, Ceb) \\ &\quad = \prod_{i=1}^{N} \sum_{k=1}^{K} p_k \prod_{rp \in \{A,B\}} \prod_{t=1}^{T/2} Sbrp_k^{Sebrp_{it}} Cbrp_k^{Cebrp_{it}} \\ &\qquad \times (1 - Sbrp_k - Cbrp_k)^{(1 - Sebrp_{it} - Cebrp_{it})}. \end{aligned} \tag{4}$$

A belief type is given by $(SbA_k, CbA_k, SbB_k, CbB_k)$ and p_k refers to the frequency of the kth belief type. Observations are counted separately when the decider is ahead and behind the receiver. That is, the 16 decision tables will be divided into two sets of 8, depending on the decider's relative position represented by rp.

The likelihood function for the k different belief types in the restricted belief-type specification, where the decider's relative position does not matter, is

$$\begin{aligned} &L(p_k, Sb_k, Cb_k | Seb, Ceb) \\ &\quad = \prod_{i=1}^{N} \sum_{k=1}^{K} p_k \prod_{t=1}^{T} Sb_k^{Seb_{it}} Cb_k^{Ceb_{it}} (1 - Sb_k - Cb_k)^{(1 - Seb_{it} - Ceb_{it})}. \end{aligned} \tag{5}$$

Now, the belief type is given by (Sb_k, Cb_k) and, as before, p_k refers to the frequency of the kth belief type. Also, the actual elicited beliefs are given by (Seb_{it}, Ceb_{it}), but now the observations will not be separated for when the decider is ahead or behind the receiver.

[23]Notice that we could consider other belief specifications. The least restricted one would consider as a belief type an action profile for each of the available actions in each of the decision tables. These alternative belief specifications are rejected by likelihood ratio tests.

The estimated belief types are summarized in Table 7. The first block of columns, Models 1–5, shows the simpler belief-type specification when the decider's relative position does not matter, while the second block of columns, Models 6–10, shows the belief-type specification when the decider's relative position does matter. The difference between Models 1–5, as well as the difference between Models 6–10, is the number of allowed belief types, which changes from 1 up to 5 belief types. Therefore, horizontally we can compare the two different specifications of belief types by keeping the number of types fixed, while vertically we can compare what we gain when we allow for heterogeneity within each belief-type specification. Likelihood ratio tests guide our decidision over the two different specifications, as well as over the number of types. Using a conservative criterion, we reject the restricted model at the 5% significance level. As becomes clear in Table 7, likelihood ratio tests persistently favor the belief-type specification where the decider's relative position is not taken into account. The only exception is when we compare Models 1 and 6. When Models 2 and 7, 3 and 8, 4 and 9, and finally 5 and 10 are compared, the likelihood ratio tests cannot reject the restricted model—the simpler belief-type specification—with p-values of 0.09, 0.12, 0.89, and 0.86, respectively. Also, when deciding about how much heterogeneity to allow for, that is, about the number of belief types to consider, a conservative reading of likelihood ratio tests favors including up to four different belief types but not the fifth one.[24] That is, the likelihood ratio test favors Model 4, which will be our focus.

According to Model 4, 24% of the subject population believes that others can only take the selfish action (with 99% frequency). A second belief type, held by 32% of the subjects, assigns very high frequency to the selfish action (80%), but also assigns some weight to the surplus-creating action (13%) and the least weight to the surplus- destroying action (7%). A third belief type, held by 12% of the subjects, assigns most of the weight to the selfish action (59%) and to the surplus-creating action (40%) while not conceiving that others might take the surplus-destroying action. Finally, 32% of the subjects believe that half of the others will take the selfish action, then the surplus destroying-action (28%) and, the least frequently, the surplus-creating action (0.19).

These four belief types represent different views about what others do. Given the actual frequencies of actions observed in Part 1 of the experiment (72% of selfish actions, 20% surplus creating, and 8% surplus destroying), the second belief type is the most accurate, followed by the third and fourth types, while the first type is the most inaccurate.[25]

Once we have selected Model 4, where there are four belief types, we can classify each individual into different identified belief types. This can be done with a likelihood function or even following a mean squared error criterion so that each individual is classified into the belief type from which her elicited beliefs deviate the least. Both methods

[24]When Models 3 and 4 are compared, the unrestricted model is favored (p-value 0.0003), suggesting it is worth considering a fourth type, but when Models 4 and 5 are compared, the restricted model is favored (p-value 0.0658), suggesting it is not worth allowing for a fifth belief type at the 5% significance level.

[25]Calculating the ASE for subjects classified under each of the four beliefs types, we find that subjects classified under belief type 1 incurred 3.26% of the maximum ASE they could have made. Subjects under belief type 2 made an ASE of 0.5% of the maximum, while subjects under belief type 3 made an ASE of 1.88% of the maximum. Finally, subjects classified under belief type 4 made an ASE of 2.18%.

TABLE 7. Belief-type identification for different specifications.

No. of Types	Specification 1						Specification 2								
	Model	p_k	Sb_k	Cb_k	Db_k	LL	Model	p_k	SbA_k	CbA_k	DbA_k	SbB_k	CbB_k	DbB_k	LL
$K=1$	1	–	0.73	0.15	0.11	−1464.4	6	–	0.72	0.17	0.10	0.74	0.13	0.12	−1460.8
							Restricted: Model 1								
							Unrestricted: Model 6		p-value = 0.0273						
$K=2$	2	0.39	0.95	0.04	0.02	−1360.8	7	0.39	0.94	0.04	0.02	0.95	0.03	0.01	−1356.8
		0.61	0.59	0.23	0.18			0.61	0.58	0.26	0.16	0.61	0.20	0.20	
	Restricted: Model 1						Restricted: Model 2								
	Unrestricted: Model 2		p-value = 0.0000				Unrestricted: Model 7		p-value = 0.0916						
$K=3$	3	0.38	0.95	0.03	0.02	−1331.5	8	0.39	0.94	0.04	0.02	0.95	0.03	0.02	−1326.5
		0.23	0.65	0.32	0.03			0.24	0.59	0.37	0.04	0.69	0.27	0.03	
		0.39	0.57	0.17	0.26			0.37	0.58	0.19	0.23	0.56	0.15	0.30	
	Restricted: Model 2						Restricted: Model 3								
	Unrestricted: Model 3		p-value = 0.0000				Unrestricted: Model 8		p-value = 0.1247						
$K=4$	4	0.24	0.99	0.00	0.00	−1322.2	9	0.24	0.99	0.00	0.00	0.99	0.00	0.00	−1320.4
		0.32	0.80	0.13	0.07			0.33	0.80	0.14	0.06	0.79	0.13	0.08	
		0.12	0.59	0.40	0.02			0.44	0.57	0.42	0.02	0.60	0.38	0.02	
		0.32	0.53	0.19	0.28			0.32	0.55	0.18	0.26	0.50	0.19	0.31	
	Restricted: Model 3						Restricted: Model 4								
	Unrestricted: Model 4		p-value = 0.0003				Unrestricted: Model 9		p-value = 0.8913						
$K=5$	5	0.24	0.99	0.00	0.00	−1318.6	10	0.24	0.99	0.00	0.00	0.99	0.00	0.00	−1316.6
		0.35	0.79	0.14	0.07			0.35	0.79	0.15	0.06	0.78	0.14	0.08	
		0.13	0.58	0.08	0.33			0.12	0.61	0.09	0.30	0.56	0.07	0.37	
		0.11	0.57	0.41	0.01			0.11	0.55	0.43	0.01	0.59	0.39	0.01	
		0.18	0.48	0.26	0.25			0.19	0.51	0.25	0.24	0.45	0.27	0.28	
	Restricted: Model 4						Restricted: Model 5								
	Unrestricted: Model 5		p-value= 0.0658				Unrestricted: Model 10		p-value= 0.8571						

Table 8. Frequency table for interdependent preferences types and belief types.

	Belief Types				
Preferences Types	Belief Type 1 (0.99, 0.00, 0.00)	Belief Type 2 (0.80, 0.13, 0.07)	Belief Type 3 (0.59, 0.40, 0.02)	Belief Type 4 (0.53, 0.19, 0.28)	Total
SF	28	18	4	5	55
SW	0	5	7	10	22
IA	1	7	6	10	24
CP	1	4	0	10	15
Total	30	34	17	35	116

give us the same classification. We can, therefore, proceed with a direct comparison between the classification of subjects by their actions (preferences-type classification in Part 1) and the classification of subjects by their beliefs.

Results are shown in Table 8. This contingency table shows the preferences-type classification by rows and the beliefs-type classification by columns. Each cell of the table contains the number of individuals classified as belonging to the preferences type represented by that particular row, who have the belief type represented by that particular column. We observe dependency between the row and the column classifications.[26] About half of the subjects classified as SF (28 out of 55) hold the first belief type. As such, these SF subjects can only conceive that other individuals take the selfish action. The other half of SF subjects hold the second belief type, assigning the highest weight to the selfish action, but also conceiving that others might take surplus-creating and -destroying actions with very low frequency. Other types of subjects have more disperse beliefs. Subjects classified as SW are behind the second (5 out of 22), third (7 out of 22), and fourth belief types (10 out of 22). SW individuals cannot conceive that others will take only the selfish action such that none of these individuals holds the first belief type. IA subjects hold the most dispersed beliefs, where the most frequent belief type held by IA subjects (10 out of 24) is the fourth type, assigning the highest frequencies to selfish and surplus-destroying actions and the least to surplus-creating actions. Finally, a large majority of subjects classified as CP (10 out of 15) have the fourth type of belief, where they believe the selfish and surplus- destroying actions are the most frequent.

Overall, Table 8 can be summarized as showing that while everyone believes the selfish action is taken with the highest frequency, SF individuals think that others will rarely incur costly actions to either increase or decrease others' actions. The SW individuals, on the other hand, cannot conceive that others will take only the selfish action. The IA individuals hold the most diverse beliefs. Finally, most CP types assign the highest frequency to both selfish and surplus-destroying actions.

[26]Association measure tests such as Goodman and Kruskal's tau and uncertainty coefficients both yield (asymptotic) p-values lower than 0.001. A chi-squared test allows us to conclude that rows are not independent (p-value < 0.01). Pairwise Fisher exact tests inform us that the distribution of belief types of subjects classified as SF and any other type is significantly different at the 1% level. Comparisons of SW versus CP and of IA versus CP are also significantly different at the 5% level.

We conclude that individuals do not have a homogeneous and accurate perception of the existent heterogeneity in actions and, therefore, in preferences. Therefore, we reject the null hypothesis of a common prior distribution among subjects. Furthermore, given the dependence found between preferences-type and belief-type classifications, different preferences types hold different beliefs about the heterogeneity in actions. While all subjects are affected by self-consensus bias, selfish subjects conceive the least heterogeneity in others' actions. In particular, half of the subjects classified under SF cannot conceive of others doing any other action than the selfish action. Other preferences types are partially affected by false-consensus bias, but believe others may take different actions than they take.

6. Results in Part 3 of the experiment: Estimation of the distribution of interdependent preferences types after elicitation of beliefs and observation of other participants' actions

In this section, we study the robustness of preferences types to repeated play after belief elicitation and the observation of others' actions.[27] In Part 3 of the experiment, subjects again make their choices over the three available actions in the 16 decision tables. Half of the subjects make their choices after having had their beliefs elicited and after having observed what the 10 participants of the other group of deciders in their same session actually chose in Part 1 of the experiment (control treatment, repeated play with social information). The other half make their choices after having had their beliefs elicited but without any information on others' choices (repeated play without social information). We look at how many subjects actually change preferences type from Part 1 to Part 3, toward which preferences type these switches occurred, and whether these changes are a consequence of social information.

We start by reclassifying subjects according to their actions in the third part of the experiment. We redo the interdependent preferences-type estimation, as explained in Section 4, for the third part of the experiment. Table A.2, available in the supplementary material, presents the individual by individual estimation and Table 9 summarizes the preferences-type distribution, both when social information is provided and when it is not. The first thing to notice in Table A.2, compared to the estimation in Part 1 shown in Table A.1, is that the noise level decreases considerably. In Part 3, almost 60% of subjects (67 out of 120) are estimated to belong to a preferences type without any error in contrast to the 38% of subjects in Part 1. These noiseless subjects' preferences types are readable directly from their actions, which are summarized in the first six columns of Table A.2. The second important thing to notice is that, as shown in Table 9, the preferences-type distribution changes slightly toward a distribution where the SF preferences type is even more prominent compared to the distribution in Part 1. An estimated 55% of subjects have SF preferences and each of the rest of the preferences types accounts for about 15% of the population. Finally, in Table 9, we see that both with and without social information, the distribution of preferences types is quite similar. This already suggests

[27]Ruström and Wilcox (2009) in a strategic environment showed that eliciting beliefs in itself, without the provision of any information, can modify individual behavior.

Table 9. Summary of individual by individual estimation (Part 3).

	Repeated Play With Social Information			Repeated Play Without Social Information		
	Number of Individuals	Frequency	$\bar{\varepsilon}_k$	Number of Individuals	Frequency	$\bar{\varepsilon}_k$
SF	32	0.55	0.04	32	0.54	0.03
SW	9	0.16	0.17	8	0.14	0.24
IA	11	0.19	0.26	9	0.15	0.21
CP	6	0.10	0.20	10	0.17	0.29
Total	58	1		59	1	

Table 10. Frequency of play separately for preferences types (Part 3) ($N = 117$ subjects).

				Decider's Position					
	Overall			Ahead			Behind		
Preference Type	*S*	*C*	*D*	*S*	*C*	*D*	*S*	*C*	*D*
				With Social Information					
SF	0.98	0.02	–	0.98	0.02	–	0.98	0.02	0.01
SW	0.22	0.78	–	0.19	0.80	–	0.25	0.75	–
IA	0.64	0.28	0.08	0.47	0.47	0.06	0.81	0.09	0.10
CP	0.55	0.01	0.44	0.60	0.02	0.38	0.5	–	0.5
				Without Social Information					
SF	0.98	0.02	–	0.98	0.02	–	0.99	0.01	0.01
SW	0.46	0.53	–	0.42	0.58	–	0.5	0.5	–
IA	0.58	0.35	0.07	0.42	0.57	0.01	0.75	0.13	0.13
CP	0.51	0.06	0.43	0.70	0.05	0.25	0.33	0.06	0.61
Total	0.76	0.16	0.07	0.74	0.21	0.05	0.78	0.12	0.10

that social information is not the main source for the change in preferences types, but having to choose a second time after having their beliefs elicited is already making some individuals change their behavior.

Table 10 shows, as Table 5 did for Part 1, the frequency of actions taken by each preferences type, both with and without social information. Comparing Table 5 and Table 10, we observe that the distribution of average frequencies with which overall the three actions were played (the last row in both tables) is practically identical in Parts 1 and 3 of the experiment. However, when we look at the average frequency with which each preferences type took each of the three actions in Part 3 of the experiment, we observe that preferences types are now more clearly separated. Again we find that SF types barely took any but the selfish action, no matter what their relative position and whether there was social information or not (98%). The SW type took the surplus- creating action with highest frequency (78% with social information; 53% without it) and while they also took the selfish action (22% with social information; 46% without it), they barely took the surplus-destroying action (only 7% did so in the control treatment without social information). The IA type mainly took selfish and surplus- creating actions when ahead (47%

and 47%, respectively, with social information; 42% and 57% without it), but they choose most frequently the selfish action and much less frequently the surplus-destroying action when behind (81% and 9%, respectively, with social information; 75% and 13% without it). Finally, the CP type almost never created surplus (1% with social information; 6% without it) and chose the selfish and surplus- destroying actions in similar percentages (55% and 44%, respectively, with social information; 51% and 43% without it).

We now check whether actions and preferences types were consistent between Part 1 and Part 3 of the experiment. As a first approximation, we find that subjects changed their action from Part 1 to Part 3 on average in 3.78 out of 16 decision tables (23.62% of the time). Subjects classified as SF according to their actions in Part 1 actually changed their actions the least (an average of 1.41 table changes or 8.83% of the time). The IA, SW, and CP subjects changed their actions more often on average, in 5.83, 5.87, and 6.2 out of 16 decision tables, respectively (36.46%, 36.69%, and 38.75%, respectively).[28]

More precisely, Table 11 presents two contingency tables, when social information was provided and when it was not, where rows refer to the preferences-type classification in Part 1 and columns refer to the preferences-type classification in Part 3. The diagonal cells of this table show the number of subjects who did not change preferences type from Part 1 to Part 3 of the experiment. Off-diagonal cells present the number of subjects who changed type from row preferences type to column preferences type. Overall, the majority of subjects (67% (38 out of 57) with social information and 68% (38 out of 56) without social information) did not change their preferences type from Part 1 to Part 3.[29]

Table 11. Interdependent preferences-type classification in Parts 1 and 3.

	Preferences Types Part 3				
Preferences Types Part 1	SF	SW	IA	CP	Total
Repeated Play With Social Information					
SF	**22**	0	0	1	23
SW	2	**8**	5	0	15
IA	4	1	**6**	3	14
CP	3	0	0	**2**	5
Total	31	9	11	6	57
Repeated Play Without Social Information					
SF	**25**	4	0	2	31
SW	1	**3**	2	1	7
IA	4	1	**4**	0	9
CP	2	0	1	**6**	9
Total	32	8	7	9	56

[28]The SF types changed their action an average of 1.5 tables with social information and 0.79 without it. Averages for other preference types were higher (6.85 and 5.4 for SW, 6.11 and 5.66 for IA, and 8.12 and 3.66 for CP, respectively, for sessions with and without social information).

[29]The kappa test, a chance-corrected measure of agreement between two classifications, yields values of 0.512 and 0.481 for with and without social information, respectively. Therefore, we conclude that agree-

The numbers in the diagonal cells are always higher than in any other cell, except for the competitive individuals with social information and the inequity averse individuals without social information. Consistently, subjects who changed type changed their actions an average of 6.51 actions from Part 1 to Part 3 (40.7%), while subjects who did not changed type changed their actions an average of 2.18 actions (13.65%).

Moreover, there are significant differences if we compare the likelihood of changing types across different rows and, therefore, across different preferences types. Consistent with the finding in changes on actions, subjects estimated as having SF preferences are the ones who least changed their preferences type from Part 1 to Part 3, both with and without social information. Only 1 subject out of 23 actually switched type when social information was provided and only 6 out of 31 when social information was not provided. On the other hand, almost half of the subjects estimated as having SW and IA preferences in Part 1 actually switched type, both when social information was provided and when it was not. Finally, as for CP individuals, when information was provided, they did change considerably, but not so when no information was provided. We can conclude that while SF preferences are very stable SW and IA show less stability, both when social information is provided and when it is not. It seems that reviewing actions after having elicited their beliefs, that is, thinking what other individuals in similar situation have done, is already making these preferences types unstable. The CP preferences types show more stability when no information is provided than when it is.[30] If we order preferences types with respect to a decreasing level of altruism (SW, IA, SF, CP), Table 11 shows that the majority of the subjects who changed type (25 out of 37) moved from a more altruistic type to a less altruistic one.

Overall, we conclude that almost 70% of subjects did not change preferences type, both when social information was provided and when it was not. Therefore, interdependent preferences types seem to show robustness to repeated play after belief elicitation and also to the provision of social information.[31] However, there are important differences across different preferences types. While selfish subjects hardly change type, other preferences types show much less stability. Our results are consistent with Cason and Mui (1998), who showed, in a twice-repeated regular dictator game in which dictators are informed of a previous choice by another single dictator, that subjects who are more self-regarding on their first decisions are less likely to change choices between their first and second decisions. Notice that this result seems intuitive to some extent: SF individuals who do not care about others' payoffs are not affected by making them think about what others have done or by others' actions, while other- regarding individuals (SW, IA, or CP) are more affected by belief elicitation of others' actions and others' actual choices. Finally, we cannot conclude that social information is the crucial element, since repeat-

ment exists between both classifications. In any case, this value partially comes from the high proportion of subjects who were consistently classified as SF in both Part 1 and Part 3 of the experiment.

[30]Pairwise Fisher exact tests that compare the classification in Part 3 of the experiment with the subjects classified under the four types in Part 1 allow us to conclude that significant differences exist between the SF and the SW types at the 1% level and between the SF and the IA types. Fisher tests also show a significant difference between the SW and the CP types at the 1% level.

[31]This robustness is compatible with previous evidence by Brandts and Fatás (2001), who observed little indication of social influence in a public good game and, thus, in a strategic environment.

ing play after belief elicitation, that is, making individuals think about others' actions, is already affecting other-regarding individuals' behavior.

7. Conclusions

This paper studies the role *beliefs* and the knowledge about others' distributional decisions, *social information*, play in interdependent preferences.

When we elicit beliefs about others' actions in a purely decision making environment, such that beliefs only capture uncertainty about others' preferences (with no strategic uncertainty), we find that different interdependent preferences types hold different beliefs about others' actions and, therefore, preferences. In particular, while selfish individuals hardly believe others incur personal costs to create or destroy surplus, individuals with interdependent preferences are aware that others who take different actions may exist. Therefore, we have found a strong correlation between preferences types and their beliefs.

Regarding social information, although interdependent preferences assume individuals care only about others' payoffs, we show that in a nonstrategic setting, although the majority of subjects show robustness in their type classification, repeated play, with and without social information, affects different types very differently. Selfish types barely change their behavior when making choices a second time, both when social information is provided and when it is not, while almost half of social welfare maximizers and inequity averse individuals actually change their type. Competitive individuals are shown to be most sensitive to repeated play and information. Overall, we show that those individuals with interdependent preferences who change their behavior tend to behave more selfishly.

The experimental results reported in this paper show that interdependent preferences that only include payoff differences among players might be too limited to capture other-regarding preferences, because beliefs, despite not directly affecting own payoffs, might actually play an important role in determining behavior. Providing an explanation for the relationship found between preference types and beliefs is a crucial step in the research agenda on other-regarding preferences.

Appendix: Experimental instructions

Below you can find a translation of the experimental instructions which were handed to Deciders sequentially and read aloud before each part. A summary of these instructions appeared on subjects' screens before each part.

Instructions read to all subjects ("Deciders" and "Receivers")

THANK YOU FOR PARTICIPATING IN OUR EXPERIMENT!

This is an experiment and thus, no talking, looking-around or walking is allowed. If you have any question or need help please raise your hand and one of the researchers will assist you. If you do not follow the indicated rules, WE WILL ASK YOU TO LEAVE THE EXPERIMENT AND YOU WILL NOT RECEIVE ANY PAYMENT. Thank you.

This experiment is about individual decisions. Both Pompeu Fabra and Autònoma de Barcelona universities have provided funds to carry it out. *You will receive* 3 *euros for having arrived on time*. Additionally, if you follow the instructions correctly you may earn more money.

The experiment has three parts. Before each part, we will let you know about the tasks you have to do and how your decisions will affect your payments. Everything you earn will be for you and paid in cash inside a closed envelope in a strictly private way at the end of the experimental session.

Each participant has a strictly confidential "Experiment Code" to guarantee that no participant can identify another one by his/her decisions nor earnings. Researchers will observe each participant's earnings at the end of the experiment but we will not associate your decisions with any participants' names.

Your Experiment Code is: XXXXX

The experiment consists of three parts. Your final payment will be the sum of a participation fee of the 3 euros plus whatever you earn in the three parts of the experiment.

Each experimental point corresponds to 25 Euro cents.

Thus, if you obtain a total of 32 points, you will receive a total of 11 euros (3 for participating and 8 from converting 32 experimental points into euros at a rate of 4 experimental point $*0.25 = 1$ Euro).

If, for example, you obtain 10 experimental points, you will receive 5.5 Euros ($10 * 0.25 = 2.5 + 3 = 5.5$).

If, for example, you obtain 70 experimental points, you will receive 20.5 Euros ($70 * 0.25 = 17.5 + 3 = 20.5$).

There are 40 participants in this experiment, 20 in the laboratory to whom we refer to as "Deciders" and 20 in a classroom to whom we refer to as "Receivers."

As you have observed, who is a "Decider" (and stayed in the laboratory) and who is a "Receiver" (and went to the classroom) has been randomly decided by extracting a paper from a bag.

"Deciders" take decisions which affect their payments and the payments of other participants in the experiment. "Receivers" do not take any decision, which affect neither their payments nor those of other participants in the experiment. When the experiment concludes, we will first pay "Deciders" in private. Once "Deciders" have left, "Receivers" will come to the laboratory and will be paid in private.

The 20 "Deciders" have been divided in two groups of 10 subjects each: "group A" and "group B."

You belong to Group A (B). If you are a "Receiver" you do not belong to any group.

PART 1 is about to start. Please wait until everyone has read these instructions.

Instructions for Deciders' Task 1. PART 1

In this part of the experiment we are going to show you 16 tables. The 16 tables the computer will show you will look as follows:

	Option 1	Option 2	Option 3
Decider	8	7	11
Receiver	17	19	13

In each of the tables you must choose between "Option 1," "Option 2" and "Option 3." Each of these 3 options describes how many experimental points a participant ("Decider") receives and how many another randomly matched participant ("Receiver") gets.

At no time a participant will know who they are matched with in any table.

When the experiment is over, the computer will randomly choose one of the 16 tables to determine the payments for PART 1.

You will receive the amount of experimental points corresponding to "Decider" in the chosen table and your matched participant will receive the number of experimental points corresponding to "Receiver" in the same table.

For example, if the chosen table was the one that appears above and you had chosen "Option 2," you would obtain 7 experimental points while your matched participant would obtain 19 experimental points.

Notice that the numbers in the example are just for illustrative purposes. They DO NOT intend to suggest how anyone may choose among the different options.

Participants in the other classroom ("Receivers") cannot take any decision which may affect your payments or their payments.

What you earn and what your matched participant ("Receiver") earns depends only on your decisions and on the randomly chosen table.

Once you have chosen your option in a particular table, please press "OK" and wait for the other participants to make their choice before moving to the next table.

Instructions for Deciders' Task 2. PART 2

In this part of the experiment the computer will show you the same 16 tables you saw in PART 1, although the tables may appear in a different order than before.

Remember that we have divided the 20 participants in the experiment in two groups of 10 people ("group A" and "group B"). In the first part of the experiment all "Deciders" have chosen among the three options having as a matched participant another subject from the other room ("Receivers").

Now you will have to guess how many out of the 10 "Deciders" from the other group ("group A"/"group B") have chosen each option ("Option 1," "Option 2," and "Option 3") in each of the 16 tables in PART 1 of the experiment.

For example, in one of the tables you may write:

Option 1: 6

Option 2: 3

Option 3: 1

This would mean that you think that in this particular table, 6 out of the 10 participants in Group B (A), chose "Option 1," 3 chose "Option 2," and 1 chose "Option 3."

Notice that the numbers in the example are just for illustrative purposes. They DO NOT intend to suggest how anyone may choose among the different options.

When the experiment is over, the computer will randomly choose one of the 16 tables to make payments for PART 2. You will receive more money the closer your guesses are to what participants from Group B (A) actually chose in PART 1.

You will be paid according to the mathematical formula which appears below. Do not worry if you do not understand the formula exactly. What is important is that you understand that the closer the numbers you write to the number of participants who actually chose each option the more money you will receive.

For example, if you write that 6 participants choose "Option 1" and actually 6 participants chose "Option 1," you will receive more money than if 5 or 7 participants chose "Option 1."

Notice that in this part of the experiment your answer can only affect your payments, and not those of any other participant, either from your group or the other group.

Here is the formula:

$$\text{Experimental Points in PART 2} = 20 - 0.1 * [(a - X)^2 + (b - Y)^2 + (c - Z)^2],$$

where:

a: Number of participants you think choose "Option 1"

b: Number of participants you think choose "Option 2"

c: Number of participants you think choose "Option 3"

X: Number of participants who actually chose "Option 1"

Y: Number of participants who actually chose "Option 2"

Z: Number of participants who actually chose "Option 3"

Please read the following examples to see how the formula works:

– In one table, you write that 6 participants choose "Option 1," 3 participants "Option 2" and 1 participant "Option 3." If, in fact 6 participants chose "Option 1," 3 participants "Option 2" and 1 participant "Option 3" you will obtain:

$$\text{Experimental Points in PART 2} = 20 - 0.1 * [(6 - 6)^2 + (3 - 3)^2 + (1 - 1)^2] = 20.$$

– In one table, you write that 2 participants choose "Option 1," 4 participants "Option 2" and 4 participants "Option 3." If, in fact in that Table 8 participants chose "Option 1," 2 participants "Option 2" and 0 participants "Option 3" you will obtain:

$$\begin{aligned}&\text{Experimental Points in PART 2}\\&\quad = 20 - 0.1 * [(2 - 8)^2 + (4 - 2)^2 + (4 - 0)^2] = 14.4.\end{aligned}$$

– In one table, you write that 0 participants choose "Option 1," 10 participants "Option 2" and 0 participants "Option 3." If, in fact in that Table 10 participants chose "Op-

tion 1," 0 participants "Option 2" and 0 participants "Option 3" you will obtain:

$$\text{Experimental Points in PART 2} = 20 - 0.1 * [(0-10)^2 + (10-0)^2 + (0-0)^2] = 0.$$

Notice that the numbers in the example are just for illustrative purposes. They DO NOT intend to suggest how anyone may choose among the different options.

Examples show that with this formula you will never lose experimental points in PART 2, and that you can obtain up to 20 experimental points in PART 2. You will earn more money the closer your guesses are to the number of participants who actually chose each option.

Once you have entered your guess in a particular table, you should press "OK" and wait for the other participants to make their guesses before moving to the next table.

Instructions for Deciders' Task 3. PART 3

In this final part of the experiment the computer will show you for the last time the 16 tables you have already seen, although they might be in a different order. For each table, you are matched with a participant from the other classroom ("Receiver") randomly chosen and different from the one in PART 1.

Your task will consist once again in deciding between the three options ("Option 1," "Option2" and "Option 3") as you did in PART 1.

The way you ("Decider") and your matched participant ("Receiver") will earn experimental points is the same as in PART 1 of the experiment. Your payments only depend on your decisions and on the randomly chosen table by the computer at the end of the experiment.

[*Experimenters' note: In the control treatment without Social Information this paragraph was omitted*]

The only novelty you will find is that when you now observe each of the tables you will see how many of the other group of "Deciders" actually chose each option ("Option 1," "Option2" and "Option 3") in PART 1 of the experiment.

Once you have chosen your option in a particular table, please press "OK" and wait for the other participants to make their choice before moving to the next table.

Thank you very much for your participation.

Anonymous questionnaire filled by all participants

My Experiment Code is: _____

1. What do you think about the experiment?
2. How did you make your choices in each part of the experiment?
3. How do you think others made their choices in each part of the experiment?
4. Are you satisfied with your earnings in the experiment?
5. Gender.
6. Age.

7. What do you study?
8. Would you like to add any other comment?

References

Akerlof, G. A. (1982), "Labor contracts as partial gift exchange." *The Quarterly Journal of Economics*, 97 (4), 543–569. [517]

Andersen, S., J. Fountain, G. Harrison, and E. Rutström (2007), "Eliciting beliefs: Theory and experiments." Working paper 07–08, Department of Economics, College of Business Administration, University of Central Florida. [522]

Andreoni, J. and J. Miller (2002), "Giving according to GARP: An experimental test of the consistency of preferences for altruism." *Econometrica*, 70, 737–753. [516, 517, 518, 528]

Bellemare, C., S. Kröger, and A. van Soest (2008), "Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities." *Econometrica*, 76 (4), 815–839. [517]

Bernheim, D. (1994), "A theory of conformity." *Journal of Political Economy*, 102 (5), 841–877. [517]

Blanco, M., D. Engelmann, and H. Normann (2011), "A within-subject analysis of other-regarding preferences." *Games and Economic Behavior*, 72 (2), 321–338. [516]

Bolton, G. and A. Ockenfels (2000), "ERC: A theory of equity, reciprocity and competition." *American Economic Review*, 90 (1), 166–193. [516]

Brandts, J. and E. Fatás (2001), "Social information and social influence in an experimental dilemma game." Report. [538]

Cabrales, A., R. Miniaci, M. Piovesan, and G. Ponti (2010), "Social preferences and strategic uncertainty: An experiment on markets and contracts." *American Economic Review*, 100, 2261–2278. [517]

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004), "A cognitive hierarchy model of one-shot games." *The Quarterly Journal of Economics*, 119 (3), 861–898. [524]

Cason, T. and V. Mui (1998), "Social influence in the sequential dictator game." *Journal of Mathematical Psychology*, 42, 248–265. [517, 538]

Charness, G. and M. Dufwenberg (2006), "Promises and partnership." *Econometrica*, 74 (6), 1579–1601. [522]

Charness, G. and B. Grosskopf (2001), "Relative payoffs and happiness: An experimental study." *Journal of Economic Behavior and Organization*, 45, 301–328. [530]

Charness, G. and M. Rabin (2002), "Understanding social preferences with simple tests." *The Quarterly Journal of Economics*, 117 (3), 817–869. [516]

Costa-Gomes, M. A. and V. P. Crawford (2006), "Cognition and behavior in two-person guessing games: An experimental study." *American Economic Review*, 96, 1737–1768. [524]

Costa-Gomes, M. A., V. P. Crawford, and B. Broseta (2001), "Cognition and behavior in normal-form games: An experimental study." *Econometrica*, 69, 1193–1235. [524, 525]

Costa-Gomes, M. and G. Weizsäcker (2008), "Stated beliefs and play in normal form games." *Review of Economic Studies*, 75, 729–762. [517]

Crawford, V. P. and N. Iriberri (2007a), "Fatal attraction: Salience, naivete, and sophistication in experimental hide-and-seek games." *American Economic Review*, 97 (5), 1731–1750. [524]

Crawford, V. P. and N. Iriberri (2007b), "Level-k auctions: Can a non-equilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions?" *Econometrica*, 75 (6), 1721–1770. [525]

Croson, R. (2000), "Thinking like a game theorist: Factors affecting the frequency of equilibrium play." *Journal of Economic Behavior and Organization*, 41, 299–314. [517]

Croson, R. and J. Shang (2008), "The impact of downward social information on contribution decisions." *Experimental Economics*, 11, 221–233. [517]

Duffy, J. and T. Kornienko (2010), "Does competition affect giving?" *Journal of Economic Behavior and Organization*, 74, 82–103. [517]

Dufwenberg, M. and U. Gneezy (2000), "Measuring beliefs in an experimental lost wallet game." *Games and Economic Behavior*, 30, 163–182. [517]

El-Gamal, M. and D. Grether (1996), "Are people Bayesian? Uncovering behavioral strategies." *Journal of the American Statistical Association*, 90 (432), 1137–1145. [524, 525]

Engelmann, D. and M. Strobel (2000), "The false consensus effect disappears if representative information and monetary incentives are given." *Experimental Economics*, 3 (3), 241–260. [530]

Engelmann, D. and M. Strobel (2004), "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments." *American Economic Review*, 94 (4), 857–869. [516]

Fehr, E. and K. Schmidt (2000), "Theories of fairness and reciprocity: Evidence and economic applications." In *Advances in Economic Theory. Eight World Congress of the Econometric Society* (M. Dewatripont, L. Hansen, and S. Turnovsky, eds.), Cambridge University Press, Cambridge. [515]

Fehr, E. and K. M. Schmidt (1999), "A theory of fairness, competition, and cooperation." *The Quarterly Journal of Economics*, 114 (3), 817–868. [516]

Fischbacher, U. (2007), "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10 (2), 171–178. [519]

Fisman, R., S. Kariv, and D. Markovits (2007), "Individual preferences for giving." *American Economic Review*, 97 (2), 153–158. [516, 517, 518, 528]

Frey, B. and S. Meier (2004), "Social comparisons and pro-social behavior: Testing 'conditional cooperation' in a field experiment." *American Economic Review*, 94 (5), 1717–1722. [517]

Gächter, S. and E. Renner (2010), "The effects of (incentivized) belief elicitation in public good experiments." *Experimental Economics*, 13 (3), 364–377. [517]

Greiner, B. (2004), "The online recruitment system ORSEE 2.0—A guide for the organization of experiments in economics." WP Series in Economics 10, University of Cologne. [519]

Harless, D. and C. Camerer (1994), "The predictive utility of generalized expected utility theories." *Econometrica*, 62, 1251–1290. [524]

Huck, S. and G. Weizsäcker (2001), "Do players correctly estimate what others do? Evidence of conservatism in beliefs." *Journal of Economic Behavior and Organization*, 3, 367–388. [522]

Iriberri, N. and P. Rey-Biel (2011), "The role of role uncertainty in modified dictator games." *Experimental Economics*, 14 (2), 160–180. [520]

Ivanov, A. (2011), "Attitudes to ambiguity in one-shot normal-form games: An experimental study." *Games and Economic Behavior*, 71, 366–394. [517]

Jones, S. (1984), *The Economics of Conformism*. Blackwell Publishing, Oxford. [517]

Kerschbamer, R. (2012), "The geometry of distributional preferences and a nonparametric identification approach." Unpublished manuscript, University of Innsbruck. [525]

Mullen, B., J. Atkins, D. Champion, C. Edwards, D. Hardy, J. Story, and M. Venderklok (1985), "The false consensus effect: A meta-analysis of 115 hypothesis tests." *Journal of Experimental Social Psychology*, 21, 263–283. [530]

Nyarko, Y. and A. Schotter (2002), "An experimental study of belief learning using real beliefs." *Econometrica*, 70 (3), 971–1005. [517]

Offerman, T. and J. Sonnemans (2001), "Is the quadratic scoring rule behaviorally incentive compatible?" Report, University of Amsterdam. [522]

Rey-Biel, P. (2009), "Equilibrium play and best response to (stated) beliefs in normal form games." *Games and Economic Behavior*, 65 (2), 572–585. [517]

Ross, L., D. Greene, and P. House (1977), "The 'false consensus effect': An egocentric bias in social perception and attribution processes." *Journal of Experimental Social Psychology*, 13, 279–301. [530]

Rutström, E. and N. Wilcox (2009), "Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test." *Games and Economic Behavior*, 67, 616–632. [535]

Selten, R. and A. Ockenfels (1993), "An experimental solidarity game." *Journal of Economic Behavior and Organization*, 34, 517–539. [530]

Sobel, J. (2005), "Interdependent preferences and reciprocity." *Journal of Economic Literature*, 43, 392–436. [515]

Stahl, D. and P. Wilson (1994), "Experimental evidence on players' models of other players." *Journal of Economic Behavior and Organization*, 25, 309–327. [524]

Stahl, D. and P. Wilson (1995), "On players' models of other players: Theory and experimental evidence." *Games and Economic Behavior*, 10, 218–254. [524]

Submitted January, 2011. Final version accepted November, 2012.