Social Preferences and Strategic Uncertainty:

An Experiment on Markets and Contracts *

Antonio Cabrales

Universidad Carlos III de Madrid

Raffaele Miniaci

Università di Brescia

Marco Piovesan

University of Copenhagen

Giovanni Ponti[†]

Universidad de Alicante

and Università di Ferrara

Abstract

This paper reports a 3-phase experiment on a stylized labor market. In the first two phases, agents face simple games, which we use to estimate subjects' social and reciprocity concerns. In the last phase, four principals compete by offering agents a contract from a fixed menu. Then, agents "choose to work" for a principal by selecting one of the available contracts. We find that *i)* (heterogeneous) social preferences are significant determinants of choices, *ii)* for both principals and agents, *strategic uncertainty aversion* is a stronger determinant of choices than *fairness* and *iii)* agents display a marked propensity to work for principals with similar distributional concerns.

KEYWORDS: Social Preferences, Team Incentives, Mechanism Design, Experimental Economics

JEL CLASSIFICATION: C90, D86

Many economists now accept that individuals have *social* (i.e. distributional) and/or *reciprocal* preferences. An important consequence of this is that they dislike inequality in rewards (Bewley, 1999). This consideration notwithstanding, Winter (2004) has shown that, in activities which display *strategic complementarities* (take, for example, Kremer's (1993) well-known "O-ring" production function), inequalities in rewards are *necessary* to implement the high-effort profile as the *unique* equilibrium of the game. This is because, "*if agents' exertion of effort induces a positive externality on the effectiveness of other agents' effort, it is optimal to promise high rewards to some agents so as to make the others confidently believe that these highly paid agents will contribute, hence allowing the planner to save resources by offering other agents substantially less*".

Winter's (2004) result abstracts from the existence of social preferences, but it adds an additional ingredient to the debate on inequality by showing that the principal faces a trade-off between fairness and robustness: *fairness can be obtained only at the expense of robustness to strategic uncertainty.*[1] In this respect, one can only expect this trade-off to be exacerbated by the presence of (inequality-averse) distributional preferences.

The aim of this paper is precisely to test experimentally the idea that workers' (heterogeneous) social preferences are crucial in determining the contracts they are offered and choose.[2] We are also interested in the way our experimental subjects resolve the trade-off between fairness and robustness, as they can choose either *i)* contracts in which -following Winter (2004)- the all-effort profile is the unique equilibrium, but inequality is enhanced; or *ii)* contracts in which the all-effort profile is not the unique equilibrium, but inequality is mitigated. In this respect, subjects more concerned with equity (and less worried about coordination failure) may find convenient to opt for the latter alternative. Finally, since another solution to the trade-off is *sorting* (agents with similar distributional concerns work for the same firms), this will also be an important element of our experimental design.[3]

With these goals in mind, we design and perform an experiment with three phases.

1. In the first phase (P_1), subjects are matched for 24 rounds with a different partner and

have to choose among four possible options involving a payoff pair -one for them, one for their matched partner- in a (Random) Dictator Game-type protocol. We use evidence from P_1 to estimate subjects' purely distributional preference parameters within the realm of Charness and Rabin's (2002, C&R hereafter) model.

2. In the second phase (P_2), subjects are again matched in pairs for 24 rounds and asked to choose among the same sets of payoff pairs. However, this time options correspond to *contracts*, as they yield a 2×2 effort game induced by Winter's (2004) technology, which subjects have then to play at a second stage. In P_2 reciprocity may play a role, since agents may condition their second-stage effort decision on their teammate's contract choice. Thus, we use P_2 to estimate subjects' C&R reciprocity parameters, together with their beliefs in the effort game.

3. Finally, in the third phase (P_3), there are 4 principals and 4 pairs ("teams") of agents. Principals offer a contract (a 2×2 game, such as those played in P_2) selected from the round choice set. The presence of several competing principals acts as a kind of menu of contracts, among which agents may sort themselves.

This three-stage experimental design (and the associated estimation strategy) is especially designed to solve the identification problem discussed by Manski (2002), as we use it to disentangle preference and belief parameters. Since in P_1 beliefs do not play any role, we use data from P_1 to identify distributional preference parameters and data from P_2 to estimate reciprocity concerns and beliefs.

We here summarize the main results of our study.

1. *Subjects display a significant degree of heterogeneity in their decisions, and thus, in estimated preferences and beliefs.*

2. *This heterogeneity explains, to a large extent, agents' behavior.* That is, preferences and beliefs which best explain agents' behavior in P_1 and P_2, also predict well their effort decisions in P_3.

3. We also observe that *equality is a less important concern than robustness*, for both principals and agents, since the egalitarian (but not robust) contract is rarely selected and, when it is selected, very often yields the (inefficient) low effort outcome. This, in turn, implies lower profits, for both principals and agents.

4. Finally, we find that *principals and agents sort themselves according with their social preferences*. An agent's probability of selecting a contract in P_3 decreases with the distance between her estimated preferences and those of the principal for whom she ends up working for. Moreover, principals usually offer contracts in tune with their own estimated distributional preferences.

The remainder of this paper is arranged as follows. Section 1 presents the experimental design, while in Section 2 we develop an econometric model to estimate distributional preferences and beliefs. Section 3 presents the experimental results and discusses our testable hypotheses. Final remarks are placed in Section 4. Three Appendices provide proofs, additional statistical evidence and the experimental instructions.

1   Experimental design

1.1   Sessions

Nine experimental sessions were conducted at the Laboratory of Theoretical and Experimental Economics (LaTEx), of the Universidad de Alicante. A total of 216 students (24 per session) was recruited among the undergraduate population of the Universidad de Alicante. The experimental sessions were computerized. Instructions were read aloud and we let subjects ask about any doubt they may have had.[4] In all sessions, subjects were divided into two *matching groups* of 12, with subjects from different matching groups never interacting with each other throughout the session.

1.2 Choice sets

Our experiment involves, for each one of the 24 rounds t constituting each phase, two subjects, 1 and 2, deciding over a set of four *options* $C_t = \{b_t^k\}, k = 1, ...4$. Each option constitutes a monetary payoff pair $b_t^k \equiv (b_{1t}^k, b_{2t}^k)$, with $b_{1t}^k \geq b_{2t}^k$ by construction. Each pair determines the payoff matrix of a simple 2×2 effort game, $G(k)$. The rules of $G(k)$ are as follows. Each agent $i = 1, 2$, has to decide, simultaneously and independently, whether to make a costly effort. We denote by $\delta_i \in \{0, 1\}$ agent i's effort decision, where $\delta_i = 1$ (0) if agent i does (not) make effort. Let also $\delta = (\delta_1, \delta_2)$ denote agents' action profile. Agent i's monetary payoff, $\pi_{it}^k(\delta)$, is described by

$$(1) \qquad \qquad \pi_{it}^k(\delta) = B + P(\delta)b_{it}^k - \delta_i c;$$

with

$$(2) \qquad \qquad P(\delta) = \begin{cases} 0 \text{ if } \delta_1 + \delta_2 = 0, \\ \gamma \text{ if } \delta_1 + \delta_2 = 1, \\ 1 \text{ if } \delta_1 + \delta_2 = 2, \end{cases}$$

where B is a fixed monetary prize independent on effort decision, c is the cost of effort and $\gamma \in (0, \frac{1}{2})$.[5] By (2), players receive their full payoff, b_{it}^k, if they both coordinate on the effort decision, and a fixed share of the latter if only one makes effort. If nobody makes effort (i.e. if $\delta_1 + \delta_2 = 0$), both agents only get B. In our experiment we fix $B = 40$, $c = 10$ and $\gamma = \frac{1}{4}$.

Put Figure 1 about here

In Figure 1 we report all payoff pairs b_t^k used in the experiment. As Figure 1 shows, these payoff pairs are drawn at random in the positive orthant, but not uniformly. Precisely, they are concentrated in two "clouds", which differ from one another by the fact that, for one of

4

them, player 1 (the "advantaged" player within the 2-member team) receives substantially more. As we explain in detail in Appendix A, these two clouds include pairs b_t^k which are the solutions of two different mechanisms design problems aimed at inducing both players to make effort. The two mechanism design problems differ in that

1. under the "*w*eak effort induc*ing*" solution (*wing* hereafter) players have a strict incentive to make effort only if the other does;

2. under the "*st*rong effort induc*ing*" solution (*sting* hereafter) player 1's payoff is sufficiently high to provide her with a strict incentive to make effort independently on what player 2 does, while player 2, like in the *wing* solution, has a strict incentive to make effort only if player 1 does.

This implies that, under the *sting* solution, the all-effort profile is the unique equilibrium of the induced game, while under the *wing* solution also the all-no-effort profile is an equilibrium.

Unlike Winter (2004), who focuses on Egoistic (i.e. non distributional) Preferences (EP), we solve the two mechanism design problems under a wide variety of distributional preferences analyzed by the literature. This explains the additional payoff variability within each cloud (where the larger points in each cloud identify the corresponding EP solutions).

The interested reader can find in Appendix A all the details. What is important to stress here is that our choice set provides sufficient variability in payoffs to estimate individual social preferences in Section 2, and that the specific variability we created (essentially, payoffs of similar magnitude for player 2, while a substantial difference in prizes in favor of player 1, depending on whether a *wing* or a *sting* solution is applied) allows us to quantify the discussion on the trade-off between equality and robustness we proposed earlier.[6]

Depending on the round t, the choice set C_t can be composed of *i)* 4 *wing* contracts; *ii)* 4 *sting*; or *iii)* 2 *wing* and 2 *sting*. We group rounds into *time intervals*. A time interval is defined as a group of three consecutive rounds: $\tau_p = \{3(p-1) < t \le 3p\}$, $p = 1, ..., 8$. Within

5

each time interval τ_p, subjects experienced each and every possible situation, *i)* to *iii)*. The particular sequence of three situations within each time interval was randomly generated. We did so to keep under control the time distance between two rounds characterized by the same situation.

1.3   Phases

Subjects played three *phases, P_1* to P_3, of increasing complexity, for a total of 72 rounds (24 rounds per phase).[7]

P_1. **Dictator Game (24 rounds).**   In this phase we use a variant of the classic protocol of the Dictator Game. The timing for each round t and matching group is as follows:

1. At the beginning of the round, six pairs are formed at random. Within each pair, another (independent and uniformly distributed) random device determines player position (i.e. the identity of the best paid agent).

2. Both agents choose their preferred option. In P_1, monetary payoffs associated to each option correspond to the all-effort profile payoff, $\pi_{it}^k(1,1)$.[8]

3. Once choices are made, another independent draw fixes the identity of the *Dictator*.

4. The Dictator's choice, k, determines monetary payoffs for that pair and round.

P_2 : **Effort Game (24 rounds).**   Stages 1 to 3 are identical to those of P_1. Instead of stage 4, we have

4 Subjects are asked to play the 2×2 effort game, $G(k)$, described in Section 1.2. Subjects' action profile determines their financial reward (1).

P_3 : **Market (24 rounds).** At the beginning of P_3, within each matching group, 4 subjects are randomly chosen to act as "principals". Then, in each round t, these 4 principals have to select one contract within the choice set C_t to be offered to the 4 teams of agents in their matching group. We denoted by $C_t^0 \subseteq C_t$ the set of contracts offered by at least one principal (this set may be a singleton, since contracts offered by principals may all coincide, as it often happened in the experiment). Agents have then to choose within this subset C_t^0. Stages 2-4 are then identical to those of P_2. The payoff for the principal -associated with each team which selected his contract- is calculated as the difference between total output, V, and total costs:

$$\pi_0^k(\delta) = P(\delta)(V - b_1^k - b_2^k),$$

with V uniformly distributed within the interval $[100, 150]$. In other words, when a single principal is matched with several teams, the payoff for that principal simply adds the ones he receives by the interaction with each team. Suppose, on the other hand, that the same contract is offered by more than one principal and that contract is accepted by some team (or teams). Then, the principal receives, for each accepting team, the payoff derived from the interaction with that team, divided by the number of principals offering that contract. Notice that this is equivalent to the payoff that would be obtained by randomly and independently matching each accepting team with one of the principals suppliers of that contract.

1.4 Player and Dictator assignment

As we just explained, player position (either 1 or 2) is assigned randomly, for each team and round. This is to fully identify the distributional parameters of our model for each individual subject participating to the experiment -see (3) below. The Dictator position is also assigned randomly, for each team and round. In this respect, we apply the so-called *Random Dictator* protocol (Harrison and McDaniel, 2008). In Section 3.4 we further discuss on the robustness of our results with respect to these two delicate design choices,

presenting evidence from two alternative treatments, in which we vary both players and Dictator assignment protocols.

1.5 Feed-back

At the end of each round, agents were obtaining full information about the other team-mate's behavior. This applies to P_1 (where only one decision is made) as well as P_2 and P_3 (where the other teammate's contract choice is communicated before agents have to make their effort decisions). As for principals in P_3, all payoff relevant information is revealed at the end of each round (how many Dictators and other principals have selected their option, together with the agents' effort profiles).

1.6 Monetary payoffs

All monetary payoffs in the experiment were expressed in Spanish Pesetas (1 euro is approx. 166 ptas.). Subjects received 1.000 ptas. just to show up, to which they summed up all their cumulative earnings throughout the $24 \times 3 = 72$ rounds of the experiment. Average earnings were about 21 euros, for an experimental session lasting for approximately 90 minutes.

1.7 Three testable questions

We are now in the position to specify the main objectives of our experiment.

Q1. *Is it inequality aversion or strategic uncertainty aversion?* Contracts have been calculated using two different mechanism design strategies, with rather different distributional properties. Two kinds of questions arise here.

Q1.1. *Which contract type (sting or wing) is chosen more often by principals and agents?* Evidence for this in **Remark 1**

8

Q1.2. *What is the role of strategic uncertainty?* That is, to which extent the (non) existence of multiple equilibria in *wing* (*sting*) affects agents' behavior in the effort game. Evidence for this in **Remark 2** and **Remark 3**.

Q2. *Do models of social preferences work?* That is, does a model with distributional and reciprocity preferences provide a reliable framework to predict principals and agents' behavior? Evidence for this in **Remark 4** and **Remark 5**.

Q3. *Does separation emerge?* That is, is market able to sort (principals and) agents according to their social preferences? Evidence for this in **Remark 6**.

2 Identifying preferences and beliefs

In what follows, i and j identify our subjects matched in pairs, and we drop the round index, t, whenever this does not create any ambiguity. We assume that our subjects' preferences follow C&R, as we explain in the following

Definition 1 (C&R Preferences)).

(3) $u_i(\delta) = \pi_i(\delta)$

$$-(\alpha_i - \theta_i\phi_j)\max\left\{\pi_j(\delta) - \pi_i(\delta), 0\right\} - (\beta_i + \theta_i\phi_j)\max\left\{\pi_i(\delta) - \pi_j(\delta), 0\right\},$$

where $\phi_j = -1$ if j "has misbehaved", and $\phi_j = 0$ otherwise (we provide an operational definition of misbehavior in equation (4) below). In words, if player j has misbehaved, player i increases her "envy" parameter α_i (or lowers her "guilt" parameter β_i) by an amount equal to θ_i. Thus, θ_i can be interpreted as player i's sensitivity to negative reciprocity. Model (3) has the useful feature that it subsumes parameters which account for subjects' distributional tastes a' la Fehr and Schmidt (1999, F&S hereafter), α_i and β_i, as well as for their tastes for reciprocity, θ_i.

9

In P_1, both agents select their preferred option, after which a random draw determines the identity of the Dictator and both agents earn the monetary payoff associated to the Dictator's choice. Thus, by design, agents in P_1 cannot reciprocate their teammate's decision. In this sense, we can use evidence from P_1 to estimate the distributional preferences parameters of (3), α_i and β_i.

In P_2, after selecting in Stage 1 their favorite contract, k, agents are asked to play the induced effort game, $G(k)$, in which they may condition their effort decision upon the (publicly known) contract choice of their teammate. This, in turn, implies that we can use evidence from P_2 to estimate subjects' reciprocal concerns. To do this, we need first to operationally identify what misbehavior means in the context of our experimental setup. In this respect, we shall use contract choice decision by j in Stage 1, k_j:

$$(4) \qquad \phi_j = \begin{cases} -1 \text{ if } b_j^{k_j} = \max_C b_j^k, \text{ and } b_i^{k_j} \neq \max_C b_i^k, \\ 0 \text{ otherwise.} \end{cases}$$

By (4), j misbehaves by choosing a contract k_j which gives him the highest possible benefit $\left(\max_C b_j^k\right)$, but does not give i her highest possible benefit $\left(\max_C b_i^k\right)$.[9]

We can now look at agents' effort decisions in P_2 as the result of a process of expected utility maximization. Individual i will choose to make effort in Stage 2 $\left(\delta_i^k = 1\right)$ if

$$(5) \qquad E_{\lambda_i^k}\left[u_i^k\left(1, \delta_j^k\right) - u_i^k\left(0, \delta_j^k\right)\right] > 0,$$

where $E_{\lambda_i^k}[\cdot]$ indicates the expected value taken with respect to player i's beliefs on j's effort decision, λ_i^k. We parametrize λ_i^k as a logistic function of the distributional features of contract k, b_j^k and $\left(b_i^k - b_j^k\right)$, and on player i's role ($D_i = 1$ if individual i is the Dictator,

and zero otherwise)

$$(6) \qquad \lambda_i^k = \frac{\exp\left(\psi_1 D_i + \psi_2 b_j^k + +\psi_3(b_i^k - b_j^k)\right)}{1 + \exp\left(\psi_1 D_i + \psi_2 b_j^k + +\psi_3(b_i^k - b_j^k)\right)}.$$

Notice that, in contrast with Bellemare, Kröger and van Soest (2008) or Iriberri and Rey-Biel (2008), we do not directly ask subjects to elicit their beliefs, λ_i^k. Instead, we derive them -indirectly- via their own effort decision. We opted for this design because it allows us to identify cleanly the distributional preferences, separating them from belief identification, without distracting the subjects with new tasks.[10]

Our belief specification (6) allows player i to anticipate that her own Dictator role in Stage 1 may affect j's willingness to make effort. In addition, ψ_2 and ψ_3 proxy the effect associated with absolute and relative payoffs. Our specification for the reciprocity parameter θ_i in (3) allows j's behavior to affect i's effort decision differently, according to $i's$ player position ($L_i = 1$ if individual i is player 2, and zero otherwise) and to the Dictator assignment ($D_i = 1$ if individual i is the Dictator, and zero otherwise):

$$(7) \qquad \theta_i = \theta_0 + \theta_1 D_i + \theta_2 L_i.$$

We exploit the observed contract choice in P_1 to estimate, for each individual, subjects' distributional preference parameters (α_i, β_i) by modelling the probability that individual i chooses option k at round t of P_1 as

$$(8) \qquad \Pr\left(y_{it} = k|\alpha_i, \beta_i, C_t = \left\{b_t^k\right\}, L_{it}\right) = \exp\left(u_{it}^k\right) / \sum_{m=1}^{4} \exp\left(u_{it}^m\right),$$

where u_{it}^m is the all-effort utility of individual i when option m is chosen at round t. Since we posit that distributional preferences estimated in P_1 are constant across phases, effort decisions taken in Stage 2 of P_2 reveal individuals' subjective belief over their teammates' effort decision (i.e. $\boldsymbol{\psi} = (\psi_1, \psi_2, \psi_3)'$) and their own sensitivity to reciprocity (i.e. $\boldsymbol{\theta} =$

$(\theta_0, \theta_1, \theta_2)')$. Consistently, our estimation strategy is a two step procedure: (i) we first get estimates of the distributional parameters, $\widehat{\alpha}_i$ and $\widehat{\beta}_i$, by maximizing the individual log-likelihood $\ln \mathcal{L}_i = \sum_{t,k} 1(y_{it} = k) \ln \Pr(y_{it} = k | \alpha_i, \beta_i, C_t, L_i)$ using data from P_1; (ii) we then model the probability that individual i provides effort in Stage 2 of P_2, given the contract k chosen by the Dictator, as

$$(9) \qquad \Pr\left(\delta_i^k = 1 | \alpha_i, \beta_i, \boldsymbol{\psi}, \boldsymbol{\theta}, (b_1^k, b_2^k), L_i, D_i, \phi_j\right) =$$

$$\frac{\exp\left(E_{\lambda_i^k}\left[u_i^k\left(1, \delta_j^k\right)\right]\right)}{\exp\left(E_{\lambda_i^k}\left[u_i^k\left(1, \delta_j^k\right)\right]\right) + \exp\left(E_{\lambda_i^k}\left[u_i^k\left(0, \delta_j^k\right)\right]\right)}$$

and we estimate - via partial maximum likelihood - the parameters of interest $\boldsymbol{\psi}$ and $\boldsymbol{\theta}$ replacing $\widehat{\alpha}_i$ and $\widehat{\beta}_i$ in (9). Given the two-step nature of the procedure, we use P_1 data to obtain $N = 150$ bootstrap estimates of (α_i, β_i) for each of the 72 subjects, and we use them to obtain a bootstrap distribution of Step 2 estimates.

Put Figure 2 about here

In Figure 2 we plot the estimated α_i and β_i of each member of our subject pool. Figure 2 is composed of two graphs:

1. In Figure 2a) each subject corresponds to a point in the (α_i, β_i) space. Figure 2a) makes clear that our subjects display significant heterogeneity in their distributional preferences. As we discuss in Appendix B1, (some subsets of) the different quadrants in Figure 2 are related to some classes of social preferences which the literature has been interested in. Our estimates can shed light of the prevalence of these different classes among our subjects (see Table B1).

2. Figure 2b) reports, together with each estimated (α_i, β_i) pair (as in Figure 2a), the corresponding 95% confidence intervals associated to each individual estimated parameter. As Figure 2b) shows, we have now many subjects whose estimated distributional pref-

12

erences fall, with nonnegligible probability, in more than one quadrant. Moreover, for some of them (about 20% of our subject pool), we cannot reject (at the 5% confidence level) the null hypothesis of Egoistic Preferences, $\alpha_i = \beta_i = 0$.

Before presenting our estimates of beliefs and reciprocity parameters, Table 1 provides a sketch of average positive effort decisions conditional on misbehavior.

Put Table 1 about here

As Table 1 shows, (i) player 2 provides always less effort than player 1 (0.54 *vs* 0.76); (ii) being or not the Dictator does not affect player 1's effort decision; while (iii) player 2 provides less effort if she is not the Dictator. Moreover, (iv) misbehavior is limited: player 2 misbehaves only 17.6% of the times, player 1 only 13.9% of the cases. As far as reciprocity is concerned, (v) player 1, basically, never reacts to misbehavior: the ratios between conditional effort rates are $0.74/0.76 = 0.97$ for non Dictators and $0.68/0.78 = 0.871\,79$ for Dictators. By (slight) contrast, (vi) player 2 reacts to misbehavior only when she isn't the Dictator ($0.38/0.51 = 0.745\,10$ and $0.57/0.6 = 0.95$, respectively). Points (iv)-(vi) are important to explain why we do not find reciprocity effects in the estimates of Table 2, where we report the estimated coefficients for belief ($\boldsymbol{\psi}$) and reciprocity ($\boldsymbol{\theta}$) parameters.

Put Table 2 about here

As for our belief specification (6), our estimates of Table 2 indicate that player i is expecting more effort the higher j's payoff ($\partial\lambda_i^k/\partial b_j^k = (\psi_2 - \psi_3)\,\lambda_i^k$, and $\widehat{\psi}_2 - \widehat{\psi}_3$ are significantly above zero) and lower effort if her teammates is player 2 ($\widehat{\psi}_3 < 0$ and $b_i^k - b_j^k > 0$). As for our account for Dictator role in i's beliefs, ψ_1, we find a not statistically significant coefficient. Also, our estimates of the three coefficients $\boldsymbol{\theta}$ are not statistically significant, and the same holds for all their relevant linear combinations.

To summarize, our estimations yield statistically insignificant reciprocity effects, and only

(absolute and relative) payoffs seem to have an impact on how subjects form their beliefs.

3   Discussion

We devote this section to provide answers to our conjectural hypotheses and discuss several methodological (as well as empirical) issues raised by our theoretical and experimental setting.

3.1 $Q1$. Is it inequality aversion or strategic uncertainty aversion?

We first analyze subjects' revealed preferences over the type of contract, *wing* or *sting*, to see how subjects resolved the tension between fairness and strategic uncertainty we discussed earlier, and how this depends on their individual social preferences. As explained in Section 1, in 8 out of 24 rounds of each phase, the choice set C_t was composed of 2 *wing* and 2 *sting* contracts, built upon two pairs of distributional preferences. In what follows, we shall refer to the latter as the "mixed rounds". Table 3 reports the relative frequency of subjects' choices of a *sting* contract in the mixed rounds.[11]

Put Table 3 about here

Remark 1 *sting* is the most frequent choice for all players and phases.

As Table 3 shows, in all phases, *sting* is by far the most popular choice, and this is particularly true for player 1 (who, in P_2, goes for *wing* only 7 out of 288 times). Principals also display a marked preference for *sting*, even though choice frequencies are much closer to those of player 2. To assess the extent to which social preferences affect the probability of choosing a *sting* contract, we need to control for the relative inequality associated with the *wing* options, which varies substantially from round to round. In Appendix B2 we run two logit regressions, whose main conclusions are:

14

1. The more "unequal" is the *wing* choice (i.e., the bigger are the payoff differences $b^k_{1t} - b^k_{2t}$ of the 2 *wing* contracts, relative to those of the 2 *sting* contracts in C_t), the more likely is the choice of a *sting* contract. On average, a 1% increase of a "relative inequality index" we build for this purpose yields an increase of the 29% of the probability of choosing *sting* for player 2, and of 14% for the principals in P_3.

2. For principals, distributional parameters are not significant to explain the choice of contract type, while for players 2 in P_2, both α and β are significant, with opposite signs.

We now look at the extent to which contract choices are able to solve the coordination problems agents face in the effort game of Stage 2. Table 4 shows that the relative frequencies of the all-effort efficient equilibrium in *sting* are about twice as large as in *wing* (about 60% vs 30%).

Put Table 4 about here

Remark 2 In *wing*, the inefficient all-no-effort equilibrium pools more than 1/3 of total observations, and it is played more frequently than the efficient all-effort equilibrium.

Notice that about 30% of total observations correspond to a (non-equilibrium) strategy profile in which only one agent makes effort. While this frequency stays basically constant over phases and mechanisms, in *sting* the relative frequency of outcomes in which only player 2 makes effort never exceeds 4% while, in *wing*, this frequency is 3 times as big. If we look at the evolution of outcomes over time, we see that, for both *wing* and *sting,* the relative frequency of efficient equilibria is falling, with this effect is much stronger in *wing*. In addition, when we compare the first and the last 12 rounds of each phase, the frequency of the inefficient no-effort equilibria almost doubles.

We now discuss agents' effort decisions in P_2 and P_3, which can be derived from Table 3. Here we see that i's willingness to make effort is higher when she faces a *sting* contract: if

we focus on P_2 we see that, under a *sting* contract, player 1 makes effort in 92% of the cases, while the same statistic drops to 51% in the *wing* cases. For player 2, the corresponding figures are much lower (62% and 43%, respectively). If we compare effort decisions between P_2 and P_3, we observe an overall reduction in average effort in P_3 for player 1 only in the *wing* case (51% vs 44%).

Remark 3 Effort is much higher in *sting* that in *wing*.

To summarize, if we look at the mechanism design problem from the principal's viewpoint, our evidence yields a clear preference for the "*sting* program". Despite its being more expensive (since the sum of benefits to be distributed is higher), the difference in average team effort is sufficient to compensate the difference in cost. In addition, in the mixed rounds of P_3, principals offering *sting* contracts were chosen with a much higher frequency. This, in turn, implies that average profits for a principal offering a *sting* contract in the mixed rounds were substantially higher, three times as much as when offering a *wing* contract (95.4 ptas. *vs* 30.1).

3.2 $Q2$. Does the social preference model work?

What we learn from the previous section is that distributional preferences play a role to resolve the trade-off implicit in the *wing-sting* choice only for player 2. Matters change in non mixed rounds, when C_t is composed of the same contract type, either *sting* or *wing,* and differences across contracts in C_t are less pronounced. In this case, the *wing-sting* trade-off is not an issue, and principals and agents may fine-tune their contract decisions to their individual distributional tastes. In Appendix B3 we show that, when we focus on relative inequality and relative total cost of chosen contracts by principals and agents (compared with the other available options in C_t), individual social preferences matter. And in the expected direction, as more inequality averse principals and agents choose, on average, contracts in which inequality is reduced. By the same token, more inequality averse principals go for

16

"more expensive" contracts (i.e. contracts in which agents' benefits are higher).

Remark 4 Estimated distributional preferences parameters account well for agents' and principals' observed contract choices in P_2 and P_3.

We also use data from P_3 to check whether our structural model (3-9) is able to predict out-of-sample agents' effort choices in P_3.[12] Once we provide agents with parameters on tastes for distribution, reciprocity and beliefs about their teammate's behavior in Stage 2, we can fully characterize agents' effort decision in P_3 *at the individual level*.

Using the evidence from P_3, each cell of Table 5 reports *a)* relative frequencies of *actual* positive effort decisions, *b)* relative frequencies of *predicted* positive effort decisions and *c)* relative frequencies of instances in which actual and predicted behavior coincide. Predicted behavior is identified by subjects' effort decision which maximizes expected utility (3) in the effort game, subject to their estimated preference parameters $(\alpha_i, \beta_i, \theta_i)$ and their subjective beliefs, λ_i^k.

<div align="center">

Put Table 5 about here

</div>

Overall, the model seems to organize subjects' decisions accurately, which justifies the following

Remark 5 Estimated preferences and beliefs predict about 80% of observed agents' effort decisions.

3.3 $Q3$. Does separation emerge?

Phase 3 evidence can also be used to see how social preferences explain the matching process. To this aim, we estimate the probability that a principal is "chosen" by an agent in each round as a (logit) function of the euclidean distance -in the (α_i, β_i) space- between

agents' and principals estimated distributional preferences:

$$\Pr\left(agent\ i\ chooses\ principal\ j|\left(\alpha_i, \beta_i\right), \left(\alpha_j, \beta_j\right), \boldsymbol{D}_c\right) = \frac{\exp\left(\psi\sigma_{ij} + \boldsymbol{\gamma}'\boldsymbol{D}_c\right)}{1 + \exp\left(\psi\sigma_{ij} + \boldsymbol{\gamma}'\boldsymbol{D}_c\right)},$$

where $\sigma_{ij} = \sqrt{\left(\alpha_i - \alpha_j\right)^2 + \left(\beta_i - \beta_j\right)^2}$ and $\boldsymbol{D}_c$ is a full set of matching group dummies. We estimate the model using only those rounds in which not all the principals offer the same contract. The estimated coefficient ψ is -0.422, (bootstrap and cluster adjusted std. err. 0.099), for a $p-$value of 0.001. This evidence justifies the following

Remark 6 Agents are more likely to choose a contract offered by a principal with more similar distributional preferences to her own.

3.4 Robustness check

We now check the stability of subjects' behavior and structural estimates across alternative design specifications. In this respect, two features of our experimental design looked particularly likely to have affected our inferences from the data.[13]

1. *In our experiment, player position assignment is the outcome of an i.i.d. draw. By recording individual choices in P_1, we are able to identify* both *distributional parameters, α and β, at the individual level. On the other hand, fixing player position across the entire experiment may bias subjects' distributional behavior. For example, inequality might be perceived as less important, since subjects could "smooth" it across rounds. This cannot be done if the player position is fixed throughout.*

2. *By the Random Dictator protocol, both agents choose their favorite contract before knowing the identity of the Dictator. We employed Random Dictator to collect observations on contract decisions for all subjects and rounds (not only for Dictators). However, in this case, fairness can be achieved in two ways: either by choosing the*

"fair" contract in each round, or by choosing the "unfair" contract in each round, letting the Random Dictator assignment provide overall fairness.

To investigate on these issues, we run two extra treatments (3 sessions each, like in the "original" treatment, TR_1 hereafter):

(i) TR_2 : where we apply the Random Dictator protocol, but we fix player position throughout the session;

(ii) TR_3: where we apply the Reversal Dictator protocol (i.i.d. Dictator assignment is revealed before option choice, with only the Dictator deciding over her favorite option); and -just like in TR_2- player position is fixed throughout.

Unlike in TR_1, for our alternative treatment conditions TR_2 and TR_3 data from P_1 do not allow to identify both distributional parameters, α and β, at the individual level, but only one, or the other, depending on player position. In addition, in TR_3 we collect exactly half of the observations over the contract decisions across all phases (since non Dictators do nothing). This, in turn, implies that we can only measure reciprocity on behalf of non Dictators.

In Section 3.4.1 we first document the (small and nonsignificant) differences in distributional choices across treatments. Then, in Section 3.4.2 we examine the observed differences in average effort rates in P_2 and P_3 (compared with those of our baseline treatment, TR_1), and evaluate potential explanations for such differences.

3.4.1 Robustness check 1: contract choices

There is very little difference across treatments in the purely distributional choices of P_1. Since, in any given round t, the available choice set C_t is constant across treatments, we can study whether P_1 choices are different across treatments using a Pearson test for the null hypothesis that, for a given choice set C_t, the distribution of the chosen contracts is

independent of the treatment considered. As Table B6 (in Appendix B) shows, there are very few rounds (3 in the comparison between TR_1 and TR_2, 2 in the comparisons between TR_1 and TR_3 and one between TR_2 and TR_3) where choices within rounds are significantly different. By the same token, contract choices are essentially constant across treatments also in P_2 and P_3 (details in Appendix B4).

3.4.2 Robustness check 2: effort decisions and parameter estimations

Contrary to contract choices, effort levels differ significantly across treatments. Table B7 (in Appendix B) shows that effort levels are rather smaller in TR_3, compared to those in TR_1 or TR_2, themselves actually similar to one another (in Appendix B6 we show how these similarities and differences translate into statistical significance).

These differences require some explanation. One could reasonably conjecture that the difference between TR_1 and TR_3 arises because Player 2, who is continuously getting the worse outcome in TR_3, tries to balance this by "shirking", whereas the switching roles (and the large number of rounds) do the balancing in TR_1. The strong similarity of effort rates between TR_1 and TR_2 disposes of this conjecture since, in TR_2, player positions are also fixed, and the differences in effort rates between TR_1 and TR_2 are much lower than those between TR_1 and TR_3 (see Table B7, Appendix B).

Another potential explanation could be that fixing player position *and* revealing Dictator role before option choice makes reciprocity more salient. Hence the "punishment" for misbehavior (i.e. lower effort) would be more frequent. The evidence from Table B8 (in Appendix B) disposes of this conjecture, too, showing that there is no more punishment to misbehavior in TR_3 than in the other treatments, *there is simply less effort.* As further confirmation of this, Table B9 compares the estimates of (a suitably modified version of) our structural model of Section 2 across treatments. As Table B9 shows, the estimates of our reciprocity parameters show little differences across treatments.

According to our estimates, the most likelihood explanation for the difference in behav-

ior comes from *beliefs*. Table B10 (in Appendix B) shows that players believe (and their experiences will confirm) that others are going to make less effort in TR_3.

4 Conclusion

Our experimental results show that strategic uncertainty should be an important concern for those in charge of designing organizational incentives. In our context, where strategic uncertainty conflicts with social preferences in terms of their respective recommendations on contract design, the former seems to be subjects' primary concern. Nevertheless, we also provide evidence showing that distributional preferences are a key determinant of contracts offered and accepted, on effort levels, as well as on how markets sorts different distributional attitudes towards different organizations.

Our experimental environment is certainly *ad-hoc* in some respects (take, for example, our decision to give to only one agent the monopolistic power to decide the ruling contract for the entire team), although our findings seem fairly stable across some important design modifications, such as those concerning player position or Dictator assignment. In this respect, our results are encouraging, because a parsimonious model of individual decision making seems capable of organizing consistently the evidence from a complex experiment, across various treatment conditions.

As for avenues for future research, it would be interesting to solve theoretically the mechanism design problem under incomplete information about agents' social preferences. From an empirical standpoint, it would also be interesting to observe the effect of having agents with different productivities, which are also private information. In this way we could see how finely and in which ways "corporate culture" partitions the agents. Also, notice that, in our setup, the numbers of principals and agents exactly balance one another. On the other hand, the effect of more intense competition on either side is another empirically interesting extension.

We conclude by mentioning a development of our setup that has already been explored. Frignani and Ponti (2008) use our experimental design (and the evidence of TR_1, as a control) to collect experimental evidence of the Random Dictator setup (P_1) in which *i)* subjects choose under the "Veil Of Ignorance" (VOI), that is, choose their favorite option knowing that they will eventually be assigned to either player position with equal probability or *ii)* our choice sets C_t correspond to binary lotteries (LOTT), in which subjects may win the high or the low prize with equal probability, but their decisions do not affect other participants. Subjects' decisions are framed within the realm of a simple mean-variance utility maximization problem, where the parameter associated to the variance is interpreted, depending on the treatment, as a measure of pure risk aversion (LOTT), pure inequality aversion (TR_1), or some combination of the two (VOI). In this respect, they find that both pooled and individual estimated parameter distributions in the VOI and LOTT treatments are remarkably similar. In stark contrast, the estimated inequality aversion in TR_1 is significantly lower. In other words, we could simply rely on risk aversion to explain subjects' behavior under the VOI, suggesting that, in the presence of both risk and distributional issues, the former seem to have a primary consideration. This result nicely complements our findings on the predominance of strategic uncertainty considerations in the *wing-sting* decision reported in Section 3.1.

REFERENCES

[1] **Bellemare, C., S. Kröger, and A. van Soest.** 2008. "Actions and Beliefs: Estimating Distribution-Based Preferences Using a Large Scale Experiment with Probability Questions on Expectations." *Econometrica* 76(4): 815-839.

[2] **Bellemare C., and B. S. Shearer.** 2006. "Sorting, Incentives and Risk Preferences: Evidence from a Field Experiment." IZA Discussion Paper No. 2227.

[3] **Bewley, T.F.** 1999. *Why Wages Don't Fall during a Recession.* Cambridge, MA: Harvard University Press,

[4] **Cabrales, A., and A. Calvó-Armengol**. 2008. "Interdependent Preferences and Segregating Equilibria." *Journal of Economic Theory*, 139(1): 99-113.

[5] **Cabrales, A., A. Calvó-Armengol and N. Pavoni**. 2008. "Social Preferences, Skill Segregation and Wage Dynamics." *Review of Economic Studies*, 75(1):65–98.

[6] **Charness, G., and M. Rabin.** 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics*, 117(3): 817-869.

[7] **Crawford, V.P.** 1995. "Adaptive Dynamics in Coordination Games." *Econometrica,* 63(1): 103-143.

[8] **Crawford, V.P., and H. Haller.** 1990. "Learning How to Cooperate: Optimal Play in Repeated Coordination Games." *Econometrica,* 58(3): 571-595.

[9] **Dohmen T., and A. Falk.** 2006. "Performance Pay and Multi-dimensional Sorting:Productivity, Preferences and Gender." IZA Discussion Paper No. 2001.

[10] **Fehr E., A. Klein, and K. Schmidt.** 2007, "Fairness and contract design." *Econometrica,* 75(1): 121-154.

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

[12] **Fershtman C., H.K. Hvide, and Y. Weiss.** 2006. "Cultural diversity, status concerns and the organization of work." *Research in Labor Economics,* 24: 361-396.

[13] **Fischbacher, U.** 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics,* 10(2): 171-178.

[14] **Frignani, N. and Ponti, G.** 2008. "Social *vs* Risk Preferences under the Veil of Ignorance", *Quaderno deit*, Dipartimento Economia Istituzioni Territorio, Università di Ferrara, n. 8/2008.

[15] **Harrison G.W. and McDaniel T.** 2008. "Voting Games and Computational Complexity", *Oxford Economic Papers*, 60(3): 546-565.

[16] **Heinemann, F., R. Nagel, and P. Ockenfels.** 2009. "Measuring Strategic Uncertainty in Coordination Games.", *Review of Economic Studies*, 76(1): 181-221.

[17] **Iriberri, N. and Rey-Biel, P.** (2008). "Elicited Beliefs and Social Information in Modified Dictator Games: What Do Dictators Believe Other Dictators Do?", Working paper, Department of Economics and Business, Universitat Pompeu Fabra, 1137.

[18] **Kosfeld, M., and F. von Siemens.** 2009. "Worker Self-Selection and the Profits from Cooperation", *Journal of the European Economic Association*, 7(2-3): 573-582.

[19] **Kremer, M.** 1993. "The O-Ring Theory of Economic Development." *Quarterly Journal of Economics*, 108(3): 551-575.

[20] **Krueger, A.B. and D. Schkade.** 2007. "Sorting in the Labor Market: Do Gregarious Workers Flock to Interactive Jobs?", *Journal of Human Resources*, 43(4): 861-885.

[21] **López-Pintado, D., G. Ponti, and E. Winter.** 2008. "Inequality or Strategic Uncertainty? An Experimental Study on Incentives and Hierarchy." In *Games Rationality and Behaviour,* ed. Alessandro Innocenti and Patrizia Sbriglia, 235-255. London: Palgrave Macmillan,

[22] **Manski, C.** 2002. "Identification of Decision Rules in Experiments on Simple Games of Proposal and Response.", *European Economic Review*, 46(4-5): 880-891.

[23] **Nyarko, Y., and A. Schotter.** 2002. "An Experimental Study of Belief Learning Using Elicited Beliefs," *Econometrica*, 70(3): 971–1005.

[24] **Rey-Biel, P.** 2008. "Inequity Version and Team Incentives.", *Scandinavian Journal of Economics*, 108 (2): 297-320.

[25] **Teyssier, S.** 2008. "Optimal Group Incentives with Social Preferences and Self-Selection," Working Paper du GATE 2007 Vol. 10.

[26] **Van Huyck, J.B., R.C. Battalio and, R.O. Beil.** 1990. "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure." *American Economic Review,* 80(1): 234-248.

[27] **Van Huyck, J.B., R.C. Battalio and, R.O. Beil**. 1991. "Strategic Uncertainty, Equilibrium Selection, and Coordination Failure in Average Opinion Games," *Quarterly Journal of Economics* 106(3): 885-910.

[28] **Winter, E.** 2004. "Incentives and Discrimination." *American Economic Review,* 94(3): 764-773.

Notes

†Corresponding author. Departamento de Fundamentos del Análisis Económico -Universidad de Alicante - 03071 Alicante - SPAIN - Voice:++34965903619 - Fax: ++ 34965903898 - e-mail:giuba@merlin.fae.ua.es.

[1] Van Huyck, Battalio and Beil (1990, 1991) and Heinemann, Nagel and Ockenfels (2009) experiments measure the extent and importance of strategic uncertainty in coordination games. Crawford (1995) and Crawford and Haller (1990) are theoretical papers partly inspired by these experimental results. López-Pintado, Ponti and Winter (2008) test directly Winter's (2004) model in the lab.

[2] See Fershtman, Hvide and Weiss (2005), Rey-Biel (2008) and Kosfeld and von Siemens (2009) for theoretical results and Fehr, Klein and Schmidt (2007) for experimental evidence.

[3] Cabrales, Calvó-Armengol and Pavoni (2008), Cabrales and Calvó-Armengol (2008) and Teyssier (2008) show that social preferences lead to more productive workers sorting themselves into different firms than the remaining workers. On the empirical literature on (social)

26

preferences and sorting, see also Dohmen and Falk (2006), Krueger and Schkade (2008) and Bellemare and Shearer (2006).

[4]The experiment was programmed and conducted with the software *z-Tree* (Fischbacher, 2007).

[5]See Winter (2004) for details.

[6]The fact that monetary payoffs are derived from a specific theoretical exercise -instead of being simply randomly generated- has no further impact on our experimental design. Subjects were not told at any time where those numbers came from. They simply had to choose, at each round, one out of four different options, with no further explanation.

[7]A new set of instructions was distributed at the beginning of each phase. In this sense, subjects were not aware at all times about the rules of the phases to follow.

[8]In P_1 agents only see the four payoff pairs $\left(\pi_{1t}^k\left(1,1\right),\pi_{2t}^k\left(1,1\right)\right),$ without any reference to the game-form that generated them, $G(k)$. By contrast, in P_2 and P_3 agents see the full payoff matrices, both when they have to choose their preferred contract, and when they have to play it. Snapshots of the user interfaces for all phases are provided in Appendix C, together with the experimental instructions.

[9]This definition has the advantage of being well-defined for the three experimental treatments we have run. Results are qualitatively similar for many other definitions we have tried.

[10]As Nyarko and Schotter (2002) acknowledge, belief elicitation has its own problems, since "*As is true of all scoring functions, while payoffs are maximized by truthful revelation of beliefs, there are other beliefs that could be stated that are more secure* [...] *If subjects were risk averse, such an action might be desirable.*"

[11]The statistics of Table 2 -as far as agents in P_3 are concerned- do not include observations from mixed rounds in which all principals were offering either a *wing* or a *sting* contract.

[12]Our behavioral model (3) provides a suitable framework to predict agents' effort decisions. To also predict contract choices, we should *i)* model agents' beliefs on the probability

of teammates' misbehavior in the contract decision (and, in consequence, principals' beliefs over those beliefs); *ii)* provide a robust model of competition among principals; and *iii)* deal with the incomplete information about agents' (and other competing principals') preferences.

[13]We thank two anonymous referees for pointing out these critical aspects of our original design.

		i is Player 1		
	% $\phi_j = -1$	$\phi_j = -1$	$\phi_j = 0$	Total
No Dictator (443)	19.6	0.74	0.76	0.75
Dictator (421)	15.4	0.68	0.78	0.76
Total (854)	17.6	0.71	0.77	0.76
		i is Player 2		
	% $\phi_j = -1$	$\phi_j = -1$	$\phi_j = 0$	Total
No Dictator (421)	11.9	0.38	0.51	0.49
Dictator (443)	15.8	0.57	0.60	0.59
Total (854)	13.9	0.49	0.55	0.54

Table 1: Incidence of misbehavior (first column) and average effort rates P_2 conditional on Dictator and player position. Number of cases in parenthesis.

Beliefs $\left(\lambda_i^k\right)$		$Coeff.$	$Std.err.$	$p-value$
D_i	ψ_1	0.3135	0.2689	0.24
b_j^k	ψ_2	0.0110	0.0093	0.237
$b_i^k - b_j^k$	ψ_3	-0.1112	0.0339	0.00
Reciprocity (θ_i)		$Coeff.$	$Std.err.$	$p-value$
$Constant$	θ_0	-0.1464	0.1114	0.19
D_i	θ_1	0.0895	0.1460	0.54
L_i	θ_2	0.1667	0.1388	0.23

Table 2: Estimated parameters of belief function and reciprocity. Bootstrap and matching group adjusted standard errors.

	Phase 2	Phase 3
Player 1	0.98	0.89
Player 2	0.68	0.76
Principals		0.75

Table 3: Relative frequencies of the *sting* choice in the mixed rounds

	$P_2, wing$				$P_2, sting$			
	None	Pl. 1	Pl. 2	Both	None	Pl. 1	Pl. 2	Both
Rounds 1-12	44 26.2%	37 22%	24 14.3%	63 37.5%	10 3.8%	83 31.4%	8 3.0%	163 61.7%
Rounds 13-24	80 46.8%	31 18.1%	19 11.1%	41 24%	19 7.3%	90 34.5%	5 1.9%	147 56.3%
Total	124 36.6%	68 20.1%	43 12.7%	104 30.7%	29 5.5%	173 33%	13 2.5%	310 59.1%
	$P_3, wing$				$P_3, sting$			
	None	Pl. 1	Pl. 2	Both	None	Pl. 1	Pl. 2	Both
Rounds 1-12	35 30.7%	22 19.3%	15 13.2%	42 36.8%	6 3.5%	46 26.4%	7 4.0%	115 66.1%
Rounds 13-24	59 54.6%	10 9.3%	15 13.9%	24 22.2%	17 9.4%	60 33.3%	2 1.1%	101 56.1%
Total	94 42.3%	32 14.4%	30 13.5%	66 29.7%	23 6.5%	106 29.9%	9 2.5%	216 61.0%

Table 4: Outcome dynamics in the effort game. Absolute values and row percentages
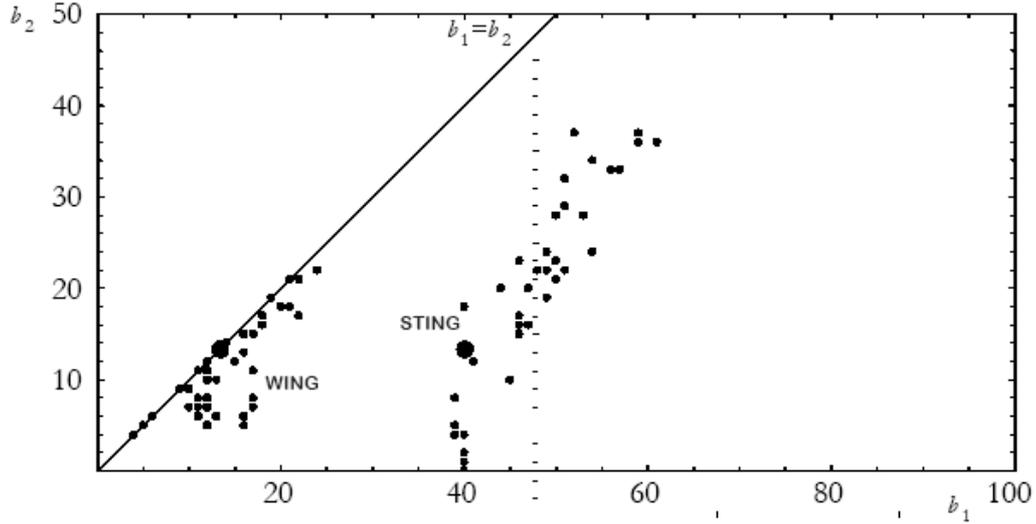
Figure 1: The experimental choice set.

	wing contracts			sting contracts		
	i is player 1					
	$\phi_j=-1$ (31)	$\phi_j=0$ (191)	Total (222)	$\phi_j=-1$ (44)	$\phi_j=0$ (310)	Total (354)
No Dictator.	.27	.49	.47	.95	.89	.90
	.24	.39	.37	.89	.92	.92
	.65	.83	.8	.67	.86	.83
Dictator	.25	.46	.42	.88	.93	.92
	.1	.38	.36	.81	.90	.88
	.8	.66	.68	.77	.83	.82
Total	.26	.47	.44	.91	.91	.91
	.19	.39	.36	.84	.92	.90
	.7	.74	.73	.73	.85	.83
	i is Player 2					
	$\phi_j=-1$ (26)	$\phi_j=0$ (196)	Total (222)	$\phi_j=-1$ (43)	$\phi_j=0$ (311)	Total (354)
No Dictator.	.1	.48	.42	.48	.65	.63
	.1	.33	.31	.54	.68	.66
	1	.68	.71	.74	.77	.77
Dictator	.18	.48	.45	.35	.68	.64
	0	.49	.44	.41	.59	.56
	.81	.74	.75	.81	.75	.77
Total	.14	.48	.43	.42	.67	.64
	.05	.40	.36	.49	.64	.61
	.90	.71	.72	.77	.77	.77

Table 5: Actual and predicted behavior in Stage 2 of P_3. For each case we report relative frequencies of actual positive effort decisions, relative frequencies of predicted positive effort decisions, and the fraction of cases for which actual and predicted effort behavior coincides. Number of cases in parenthesis.
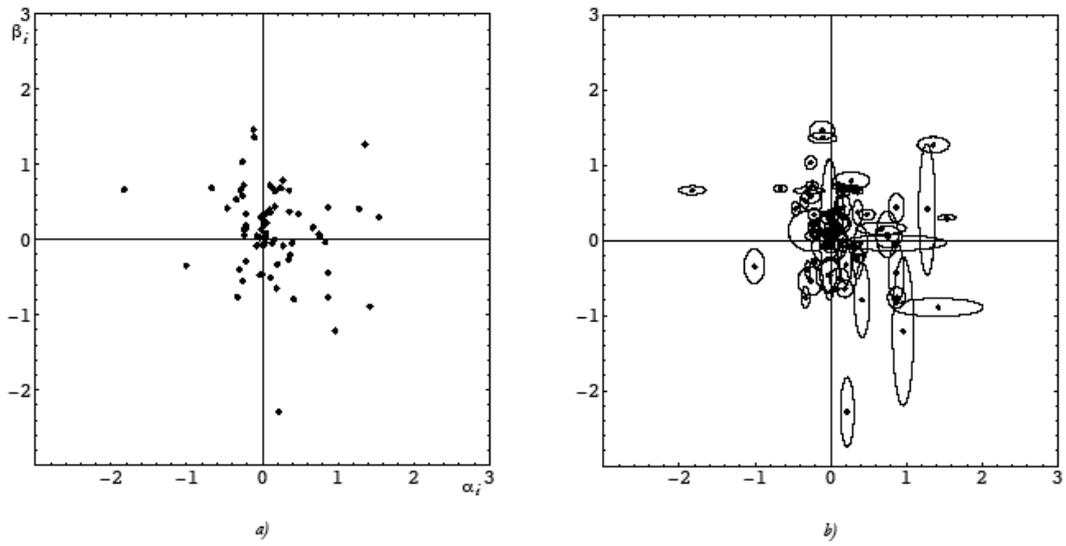
Figure 2: Estimating distributional preferences
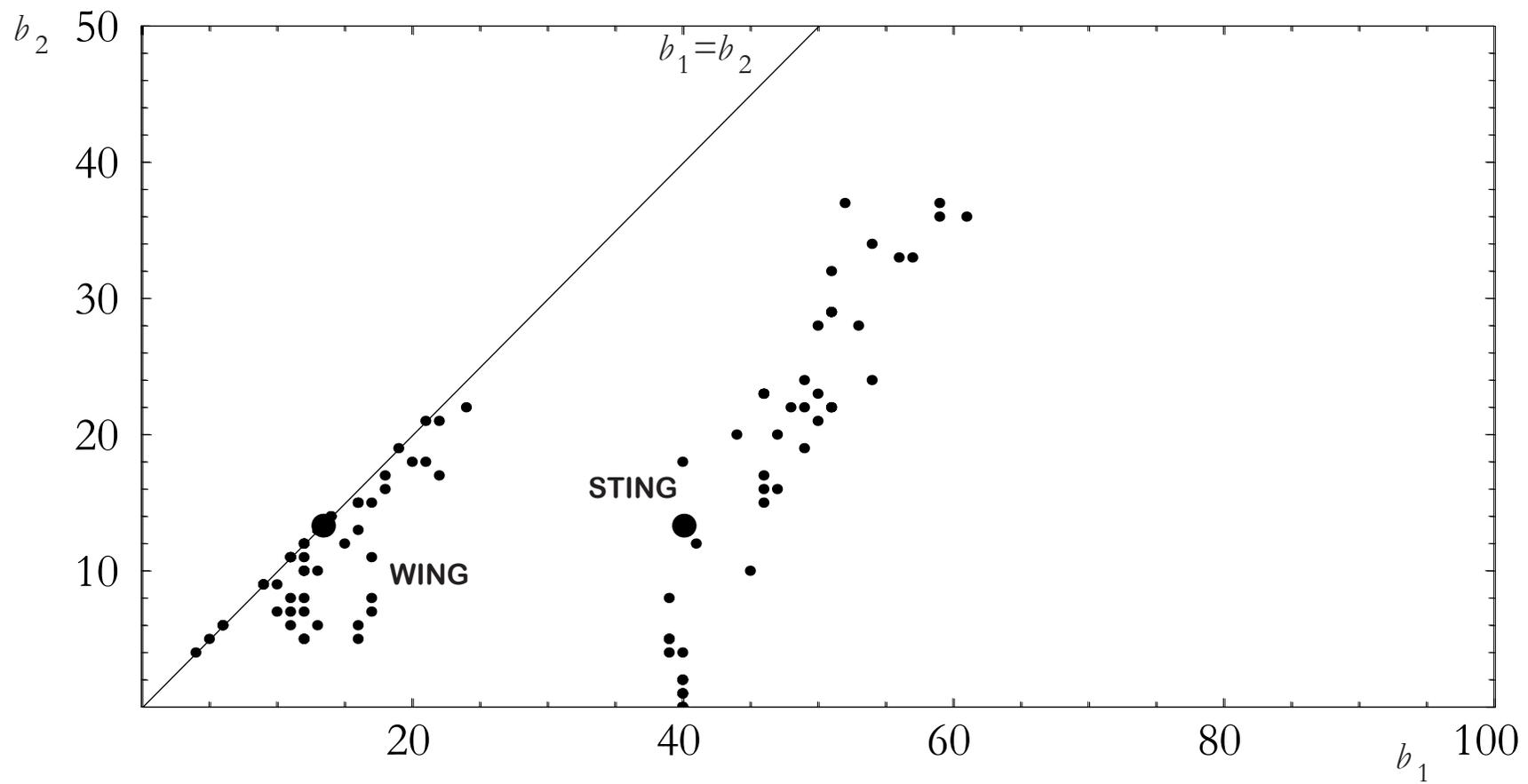
Fig. 1
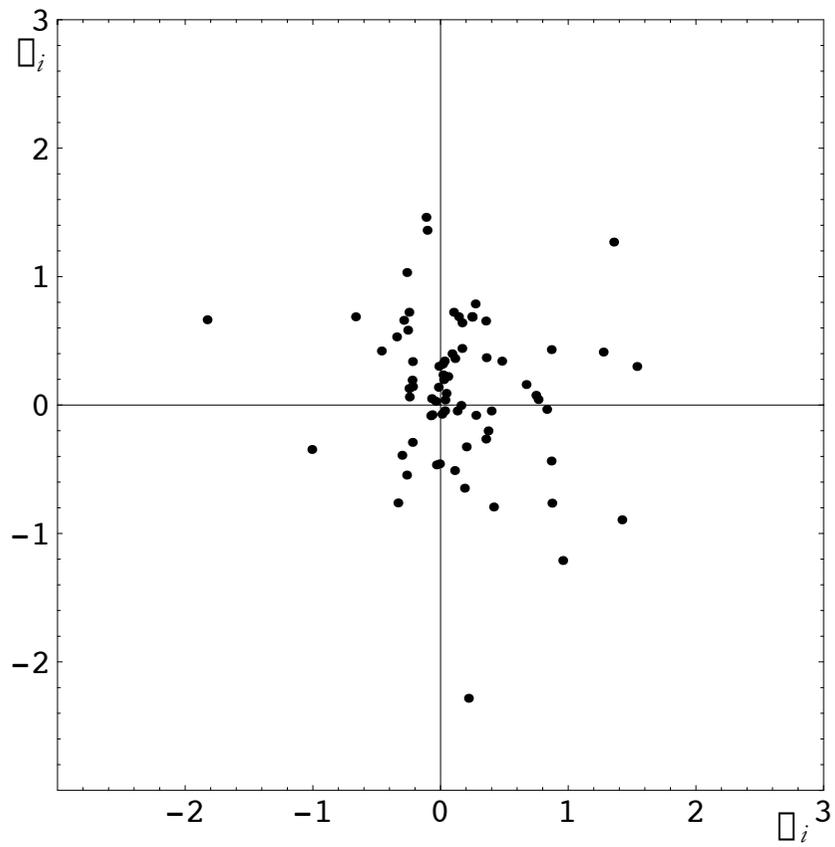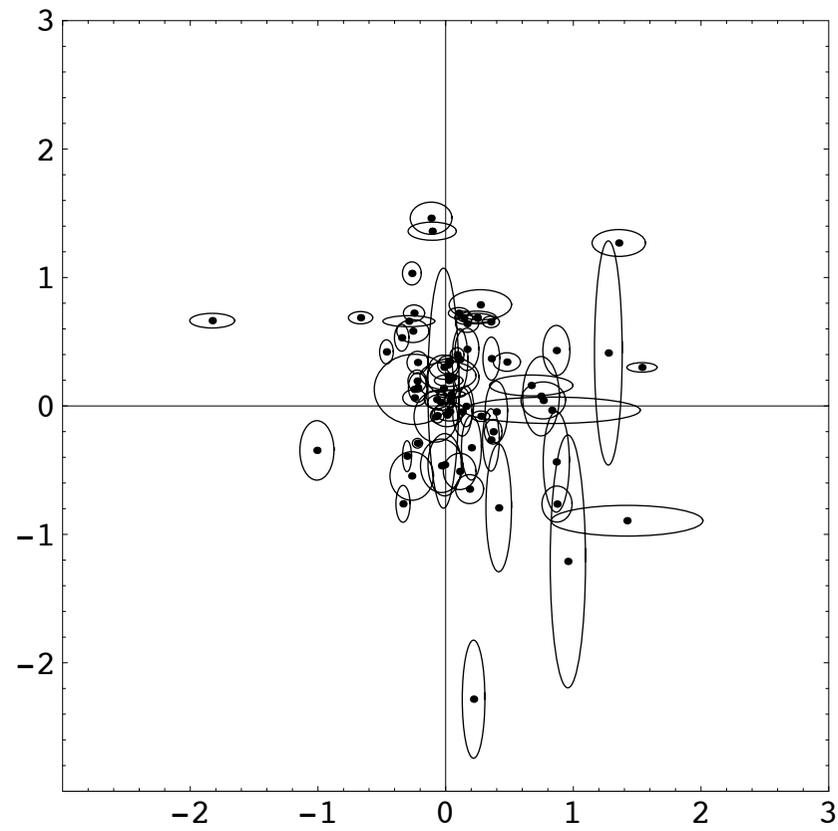
Fig. 2

Social Preferences and Strategic Uncertainty: an Experiment on Markets and Contracts
Appendix A: The Mechanism Design Problem

NOT FOR PUBLICATION

1   Two mechanism design problems

1.1   Production technology

Technology closely follows Winter's (2004) model of moral hazard in teams. Let $G(b)$ define the game-form associated with a given benefit profile, $b = (b_1, b_2)$. The rules rules of the game-form are the following. Each agent $i = 1, 2$, has to decide, simultaneously and independently, whether to make a costly effort. We denote by $\delta_i \in \{0, 1\}$ agent i 's effort decision, where $\delta_i = 1(0)$ if agent i does (does not) make effort. Let also $\delta = (\delta_1, \delta_2) \in \{0, 1\}^2$ denote the agents' action profile. The cost of effort c is assumed to be constant across agents. Team activity results in either success or failure. Let $P(\delta)$ define production as the probability of success as a function of the number of agents in the team who have put effort:

$$P(\delta) = \begin{cases} 0 \text{ if } \delta_1 + \delta_2 = 0, \\ \gamma \text{ if } \delta_1 + \delta_2 = 1, \\ 1 \text{ if } \delta_1 + \delta_2 = 2, \end{cases} \tag{1}$$

with $\gamma \in (0, \frac{1}{2})$.[1]

If the project fails, then all (principal and agents) receive a payoff of zero. If the project succeeds, then agent i receives a *benefit*, $b_i^k > 0$. Agent i's expected monetary profit associated to contract k is given by

$$\pi_i^k(\delta) = P(\delta)b_i^k - \delta_i c. \tag{2}$$

The expected monetary payoff for the principal is the difference between expected revenues, for a given (randomly generated) value for the project $V \sim U[A, B]$, and expected costs:

$$\pi_0^k(\delta) = P(\delta)(V - b_1^k - b_2^k).$$

---

[1] This is how Winter (2004) models moral hazard: agents' effort affect the overall probability of success of the project. However, since risk neutrality is assumed on agents' behalf, the fact that technology follows a random -as opposed to deterministic, as in our design- process has no impact in the solution of the mechanism design problem.

Assume a principal who wishes to design a mechanism that induces all agents to exert effort in (some) equilibrium of the game induced by $G(b)$, which we denote by $\Gamma(b)$. A mechanism is an allocation of benefits in case of success, i.e., a vector b that satisfies this property at the minimal cost for the principal. Following Winter (2004), the principal may consider mechanisms that *strongly* or *weakly* implement the desired solution, depending of how concerned he is about equilibrium multiplicity. More precisely:

Definition 1 (*sting* contracts) *The contract b is strongly effort-inducing (sting) if all Nash Equilibria (NE) of $\Gamma(b)$ entail effort by all agents with minimal benefit distribution, $b_1 + b_2$.*

Definition 2 (wing contracts) *The contract b is weakly effort-inducing (wing) if there exists at least one NE of $\Gamma(b)$ such that $\delta = (1,1)$, with minimal benefit distribution.*

2 The solutions

By analogy with our experimental conditions (and without loss of generality), we assume $b_1 \geq b_2$.

There are four relevant subsets of parameters for the utility function, which we now describe. All these specifications do not consider reciprocal motives (i. e., it is always assumed $\theta_i = 0$), and, in this sense, define purely "distributional" preferences.

$$\text{Egoistic Preferences (EP): } \alpha_i = \beta_i = 0. \tag{3}$$

$$\text{Inequality Averse Preferences (IAP): } 0 \leq \beta_i < 1, \alpha_i \geq \beta_i. \tag{4}$$

$$\text{Status Seeking Preferences (SSP): } \alpha_i \in [0,1), \beta_i \in (-1,0], |\alpha_i| \geq |\beta_i| \tag{5}$$

$$\text{Efficiency Seeking Preferences (ESP): } \alpha_i \in (-\frac{1}{2}, 0], \beta_i \in [0, \frac{1}{2}), |\beta_i| \geq |\alpha_i| \tag{6}$$

In what follows, we shall assume that both agents hold either EP (as in Winter, 2004), or IAP, SSP and ESP, respectively. We allow for heterogeneous preferences, provided they belong to the same preference class.

2.1 Solution of the mechanism design problem under the *wing* program

In the case of *wing*, the search of the optimal mechanism corresponds to the following linear program:

$$b^* \equiv (b_1^*, b_2^*) \in \arg \min_{\{b_1, b_2\}} [b_1 + b_2] \text{ sub} \tag{7}$$

$$u_1(1,1) \geq u_1(0,1) \tag{8}$$

$$u_2(1,1) \geq u_2(1,0) \tag{9}$$

$$b_1 \geq b_2 \geq 0 \tag{10}$$

2

Assumption (10) is wlog. To solve the problem (7)-(10), we begin by partitioning the benefit space $B = \left\{ (b_1, b_2) \in \Re_+^2, b_1 \geq b_2 \right\}$ in two regions, which specify the payoff ranking of each strategy profiles in $G(b)$. This partition is relevant for our problem, since it determines whether in (1,0) - player 1 exerts effort and player 2 does not - whether it is player 1 or 2 the one who experiences envy (guilt):

$$
\begin{aligned}
R_1 &= \left\{ b \in B : b_2 \leq b_1 - \frac{c}{\gamma} \right\}; \\
R_2 &= \left\{ b \in B : b_1 - \frac{c}{\gamma} \leq b_2 \leq b_1 \right\}.
\end{aligned}
$$

Let $g^1(b_1) = b_1 \left(g^2(b_1) = b_1 - \frac{c}{\gamma} \right)$ define the two linear constraints upon which our partition is built. The strategy proof is as follows. We shall solve the linear program (7)-(10) in the two regions independently (since, within each region, social utility parameters are constant for each agent and strategy profile), checking which of the two solutions minimizes the overall benefit sum $b_1 + b_2$, and determining the constraints on preferences which determine the identity of the best-paid player 1.

2.1.1 *Wing* under EP

As for the solution of *wing* under EP (i.e. with $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0$), the linear program (7)-(10) simplifies to the following:

$$\min b_1 + b_2$$

subject to:

$$
\begin{aligned}
b_1 - c &\geq \gamma b_1 \\
b_2 - c &\geq \gamma b_2 \\
b_i &\geq 0; \quad \text{with } i = 1, 2
\end{aligned}
$$

In this case, the solution of the problem is problem is trivial:

$$b_1^* = b_2^* = \tfrac{c}{1-\gamma}.$$

2.1.2 *Wing* under IAP

As for the solution of *wing* under IAP, we need to add to the basic linear program (7)-(10) the IAP constraint.

Proposition 3 (winiIAP) *The optimal wing mechanism under IAP is as follows:*

$$b_1^* = \left(\begin{array}{c} \frac{c(-1+\alpha_2(-1+\beta_1)+2\beta_1+\gamma(-1+2\beta_1)(-1+\beta_2)-\beta_1\beta_2)}{(-1+\gamma)(1+\alpha_2-\beta_1+\gamma(-1+\beta_1+\beta_2))}, \\ \frac{c(-1+\beta_1)(-1+\alpha_2-\beta_2+\gamma(-1+2\beta_2))}{(-1+\gamma)(1+\alpha_2-\beta_1+\gamma(-1+\beta_1+\beta_2))} \end{array} \right) \quad \text{if } \beta_1 < \frac{1}{2}; \tag{11}$$

$$b_2^* = \left(\frac{c(1-\beta_1)}{1-\gamma}, \frac{c(1-\beta_1)}{1-\gamma} \right) \quad \text{if } \beta_1 \geq \frac{1}{2}, \tag{12}$$

with $\beta_1 \leq \beta_2$.

To prove Proposition 3, some preliminary lemmas are required. Let $\hat{b}^k \equiv (\hat{b}_1^k, \hat{b}_2^k)$ define the solution of the linear program (7-10) in R_k.

Lemma 4

$$\hat{b}^1 = \left(\frac{c(1+\alpha_2)}{(1-\gamma)\gamma}, \frac{c(\gamma+\alpha_2)}{(1-\gamma)\gamma} \right) \tag{13}$$

.

Proof. In R_1, agent 1's monetary payoff, as determined by $G(b)$, is always higher (i.e. $\pi_1(\delta) \geq \pi_2(\delta), \forall \delta$). This, in turn, implies that constraints (8)-(9) correspond to

$$b_1 \geq f_1^1(b_1) \equiv \frac{c(1-\beta_1)}{(1-\gamma)\beta_1} - \frac{1-\beta_1}{\beta_1}b_1; \tag{14}$$

$$b_2 \geq f_2^1(b_1) \equiv \frac{c}{1-\gamma} + \frac{\alpha_2}{1+\alpha_2}b_1. \tag{15}$$

Let x_i^k define the value of b_1 such that $f_i^k(b_1) = 0$. By the same token, let y_i^k denote the intercept of $f_i^k(b_1)$, i.e. $f_i^k(0)$. Finally, let τ_i^k denote the slope of $f_i^k(b_1)$. We then have $x_1^1 = \frac{c}{1-\gamma}$ and $x_2^1 = -\frac{c(1+\alpha_2)}{(1-\gamma)\alpha_2}$. Also notice that $0 \leq \tau_2^1 = \frac{\alpha_2}{1+\alpha_2} < 1$ and $y_2^1 = \frac{c}{1-\gamma} > 0$. This implies that $f_2^1(b_1)$ and $g^2(b_1)$ intersect in the first quadrant of the $b_1 \times b_2$ space. On the other hand, $f_1^1(b_1)$ is never binding in this case, since $\tau_1^1 = -\frac{1-\beta_1}{\beta_1} < 0$ and $x_1^1 = \frac{c}{1-\gamma} < \frac{c}{\gamma}$ since $\gamma < \frac{1}{2}$. This implies that $b_1 + b_2$ is minimized where $f_2^1(b_1)$ and $g^2(b_1)$ intersect, i.e. when $\hat{b}_1^1 = \frac{c(1+\alpha_2)}{(1-\gamma)\gamma}$ and $\hat{b}_2^1 = \frac{c(\gamma+\alpha_2)}{(1-\gamma)\gamma}$. $\blacksquare$

Lemma 5 *In R_2, the optimal wing contract under IAP is (11) when $\beta_1 < \frac{1}{2}$, and (12) when $\beta_1 \geq \frac{1}{2}$, with $\beta_1 < \beta_2$.*

Proof. In the case of R_2, constraints (8)-(9) correspond to

$$b_1 \geq f_1^2(b_1) \equiv \frac{c(1-\beta_1)}{(1-\gamma)\beta_1} - \frac{1-\beta_1}{\beta_1}b_1; \tag{16}$$

$$b_2 \geq f_2^2(b_1) \equiv \frac{c(1-\beta_2)}{1+\alpha_2-\gamma(1-\beta_2)} + \frac{\alpha_2+\gamma\beta_2}{1+\alpha_2-\gamma(1-\beta_2)}b_1. \tag{17}$$

This implies that $f_1^1(b_1) = f_1^2(b_1)$ (i.e. the Nash equilibrium condition for player 1 remains unchanged in both R_1 and R_2), $\tau_1^2 = -\frac{1-\beta_1}{\beta_1} < 0$ (i.e. $\left|\tau_1^2\right| > 1$ if $\beta_1 < \frac{1}{2}$), and $0 \leq \tau_2^2 = \frac{\alpha_2+\gamma\beta_2}{1+\alpha_2-\gamma(1-\beta_2)} < 1$.

4

We first show that $\beta_1 \leq \beta_2$. Let $\check{\beta} = \min\{\beta_1, \beta_2\}$. If $\beta_1 > \beta_2$, then the optimal solution in R_2 would be $\hat{b}_i^1 = \hat{b}_i^2 = \frac{c(1-\check{\beta})}{1-\gamma}$ (i.e. $\hat{b}_i^1 + \hat{b}_i^2 = 2\frac{c(1-\check{\beta})}{1-\gamma}$). On the other hand, if $\beta_1 \leq \beta_2$, then $\hat{b}_i^1 + \hat{b}_i^2 \leq 2\frac{c(1-\check{\beta})}{1-\gamma}$. More precisely, if $\beta_1 < \frac{1}{2}$, the optimal solution is (11), that is, the intersection between $f_1^2(b_1)$ and $f_2^2(b_1)$; if $\beta_1 \geq \frac{1}{2}$, the solution is (12), that is, the intersection between $f_1^2(b_1)$ and $g^1(b_1)$. $\blacksquare$

We are in the position to prove Proposition 3.

Proof. [Proof of Proposition 1]. To prove the proposition, it is sufficient to show that $\hat{b}_i^1 > \hat{b}_i^2, i = 1, 2$. To see this, remember that $f_1^1(b_1) = f_1^2(b_1)$. Also remember that $f_1^k(b_1)$ is (not) binding for both $k = 1$ and $k = 2$. If x_i^{kl} solves $f_1^k(x) = g^l(x)$, then $x_2^{12} = x_2^{22} = \frac{c(1+\alpha_2)}{\gamma(1-\gamma)}$, which, in turn, implies

$$\hat{b}_1^1 = \frac{c(1+\alpha_2)}{\gamma(1-\gamma)} > x_1^{21} = \frac{c(1-\beta_1)}{1-\gamma} \geq \hat{b}_1^2 \text{ and}$$

$$\hat{b}_2^1 = \frac{c(\gamma+\alpha_2)}{\gamma(1-\gamma)} > x_1^{21} = \frac{c(1-\beta_1)}{1-\gamma} \geq \hat{b}_2^2.$$

$\blacksquare$

2.1.3 *Wing* with SSP

As for the solution of *wing* under SSP, we need to add to the basic linear program (7)-(10) the SSP constraint.

Proposition 6 (winiSSP) *The optimal wing mechanism under SSP is (11), with $\beta_1 \leq \beta_2$.*

Proof. We begin by showing that, as in the case of IAP, the optimal *wing* contract in R_1 is (13). This is because, also in this case, $f_1^1(b_1)$ is not binding, since $\tau_1^1 = -\frac{1-\beta_1}{\beta_1} > 1$ and $x_1^1 = \frac{c}{1-\gamma} < \frac{c}{\gamma}$.

On the other hand, the optimal *wing* contract in R_2 is (11), independently of the value of β_1. This is because, given $-1 < \gamma_i < 0$, both τ_1^2 and τ_2^2 are positive. Since $\tau_1^2 = -\frac{1-\beta_1}{\beta_1}$; $|\tau_1^2| > 1$ (i.e., as before, $f_1^2(b_1)$ and $f_2^2(b_1)$ intersect in the first quadrant. Also notice that, given $\beta_i < 0, i = 1, 2, y_1^2 = \frac{c(-1+\beta_1)}{(1-\gamma)\beta_1} < 0$. Two are the relevant cases:

1. If $\beta_1 > \beta_2$, then $f_1^2(b_1)$ and $f_2^2(b_1)$ intersect outside R_2, and the optimal solution would be $b_1 = b_2 = \frac{c(1-\check{\beta})}{(1-\gamma)}$.

2. If $\beta_1 < \beta_2$, then the solution is (11) which overall cost is never greater than $\frac{2c1-\check{\beta})}{(1-\gamma)}$.

We complete the proof by noticing, by analogy with the Proof of Proposition 3, that the optimal solution lies in R_2, rather than in R_1. $\blacksquare$

2.1.4 *Wing* with ESP

In the case of *wing* with ESP, we need to add to the basic linear program (7)-(10) the ESP constraint.

Proposition 7 (winiESP) *The optimal wing mechanism under ESP is (11), with $\beta_1 \leq \beta_2$.*

Proof. We begin by showing that here the optimal *wing* contract in R_1 is (13) if $|\alpha_2| < \gamma$ and $\hat{b}^1 = \left\{\frac{c}{\gamma}, 0\right\}$ if $\beta_2 \geq \gamma$. This is because, like in the previous cases, $f_1^1(b_1)$ is never binding, since $x_1^1 = \frac{c}{1-\gamma} < \frac{c}{\gamma}$ and $\tau_1^1 = -\frac{1-\beta_1}{\beta_1} < 0$. On the other hand, given that $x_2^1 = -\frac{c(1+\alpha_2)}{\alpha_2(1-\gamma)}$ and $0 \leq \tau_2^1 \leq \frac{1}{2}$, $f_2^1(b_1)$ is binding if and only if $|\alpha_2| < \gamma$ (i.e. if $x_2^1 > \frac{c}{\gamma}$).

As for R_2, we begin to notice that $\tau_1^2 = -\frac{1-\beta_1}{\beta_1} \geq -1$ (since $|\beta_1| < \frac{1}{2}$) and that $0 \leq \tau_2^2 = \frac{\alpha_2 + \gamma\beta_2}{1 + \alpha_2 \grave{a}\gamma(1-\beta_2)} < 1$. This implies, like before, that $f_1^2(b_1)$ and $f_2^2(b_1)$ intersect in the first quadrant. The rest of the proof is identical of that of Proposition 6. ∎

2.2 Solution of the mechanism design problem under the *sting* program

In the case of *sting*, the search of the optimal mechanism corresponds to the *wing* linear program (7)-(10) with an additional constraint (implementation with a unique equilibrium):

$$u_1(1,0) \geq u_1(0,0). \tag{18}$$

The constraint (18) makes, on behalf of player 1, the choice of putting effort a weakly dominant strategy.

2.2.1 *Sting* under EP

The solution of *sting* under EP is as follows (see Winter, 2004):

$$b_1^* = \frac{c}{\gamma},$$
$$b_2^* = \frac{c}{1-\gamma}.$$

2.2.2 *Sting* under IAP

Proposition 8 *The optimal sting mechanism under IAP is*

$$\begin{cases} b_1^* = \frac{c((1+\alpha_1)(1+\alpha_2) - \gamma(1-\beta_2))}{\gamma(1+\alpha_1+\alpha_2 - \gamma(1+\alpha_1-\beta_2))}, \\ b_2^* = \frac{c(1+\alpha_1)(\gamma+\alpha_2)}{\gamma(1+\alpha_1+\alpha_2 - \gamma(1+\alpha_1-\beta_2))}. \end{cases} \tag{19}$$

To prove Proposition 8, we follow the same strategy as before.

Lemma 9 $\hat{b}^1 = \left(\frac{c(1+\alpha_2)}{(1-\gamma)\gamma}, \frac{c(\gamma+\alpha_2)}{(1-\gamma)\gamma} \right).$

Proof. In R_1, the constraints for agent 1 and 2 correspond to:

$$b_1 \geq f_1^1(b_1) \equiv \frac{c(1-\beta_1)}{(1-\gamma)\beta_1} - \frac{1-\beta_1}{\beta_1}b_1, \tag{20}$$

$$b_1 \geq f_3^1(b_1) \equiv \frac{c(1-\beta_1)}{\gamma(1-\gamma)\beta_1} - \frac{1-\beta_1}{\beta_1}b_1, \tag{21}$$

$$b_2 \geq f_2^1(b_1) \equiv \frac{c}{1-\gamma} + \frac{\alpha_2}{1+\alpha_2}b_1, \tag{22}$$

Let x_i^{kl} solves $f_i^k(x) = g^l(x)$. We first notice that (20) is not binding. This is because (20) defines a constraint which is parallel to (21), but with a smaller intercept ($y_1^1 < y_3^1$, since $\gamma < 1$). Also notice that, in this case, (21) is not binding either. This is because, $\tau_3^1 < 0$, $\tau_2^1 > 0$, and $x_3^{12} = \frac{c(1-\gamma\beta_1)}{\gamma(1-\gamma)} < x_2^{12} = \frac{c(1+\alpha_2)}{\gamma(1-\gamma)}$.

This implies that, in R_1, $(b_1 + b_2)$ is minimized (like in *wing*) where $f_2^1(b_1)$ and $g^2(b_1)$ intersect, i.e. when $\hat{b}_1^1 = \frac{c(1+\alpha_2)}{(1-\gamma)\gamma}$ and $\hat{b}_2^1 = \frac{c(\gamma+\alpha_2)}{(1-\gamma)\gamma}$. ∎

Lemma 10 *The optimal sting contract in R_2 is (19).*

Proof. R_2, the relevant constraints are as follows:

$$b_1 \geq f_1^2(b_1) \equiv \frac{c(1-\beta_1)}{(1-\gamma)\beta_1} - \frac{1-\beta_1}{\beta_1}b_1 \tag{23}$$

$$b_1 \geq f_3^2(b_1) \equiv -\frac{c(1+\alpha_1)}{\gamma\alpha_1} + \frac{1+\alpha_1}{\alpha_1}b_1 \tag{24}$$

$$b_2 \geq f_2^2(b_1) \equiv \frac{c(1-\beta_2)}{1+\alpha_2-\gamma(1-\beta_2)} - \frac{\alpha_2+\gamma\beta_2}{1+\alpha_2-\gamma(1-\beta_2)}b_1. \tag{25}$$

Notice that, by analogy with R_1, condition (23) is not binding since $\tau_1^2 < 0$, $\tau_3^2 > 0$ and $x_1^1 = \frac{c}{1-\gamma} < x_3^3 = \frac{c}{\gamma}$. Also notice that $0 < x_2^{21} = \frac{c(1-\beta_2)}{1-\gamma} < x_3^{21} = \frac{c(1+\alpha_1)}{\gamma}$ and $x_2^{22} = \frac{c(1+\alpha_2)}{\alpha(1-\gamma)} > x_3^{22} = \frac{c}{\gamma}$. This, in turn, implies that, $f_3^2(b_1)$ and $f_2^2(b_1)$ always intersect in the interior of R_2, which implies the solution.[2] ∎

We are in the position to prove Proposition 8.

Proof. To close the proposition, it is sufficient to show that $\hat{b}_i^1 \geq \hat{b}_i^2, i = 1, 2$. To see this, notice that $x_2^{12} = x_2^{22} = \frac{c(1+\alpha_2)}{\gamma(1-\gamma)}$ (i.e. $f_2^2(b_1)$ and $f_2^2(b_1)$ cross exactly at the intersection with $g^2(b_1)$). Since $\tau_2^2 = \frac{\alpha_2+\gamma\beta_2}{1+\alpha_2-\gamma(1-\beta_2)} > 0$ and $\hat{b}^2$ is interior to R_2, the result follows. ∎

2.2.3 *Sting* under SSP

Proposition 11 *The optimal sting mechanism under SSP is (19).*

Proof. By analogy with the IAP case, in R_1, (20) is not binding. Also notice that $\tau_3^1 = -\frac{1-\beta_1}{\beta_1} > \tau_2^1 = \frac{\alpha_2}{1+a_2} > 0$. Two are the relevant cases:

---

[2] As it turns out, unlike the *wini* case, the search for the appropriate conditions on preferences to identify player 1 has no (algebraically manageable) closed-form solution, but it has to be evaluated numerically (as we did in the calibration of our experimental conditions).

1. if $\alpha_2 \geq -\gamma\beta_1$, (i.e. if $x_3^{12} = \frac{c(1-\gamma\beta_1)}{\gamma(1-\gamma)} \leq x_2^{12} = \frac{c(1+\alpha_2)}{\gamma(1-\gamma)}$), then (24) is not binding, and the optimal solution is the intersection between $f_2^1(b_1)$ and $g_3(b_1)$, that is, $\hat{b}^1 = \left(\frac{c(1+\alpha_2)}{\gamma(1-\gamma)}, \frac{c(\gamma+\alpha_2)}{\gamma(1-\gamma)} \right)$;

2. if $\alpha_2 < -\gamma\beta_1$, then the optimal solution is the intersection between $f_2^1(b_1)$ and $f_3^1(b_1)$, that is, .

$$\hat{b}^1 = \left(\frac{c(1+\alpha_2)(1-\beta_1(1+\gamma))}{\gamma(1-\gamma)(1+\alpha_2-\beta_1)}, \frac{c(\alpha_2+\gamma(1+\alpha_2))(1-\beta_1)}{\gamma(1-\gamma)(1+\alpha_2-\beta_1)} \right).$$

As for R_2, the optimal *sting* contract is, again, (19). This is because, by analogy with the IAP case, conditions (23) and $g^2(b_1)$ are not binding. Also notice that $x_2^{22} = \frac{c(1+\gamma_2)}{\gamma(1-\gamma)} > 0$ and $0 \leq \tau_2^2 = \frac{\alpha_2+\gamma\beta_2}{1+\alpha_2-\gamma(1-\beta_2)} < 1$. This, in turn, implies that, in $R_2, (b_1 + b_2)$ is minimized where $f_3^2(b_1)$ and $f_2^2(b_1)$ intersect, which implies the solution. ∎

2.2.4 *Sting* under ESP

Proposition 12 *The optimal sting mechanism under ESP is (19).*

Proof. By analogy with the previous cases, in R_1, (20) is not binding. Also notice that, in this case, (23) is not binding either, since $\tau_2^1 < 0$ and $x_2^{12} = \frac{c(1+a_2)}{\gamma(1-\gamma)} < \frac{c}{\gamma}$. Since ESP imply $\gamma_1 \leq \frac{1}{2}$, the unique solution in this case is $\hat{b}^1 = \left(\frac{c}{\gamma(1-\gamma)}, 0 \right)$. As for R_2, we first notice that, given that $|\beta_1| \leq \frac{1}{2}, \tau_3^2 > 1$. Since ESP also imply $|\alpha_2| < \gamma$ (i.e. $x_2^2 > \frac{c}{\gamma}$), then the optimal solution is the intersection between $f_2^1(b_1)$ and $g_3(b_1)$, that is, (19). ∎

Social Preferences and Strategic Uncertainty: an Experiment on Markets and Contracts
Appendix B: Additional Experimental Evidence

1 Identifying preferences and beliefs

There are four relevant subsets of parameters studied in the literature, which we now describe. Neither of these specifications consider reciprocal motives (i.e., they always assume $\theta_i = 0$), and, in this sense, they define purely "distributional" preferences.

$$\text{Egoistic Preferences (EP): } \alpha_i = \beta_i = 0. \tag{1}$$

$$\text{Inequality Averse Preferences (IAP): } 0 \le \beta_i < 1, \alpha_i \ge \beta_i. \tag{2}$$

$$\text{Status Seeking Preferences (SSP): } \alpha_i \in [0,1), \beta_i \in (-1,0], |\alpha_i| \ge |\beta_i| \tag{3}$$

$$\text{Efficiency Seeking Preferences (ESP): } \alpha_i \in (-\frac{1}{2}, 0], \beta_i \in [0, \frac{1}{2}), |\beta_i| \ge |\alpha_i| \tag{4}$$

Inequality Averse Preferences (2) were first proposed by Fehr and Schmidt (1999, F&S). They assume that an increased in the difference in payoffs is always disliked, independently on relative positions. Status Seeking Preferences (SP, Rey-Biel, 2008) assume that an increase in the other player's monetary payoff is always disliked, independently of relative positions. Efficiency Seeking Preferences (ESP, Engelmann and Strobel, 2004) assume that a payoff reduction is acceptable only if it is accompanied by an increase (at least of the same amount) in the other player's payoff. Even though C&R follow F&S in only considering IAP, with a slight abuse of terminology we jointly call "C&R distributional preferences" the four types of preferences (1)-(4).

As Figure 2 (in the main text) shows, we do not constrain our distributional parameters (α_i, β_i) to lie in any particular subspace of $\mathbb{R}^2$. Therefore, our estimated pairs $\left(\widehat{\alpha}_i, \widehat{\beta}_i\right)$ can potentially cover all the $\mathbb{R}^2$ space. In this respect, our estimates display significant heterogeneity. In many cases, the constraints on absolute values are violated (in particular, in the case of IAP). This is the reason why we shall identify each distributional preference type to the corresponding quadrant in Figure 2a).

In Table B1 we assign each subject to the quadrant (Q_1 to Q_4) of the $\mathbb{R}^2$ space in which her estimated parameters are most likely to fall considering the confidence intervals of Figure

2b). We also group in an additional EP category those subjects whose estimated α_i and β_i are jointly not significantly different from zero (at the 10% confidence level). Subjects with IAP preferences are a subset of those included in the first quadrant ($\alpha_i > 0, \beta_i > 0$; 19.4% of all the subjects), the pool in Q_2 ($\alpha_i > 0, \beta_i < 0$; 22.2%) includes agents with SSP preferences, while those with ESP preferences fall in Q_4 ($\alpha_i < 0, \beta_i > 0$, 29.2%). For 19.4% of the subjects we cannot reject the null hypothesis of EP.

	EP $\alpha=\beta=0$	Q_1 $\alpha,\beta>0$	Q_2 $\alpha>0,\beta<0$	Q_3 $\alpha,\beta<0$	Q_4 $\alpha<0,\beta>0$
Agents	11 22.9%	8 16.7%	10 20.8%	6 12.5%	13 27.1%
Principals	3 12.5%	6 25%	6 25%	1 4.3%	8 33.3%
Total	14 19.4%	14 19.4%	10 22.2%	7 9.8%	21 29.2%

Table B1. Preference types of agents and principals

In this respect, the (slight) predominance of ESP subjects, is in line to Fisman, Kariv and Markovits' (2007) evidence, when they find a majority of subjects having estimated parameters that indicate a preference for increasing total payoffs rather than reducing differences in payoffs.

2 $Q1$. Is it inequality aversion or strategic uncertainty aversion?

We construct a measure of relative inequality associated to each contract $\bar{k}$ in C_t, relative to the other available options in C_t:

$$\sigma_{\overline{k}} = \frac{\left(b_1^{\overline{k}} - b_2^{\overline{k}}\right) - \min_k \left[b_1^k - b_2^k\right]}{\max_k \left[b_1^k - b_2^k\right] - \min_k \left[b_1^k - b_2^k\right]}, k = 1, ..., 4. \tag{5}$$

By (5), $\sigma_{\bar{k}} \in [0,1]$, i.e. we normalize the inequality induced by each contract relative to the choice set C_t. We thus define $\omega_t = \frac{\sum_{k \in wing} \sigma_k}{\sum_{k \in sting} \sigma_k}$ as a "relative inequality index" associated with the choice of a *wing* vs a *sting* contract in C_t. We then estimate the following logit function:

$$\Pr\left(k_{it} \in sting | \alpha_i, \beta_i, \omega_t\right) = \frac{\exp\left(\psi_0 + \psi_1 \alpha_i + \psi_2 \beta_i + \psi_3 \omega_t\right)}{1 + \exp\left(\psi_0 + \psi_1 \alpha_i + \psi_2 \beta_i + \psi_3 \omega_t\right)},$$

where k_{it} identifies the contract choice of individual i at round t. For players 2 (principals), we use observations from P_2 (P_3).[1] We do so to frame the contract choice problem over the same choice sets, C_t, since in P_3 3 agents' choice sets are determined by principals' decisions. In Table B2 we report the partial maximum likelihood estimates of ψ_0 to ψ_3 with bootstrap standard errors.

---

[1] Player 1 goes for *wing* in P_2 only 7 times out of 288, so that the predicted probability is basically one.

	Phase 2, *player* 2			Phase 3, *Principals*		
	Coeff.	*Std.err.*	*p-val*	*Coeff.*	*Std.err.*	*p-val*
$Constant \rightarrow \psi_0$	-0.060	0.215	0.779	0.493	0.250	0.048
$\alpha_i \rightarrow \psi_1$	-0.864	0.338	0.011	0.329	0.276	0.234
$\beta_i \rightarrow \psi_2$	0.700	0.349	0.045	0.311	0.389	0.424
$\omega_t \rightarrow \psi_3$	21.248	4.919	0.000	11.979	5.269	0.023
Obs.		288			192	

Table B2. S*ting vs. wing* choice in the "mixed" rounds, logit regression

Notice that:

1. Estimated ψ_3 are always positive and significant: the more unequal is the *wing* choice, the more likely is the choice of a *sting* contract. On average, a 1% increase of the relative inequality index ω_t yields an increase of 29% in the probability of choosing *sting* for player 2, and of 14% for the principals. These results are maintained (both in sign and magnitude) if we use a fixed-effects logit model.

2. For principals, distributional parameters are not significant to explain the choice of contract type, while for player 2 in P_2, both α and β are significant, with opposite sign.

3 $Q2$. **Does the social preference model work?**

We now look at how principals' and agents' estimated preferences explain their contract decision, with respect to the two natural dimensions for the problem at stake: *a)* the total cost of the contract $(b_1 + b_2)$ and, *b)* its induced inequality $(b_1 - b_2)$. By analogy with $\sigma_{\overline{k}}$, we define, for each choice set C_t, the following two variables:

$$\tau_{\overline{k}} = \frac{\left(b_1^{\overline{k}} + b_2^{\overline{k}}\right) - \min_k \left[b_1^k + b_2^k\right]}{\max_k \left[b_1^k + b_2^k\right] - \min_k \left[b_1^k + b_2^k\right]}, \overline{k} = 1, ..., 4, \text{ and} \tag{6}$$

$$\rho_{\overline{k}} = \frac{1 + \sigma_{\overline{k}}}{1 + \tau_{\overline{k}}}. \tag{7}$$

We interpret τ, as a measure of *relative efficiency* (or relative cost, from the principal's viewpoint). Consequently, $\rho_{\overline{k}}$ proxies the trade-off agents (principals) face between inequality and efficiency (total costs).

We study principals' contract decisions by regressing $\rho_{\overline{k}}$ and $\tau_{\overline{k}}$ against their own distributional parameters, α_i and β_i. Given that, in both regressions, the dependent variable is bounded both above and below (with round dependent upper and lower bounds), we use a double censored tobit model:

$$y_{it} = \psi_1 \alpha_i + \psi_2 \beta_i + \psi_3 V_{it} + \psi'_4 \boldsymbol{D}_t + v_{it}, \tag{8}$$

where the dependent variable y_{it} refers, alternatively, to the corresponding $\rho_{\overline{k}}$ and $\tau_{\overline{k}}$ induced by the contract choice $\overline{k}$ made by individual i at round t, V_{it} is the randomly generated value for the principal, and $\boldsymbol{D}_t$ is a full set of round dummy variables. Tables B3-4 reports the partial maximum likelihood estimates of the parameters for principals and agents, respectively, with bootstrap standard errors. We estimate the parameters separating the rounds in which the contract menu includes both *sting* and *wing* contracts ("mixed" rounds) from the others ("non mixed").

	Mixed			Non mixed		
Dep.var.:$\tau_{\overline{k}}$	*Coeff.*	*Std.err.*	*p-value*	*Coeff.*	*Std.err.*	*p-value*
$\alpha_i \rightarrow \psi_1$	0.119	0.093	0.201	0.294	0.104	0.005
$\beta_i \rightarrow \psi_2$	0.206	0.143	0.149	0.276	0.184	0.134
$V_{it} \rightarrow \psi_3$	0.002	0.004	0.673	-0.004	0.005	0.472
Left censored	16 (8.6%)			30 (7.8%)		
Uncensored	76 (39.6%)			90 (23.4%)		
Right censored	100 (52.1%)			264 (68.8%)		
Dep.var.:$\rho_{\overline{k}}$	*Coeff.*	*Std.err.*	*p-value*	*Coeff.*	*Std.err.*	*p-value*
$\alpha_i \rightarrow \psi_1$	-0.061	0.034	0.075	-0.191	0.073	0.009
$\beta_i \rightarrow \psi_2$	-0.084	0.061	0.168	-0.203	0.119	0.088
$V_{it} \rightarrow \psi_3$	-0.001	0.002	0.495	0.003	0.003	0.316
Left censored	85 (44.3%)			218 (56.8%)		
Uncensored	68 (35.4%)			138 (35.9%)		
Right censored	39 (20.3%)			28 (7.3%)		

Table B3. Relative cost choice and inequality-total costs trade-off for principals

As Table B3 shows, principals opt for the most expensive contract available more than 50% of the cases (the latter corresponds to the right-censored observations), and more that 2/3 of the cases in the non-mixed rounds. By contrast, less than 10% go for the cheapest one. We explain this evidence by the effect of competition among principals, and consequently the fear of having their offered contract not chosen by any agent. Also notice also that, in the mixed rounds, principals' distributional parameters are only marginally significant in explaining the choice of $\rho_{\overline{k}}$ and $\tau_{\overline{k}}$. This is further indirect evidence of the predominance of the search for robustness we already observed in the *wing/sting* choice. By contrast, in the non-mixed rounds, we see that both principals' distributional parameters significantly explain their preferred $\rho_{\overline{k}}$. They do so in the natural direction: the highest the (inequality-averse) distributional concerns, the lowest the relative inequality, and the highest the relative cost for the principal.

	Mixed			Non mixed		
Player 1	*Coeff.*	*Std.err.*	*p-value*	*Coeff.*	*Std.err.*	*p-value*
$\alpha_i \to \psi_1$	-0.030	0.015	0.048	0.070	0.064	0.272
$\beta_i \to \psi_2$	-0.041	0.021	0.050	-0.381	0.114	0.001
Left censored	101 (35.1%)			313 (54.3%)		
Uncensored	139 (48.3%)			209 (36.3%)		
Right censored	48 (16.7%)			54 (9.4%)		
Player 2	*Coeff.*	*Std.err.*	*p-value*	*Coeff.*	*Std.err.*	*p-value*
$\alpha_i \to \psi_1$	-0.031	0.023	0.178	-0.185	0.090	0.040
$\beta_i \to \psi_2$	-0.043	0.020	0.034	-0.181	0.097	0.062
Left censored	168 (81.1%)			467 (81.1%)		
Uncensored	109 (37.8%)			97 (16.8%)		
Right censored	11 (3.8%)			12 (2.1%)		

Table B4. Inequality-inefficiency trade-off for players in P_2.Dep.var.: $\rho_{\overline{k}}$

As for the agents, we use a regression similar to (8) - where we drop V_{it}, which plays no role here- to study their choice about the inequality/inefficiency trade-off ($\rho_{\overline{k}}$) in P_2. Estimation results, conditional on player positions, are shown in Table B4. Here we generally find -as intuition would suggest- a (negative and significant) relation between distributional concerns and relative inequality.

Moving to stage 2, Table B5 reports the relative frequencies of positive effort decisions in P_2, conditional on subjects' behavior in Stage 1.

	wing contracts			*sting* contracts		
	i is player 1			*i* is player 1		
	$\phi_j = -1$ (51)	$\phi_j = 0$ (288)	*Total* (339)	$\phi_j = -1$ (101)	$\phi_j = 0$ (424)	*Total* (525)
No Dictator	0.48	0.54	0.52	0.91	0.93	0.93
Dictator	**0.11**	0.55	0.49	0.91	0.92	0.92
Total	0.34	0.54	0.51	0.91	0.92	0.92
	i is player 2			*i* is player 2		
	$\phi_j = -1$ (34)	$\phi_j = 0$ (305)	*Total* (339)	$\phi_j = -1$ (86)	$\phi_j = 0$ (439)	*Total* (525)
No Dictator	**0.17**	0.42	0.39	0.50	0.55	0.55
Dictator	0.31	0.48	0.47	0.65	0.70	0.69
Total	0.24	0.46	0.43	0.59	0.62	0.62

Table B5. Relative frequencies of positive effort decisions in P_2

Notice that player 1 misbehaves ($\phi_1 = -1$) in about 14% ($= (34 + 86) / (339 + 525)$) of the times. This frequency is higher in *sting* (16% *vs* 10%). On the other hand, player 2

misbehaves ($\phi_2 = -1$) in 15% (=51/339) of the times in *wing* and in 19% (101/525) in *sting*. Actions following misbehavior are heterogeneous, with instances of reciprocity only in *wing*.[2]

We now compare effort levels across phases. As Table 4 in the main text shows, average effort rates in P_3 are similar to those observed in P_2 (see also Table 3 in the main text): in *sting* player 1 puts effort in 91% of the cases, where the same frequency drops to 64% for player 2. By contrast, in *wing*, both players put effort about 43% of the times. The only noticeable difference between P_2 and P_3 refers to player 1 in *wing*, with 51% of effort decisions in P_2, *vs* 44% in P_3.

4 Robustness check

4.1 Robustness check 1: contract choices

We test whether P_1 choices are different across treatments for a given round. Remember that, in any given round t, the choice set C_t is the same for all treatments. In Table B6 we show the $p - values$ of the Pearson tests for the null hypothesis that, for a given choice set C_t, the distribution of the chosen contracts is independent of the treatment, with statistically significant differences highlighted in boldface. As Table B6 shows, there are very few rounds (3 in the comparison between TR_1 and TR_2; 2 between TR_1 and TR_3; 1 between TR_2 and TR_3) where the choices within rounds are significantly different in any pairwise comparison between treatments.

---

[2]As for *wing*, formal tests of mean equality conditional on player position reject the null at $\alpha = 5\%$ when player 1 is the Dictator and when player 2 is not the Dictator. As for *sting*, the same tests never reject the null.

Round	$H_0{:}TR_1{=}TR_2$	$H_0{:}TR_1{=}TR_3$	$H_0{:}TR_2{=}TR_3$
1	**.0018**	.3854	.1347
2	**.0235**	.1715	.0639
3	.3004	.1106	.1100
4	.5505	.4651	.4138
5	.3236	.8938	.8539
6	.3008	.4241	.6203
7	.5701	.5457	.5461
8	.5522	.6892	.5241
9	.3202	.8168	.7193
10	.1252	.8065	.4702
11	.4208	**.0225**	**.0034**
12	.1545	.6964	.2285
13	.5007	.2638	.2869
14	.3653	.3606	.4578
15	.5914	.1759	.1674
16	.1136	.2497	.9410
17	.3549	.0782	.4753
18	.2710	.1713	.2004
19	.3320	.9735	.7204
20	.2551	.6341	.4839
21	**.0014**	**.0052**	.9137
22	.1392	.6692	.0868
23	.4495	.4832	.9969
24	.3807	.2689	.2435

Table B6. Contract choices within rounds across treatments.

H_0 is that the contract choices in the two treatments under consideration are equal.

In each cell the $p-value$ of the corresponding Pearson test.

Contract choices are basically constant across treatments also in P_2 and P_3. In fact, using the same strategy adopted for P_1, contract choices in Stage 1 of P_2 differ in only 3 rounds in the comparison between TR_1 and TR_2 , 1 between TR_1 and TR_3 and 2 between TR_2 and TR_3. As for principals, choice distributions differ across treatments in 1 round between TR_1 and TR_2 and between TR_2 and TR_3, and in 3 rounds between TR_1 and TR_3.[3]

4.2   Robustness check 2: effort decisions and parameter estimations

In Table B7 we compare effort decisions (Panel A) and outcome distributions (Panel B) across treatments. As Table B7 shows, moving from TR_1 to TR_3 yields an overall decrease

---

[3]Estimation results are not reported here, but are available upon request.

in average effort levels, so that the relative frequency of (in)efficient outcomes gradually goes down (up).

Panel A											
Phase 2											
	Non Dictator						Dictator				
	Player 1			Player 2			Player 1			Player 2		
	wing	sting	Tot	wing	sting	Tot	wing	sting	Tot	wing	sting	Tot
TR_1	0.46	0.74	0.62	0.47	0.72	0.63	0.54	0.84	0.73	0.43	0.77	0.63
TR_2	0.44	0.85	0.66	0.42	0.50	0.47	0.46	0.93	0.76	0.43	0.60	0.52
TR_3	0.34	0.78	0.57	0.27	0.42	0.36	0.34	0.83	0.65	0.37	0.47	0.42
Phase 3											
	Non Dictator						Dictator				
	Player 1			Player 2			Player 1			Player 2		
	wing	sting	Tot	wing	sting	Tot	wing	sting	Tot	wing	sting	Tot
TR_1	0.47	0.76	0.66	0.42	0.75	0.61	0.46	0.81	0.66	0.40	0.77	0.64
TR_2	0.29	0.88	0.64	0.35	0.54	0.47	0.35	0.90	0.69	0.32	0.64	0.51
TR_3	0.25	0.82	0.59	0.32	0.48	0.42	0.18	0.84	0.58	0.20	0.51	0.38

Panel B								
Phase 2								
	wing				sting			
	None	Pl. 1	Pl. 2	Both	None	Pl. 1	Pl. 2	Both
TR_1	0.37	0.20	0.13	0.31	0.06	0.33	0.03	0.59
TR_2	0.37	0.21	0.19	0.24	0.08	0.38	0.03	0.52
TR_3	0.45	0.22	0.21	0.12	0.13	0.43	0.07	0.37
Phase 3								
	wing				sting			
	None	Pl. 1	Pl. 2	Both	None	Pl. 1	Pl. 2	Both
TR_1	0.42	0.14	0.14	0.30	0.07	0.30	0.03	0.61
TR_2	0.52	0.15	0.17	0.16	0.08	0.33	0.03	0.56
TR_3	0.63	0.12	0.16	0.10	0.13	0.38	0.04	0.45

Table B7. Panel A: relative frequencies of positive effort decisions

in Phase 2 and 3 of TR_1 to TR_3.

Panel B: outcome distribution across treatments

As for P_2, player 1 reduces her aggregate effort rate by 15% from TR_1 to TR_2 only in *wing* when she is the Dictator (from 0.54 to 0.46, see Panel A). Player 2, instead, reduces her effort rates more in *sting*: -22% (from 0.77 to 0.60) when she is the Dictator and - 31% (from 0.72 to 0.5) when she is not. Given that, moving from TR_1 to TR_2, player 1's average effort

does not fall significantly, we have a moderate reduction of the cases in which both players put effort (see Panel B): from 30.7% to 24.1% in *wing* (-21%), and from 59.1% to 51.8% in *sting* (-12%). It is in TR_3 where we detect a substantial reduction in the effort rates with respect to TR_1 also for player 1, whose average effort rates are now 26% (37%) lower than in TR_1 in *wing* when she is (not) the Dictator, respectively. In consequence, the all-effort equilibrium is played much less frequently in TR_3, being the outcome of a mere 11.7% of the *wing* cases (-61% wrt TR_1 and -50% wrt TR_2) and 37% of the *sting* cases (-37% wrt TR_1 and - 29% wrt TR_2).

In P_3 player 1's effort rates are constant across treatments in *sting*, while, in *wing,* player 1 reduces her effort rate by 24% (38%) in TR_2 and by 61% (47%) in TR_3 when she is (not) the Dictator, respectively. Just like in P_2, also in P_3 player 2 is the one who reduces average effort the most. In fact, player 2 provides less effort both in *wing* and in *sting*: effort rates drop from about 0.75 in TR_1 to about 0.50 in TR_3 in *sting*, and from 0.40 (0.42) in TR_1 to 0.20 (0.32) in TR_3 in *wing* when she is (not) the Dictator, respectively.

To further investigate on treatment effects in aggregate effort rates, we run the following regression:

$$\delta_{ilt}^k = \gamma_0 + TR_l\gamma_1 + wing_{lt}^k\gamma_2 + \left(TR_j \times wing_{lt}^k\right)\gamma_3 + \varepsilon_i + u_{it} \tag{9}$$

where i identifies the individual; $l = 1, 2, 3$ the treatment, $t = 1, \ldots, 24$ the round; k the selected contract; δ_{ilt}^k is equal one if individual i puts efforts in round t of treatment l under contract k; $wing_{lt}^k$ equals one only if the played contract k is a *wing* one; and TR_l is a treatment dummy. We estimate (9) on P_2 and P_3 effort data, using OLS and clustered standard errors, separately for each possible combination of player position, Dictator assignment, and pairwise comparisons between treatments (that is, we run $2 \times 2 \times 3 = 12$ regressions for each phase).[4] As for P_2, the estimated $\gamma_1 s$ are significantly different from zero only when comparing TR_1 and TR_3 for player 2. In all the other cases, γ_1 and γ_3 are not significantly different from zero. For P_3, we detect significant differences between TR_1 and TR_3 for all player positions and Dictator roles (with player 1 reducing effort the most), while the differences between TR_1 and TR_2 are significant only in the case of player 1 in the role of a Dictator in *sting*.

We now proceed to check different explanations for the fact that, in TR_3, agents put considerably less effort than in TR_1, or TR_2. One potential explanation could be that the situation experienced in TR_3 makes subjects more prone to reciprocate. In Table B8 panel A we first check whether misbehavior differs across treatments. Such a comparison can only be done for non Dictators since, in TR_3, non Dictators' misbehavior is not defined. Panel A shows that misbehavior is more diffused in TR_2 and TR_3, rather than in TR_1. In Panel B we can observe that, in P_2, there is no more punishment in TR_3 than in the other treatments. In fact, the ratio between effort rates in case of mis/well behavior is constant across treatments (about 0.7 with *wing* and 1 with *sting*). Only in the case of *sting* in P_3, where the same ratios go from 0.91 (TR_1) to 0.79 (TR_3), there is some evidence of more reciprocity in TR_3

---

[4]Estimation results are not reported here, but are available upon request.

than in TR_1. In this sense, reciprocity alone cannot explain player 2's lower effort levels in P_2 of TR_3, nor the Dictators' effort reduction in P_3 of TR_3 (see Table B7).

	Panel A: incidence of misbehavior			
	Phase 2			
	wing contracts		sting contracts	
	# cases	% of $\phi_j=-1$	# cases	% of $\phi_j=-1$
Treatment 1	339	14.7	525	16.6
Treatment 2	345	18.0	519	23.3
Treatment 3	359	18.1	505	24.4
	Phase 3			
	wing contracts		sting contracts	
	# cases	% of $\phi_j=-1$	# cases	% of $\phi_j=-1$
Treatment 1	222	11.7	354	12.1
Treatment 2	227	21.1	349	20.3
Treatment 3	233	16.3	343	20.1

	Panel B: effort rates (%)							
	Phase 2				Phase 3			
	wing contracts		*sting* contracts		*wing* contracts		*sting* contracts	
	$\phi_j=-1$	$\phi_j=0$	$\phi_j=-1$	$\phi_j=0$	$\phi_j=-1$	$\phi_j=0$	$\phi_j=-1$	$\phi_j=0$
Treatment 1	34.0	48.4	75.9	72.4	11.5	48.5	69.8	76.5
Treatment 2	30.6	45.6	71.1	64.3	16.7	35.8	52.1	75.9
Treatment 3	23.1	32.7	59.3	58.4	13.2	31.3	53.6	67.9

Table B8. Incidence of misbehavior and effort rates in P_2 and P_3 of TR_1 to TR_3.

Finally, we investigate on the existence of treatment effects by estimating (a suitably modified version of) the structural model of Section 2 in the main text.

Since, in TR_2 and TR_3, player position is constant, we cannot apply the same estimation strategy adopted in Section 2 for TR_1, which exploits P_1 data to identify at the individual level both α and β. Therefore, our estimates here are obtained under the constraints that α and β are constant across individuals, and we combine observations from contract choices in P_1, contract choices (of Stage 1) and effort decision (of Stage 2) in P_2 in a single step maximum likelihood estimation procedure. Furthermore, in TR_3 reciprocity can be estimated only for non Dictators. In other words, the estimates for TR_1 we present here cannot be directly compared with those presented in Section 2 of the main text, both because of the neglected heterogeneity in the social preference parameters (α, β), and because of the different information set used. In addition, constraining the distributional parameters α and β to be constant across individuals remarkably reduces the number of parameters to be estimated (from $72 \times 2 + 3 + 3 = 150$ to $2 + 3 + 3 = 8$). Thus, we expect the estimates to be more precise. Given the experimental design of TR_2 and TR_3 these are the best estimates we can produce

of our structural model to investigate on treatment effects in our estimated parameters.

Distributional preferences	TR_1	TR_2	TR_3
α	0.0739** (0.0363)	−0.1249 (0.1863)	−0.1210* (0.0694)
β	0.6194** (0.0257)	0.6548** (0.0249)	0.6434** (0.0263)
Beliefs $\left(\lambda_i^k\right)$			
$D_i \rightarrow \psi_1$	0.2133** (0.0838)	0.5044** (0.1742)	0.8089** (0.4035)
$b_j^k \rightarrow \psi_2$	−0.0064** (0.0026)	−0.0215** (0.0105)	−0.0333** (0.0087)
$b_i^k - b_j^k \rightarrow \psi_3$	−0.0278** (0.0072)	−0.0152 (0.0121)	−0.0290** (0.0061)
Reciprocity (θ_i)	Treatment 1	Treatment 2	Treatment 3
θ_0	−0.1520** (0.0355)	−0.1925** (0.0389)	−0.2069** (0.0566)
$D_i \rightarrow \theta_1$	0.0562** (0.0218)	0.0718 (0.0463)	
$L_i \rightarrow \theta_2$	0.1664** (0.0402)	0.2110** (0.0559)	0.1863* (0.0979)
$\theta_0 + \theta_1$	−0.0957** (0.0435)	−0.1207** (0.0328)	
$\theta_0 + \theta_2$	0.014 (0.0174)	0.0185 (0.0618)	−0.0206 (0.0652)
$\theta_0 + \theta_1 + \theta_2$	0.0706** (0.0089)	0.0903** (0.0373)	

Table B9. Estimated parameters of belief function and reciprocity by treatment

Cluster adjusted standard errors in parenthesis. * $p-value < 0.1$, ** $p-value < 0.05$

Table B9 and formal tests show that:

1. keeping the player position constant affects the estimate of α, but not that of β. The null hypothesis of a constant β across treatments is not rejected, while αs in TR_2 and TR_3 are significantly lower than that of TR_1. Since the estimate of α is mainly determined by player 2's behavior (and we already know from Section 4.1 that P_1 choices do not differ much across treatments), this is consistent with the evidence of Table B7, where we show that player 2 is the one who changes her effort the most between TR_1 and the other two treatments.

2. The estimated parameters for the difference $b_i^k - b_j^k$ (ψ_3) are always negative: player 1's beliefs are lower than player 2's. This evidence seems fairly robust, as statistical tests never reject the hypothesis that ψ_3s are constant across treatments.

3. In all treatments, Dictators have higher beliefs than their teammates (i.e. $\widehat{\psi}_1 > 0$). Point estimates increase going from TR_1 to TR_3, although we cannot reject the hypothesis that ψ_1 is constant across treatments.

11

4. The effect of teammate's payoff on beliefs $\left(\partial \lambda_i^k / \partial b_j^k = (\psi_2 - \psi_3) \lambda_i^k\right)$ is positive in TR_1, while it is not statistically different from zero in TR_2 and TR_3 (that is, $H_0 : \psi_2 - \psi_3 = 0$ is not rejected for TR_2 and TR_3).

5. The estimated reciprocity parameters are constant across treatments.

To better appreciate the differences in beliefs across treatments, Table B10 reports the predicted beliefs $\lambda_i^k \left(D_i, b_i^k, b_i^k - b_j^k; \widehat{\psi}\right)$ evaluated at the average values of b_i^k and $b_i^k - b_j^k$.

	Treatment 1		Treatment 2		Treatment 3	
	Non Dictator	Dictator	Non Dictator	Dictator	Non Dictator	Dictator
Player 1	0.3010** (0.0374)	0.3477** (0.0312)	0.1838** (0.0692)	0.2717** (0.0611)	0.0812** (0.0380)	0.1656** (0.0230)
Player 2	0.4972** (0.0572)	0.5503** (0.0632)	0.2046 (0.1454)	0.2988* (0.1559)	0.1150* (0.0617)	0.2258** (0.0474)

Table B10: predicted expectations $\lambda_i^k \left(\cdot; \widehat{\psi}\right)$ at the average levels.

For player 1, $\left(\overline{b}_i^k = 75, \overline{b}_i^k - \overline{b}_j^k = 17\right)$, for player 2 $\left(\overline{b}_i^k = 58, \overline{b}_i^k - \overline{b}_j^k = -17\right)$

From Table B10 we learn that subjects' beliefs systematically decrease going from TR_1 to TR_3, for both player positions and Dictator roles. This provides an explanation to the documented reduction in effort rates: subjects reduce their effort essentially because they expect their teammates to put less effort. Consistently with the descriptive statistics in Table B7, Player 2 is the one who decreases her expectations the most going from TR_1 to TR_2, while in TR_3 both Player 1 and 2 reduce their expectations remarkably (by about 75% for the non Dictator and 55% for the Dictator).

References

[1] **Engelmann, D., and M. Strobel.** 2004. "Inequality Aversion Efficiency and Maximum Preferences in Simple Distribution Experiments." *American Economic Review,* 94(4): 857-869.

[2] **Fehr, E. and K.M.Schmidt**. 1999. "A theory of fairness,competition and cooperation." *Quarterly Journal of Economics*, 114: 817-868.

[3] **Fisman, R., Kariv, S. and D. Markovits**. 2007. "Individual Preferences for Giving". *American Economic Review* 97(5): 1858-76.

[4] **Iriberri, N. and Rey-Biel, P.** (2008a). *The Role of Role Uncertainty in Modified Dictator Games*, mimeo.

[5] **Rey-Biel, P.** 2008. "Inequity Version and Team Incentives.", *Scandinavian Journal of Economics*, 108 (2): 297-320.

Appendix C
Experimental Instructions

NOT FOR PUBLICATION

NOTE: In the experiment, the instruction for each PHASE were given only after subjects had played the previous phases.

WELCOME TO THE EXPERIMENT!

- This is an experiment to study how people make decisions. We are only interested in what people do on average.
- Please, do not think we expect a particular behavior from you. On the other hand, keep in mind that your behavior will affect the amount of money you can win.
- In what follows you will find the instructions explaining how this experiment runs and how to use the computer during the experiment.
- Please do not bother the other participants during the experiment. If you need help, raise your hand and wait in silence. We will help you as soon as possible.

THE EXPERIMENT

- In this experiment, you will play for 72 subsequent rounds. These 72 rounds are divided in 3 PHASES, and every PHASE has 24 rounds.

PHASE 1

- In each of the 24 rounds of PHASE 1, you will play with ANOTHER PLAYER in this room.
- The identity of this person will change from one round to the next. You will never know if you interacted with the OTHER PLAYER in the past, nor the OTHER PLAYER will ever know if he has interacted with you. This means your choices will always remain **anonymous**.
- At each round of PHASE 1, the computer will first randomly choose 4 different OPTIONS, that is, four monetary payoff pairs, one for you and one for the OTHER PLAYER. Every OPTION will always appear on the left of the screen.
- Then, you and the OTHER PLAYER have to choose, simultaneously, your favourite OPTION.
- Once you and the OTHER PLAYER have made your decision, the computer will randomly determine who (either you or the OTHER PLAYER) will decide the OPTION for the pair.
- We will call this player the CHOOSER of the game.
- The identity of the CHOOSER will be randomly determined in each round.
- **On average half of the times you will be the CHOOSER and half of the time the OTHER PLAYER will be the CHOOSER.**
- Thus, in each round, the monetary payoffs that both players receive will be determined by the choice of the CHOOSER.

PHASE 2

- In the following 24 rounds of PHASE 2, you will participate in a game similar to the previous one, with some modifications.
- In STAGE 1 of PHASE 2, a payoff matrix will be chosen, and in STAGE 2 of PHASE 2, each pair will face this payoff matrix, which will appear on the left of the screen.

BID	NO	YES
NO	40,40	40+b1/4, 30+b2/4
YES	30+b1/4, 40+b2/4	30+b1, 30+b2

What does this matrix mean?

- In each round, you and the OTHER PLAYER will receive an **initial endowment** of **40** pesetas.
- In each round, you and the OTHER PLAYER have to choose, simultaneously, whether to BID or NOT TO BID.
- Bidding costs 10 pesetas, not bidding does not cost anything.
- You choose the ROW, the OTHER PLAYER chooses the COLUMN.
- Every cell of the matrix (which depends on the monetary payoffs b1 and b2 and your decisions on whether or not to bid) contains two numbers.
- The first number (on the left) is what you win in this round. The second (on the right) is what the OTHER PLAYER wins in this round. There are four possibilities:

  1. If both players bid, both add to their initial endowment their ENTIRE MONETARY PAYOFF b1 or b2 (to which the 10 pesetas cost of bidding will be subtracted).
  2. If you bid, and the OTHER PLAYER does not, **both** players add to their endowment ONE FOURTH of the monetary payoff b1 or b2 (and the cost of bidding will be subtracted from you only);
  3. If the OTHER PLAYER bids, and you don't, **both** players add to their endowment ONE FOURTH of their monetary payoff b1 or b2 (and the cost of bidding will be subtracted from the OTHER PLAYER only);
  4. If nobody bids, you and the OTHER PLAYER will only obtain the 40 pesetas endowment.

PHASE 2 is composed of 2 STAGES:

- In STAGE 1, you and the OTHER PLAYER have to choose your favorite OPTION, that is, the game that you would like to play in STAGE 2.
- After you and the OTHER PLAYER have made your decision, the computer will randomly determine who (either you or the OTHER PLAYER) will be the CHOOSER of the game. That is, the OPTION selected by the CHOOSER in STAGE 1 is the one played in STAGE 2.
- Like in PHASE 1, the identity of the CHOOSER, will be randomly determined in each round.
- **On average, half of times you will be the CHOOSER and half of times the OTHER PLAYER will be the CHOOSER.**
- Once the CHOOSER has determined the option that will be played in this round, you and the other player have to choose whether TO BID or NOT TO BID and the monetary consequences of your decisions are exactly those we just explained.

SUMMING UP

- In each of the 24 rounds of PHASE 2, you will play with ANOTHER PLAYER of this room.

- In STAGE 1, you and the other player, like in PHASE 1, have to choose simultaneously your favorite OPTION.
- After you and the OTHER PLAYER have made your decisions on the OPTION, the computer will randomly determine which one of those OPTIONS is the game that you will play in STAGE 2. That is, the computer designs a CHOOSER.
- In STAGE 2 you and the OTHER PLAYER have to simultaneously DECIDE whether to bid or not to bid. The payoffs of each round depend on your initial endowment of 40 pesetas, on both your choices (to bid or not to bid), on the OPTION chosen by the CHOOSER and on the cost of bidding of 10 pesetas.
- The PAYOFF MATRIX (which will always appear on the left of your screen) sums up, in a compact form, the monetary consequences of your choices.

PHASE 3

- In the last 24 rounds of PHASE 3, you will play in a game similar to the one in PHASE 2 but with some differences.
- Within the 24 persons in this room, the computer will randomly choose two groups of 12.
- In each group of 12 people, the computer will randomly determine 8 **PLAYERS** and 4 **REFEREES.**
- The identity of PLAYERS and REFEREES is randomly determined at the beginning of PHASE 2 and it will remain the same for the rest of the experiment.

PHASE 3 has 3 STAGES.

- Like in the previous PHASES, in STAGE 1 the computer randomly selects 4 OPTIONS, (that is, 4 pairs of monetary payoffs $(b1, b2)$ for the players.
- In addition, in STAGE 1, each REFEREE picks an OPTION within the 4 available for that round (which may be the same or different among them).
- Thus, the 4 OPTIONS selected by the four REFEREES will be proposed to the 8 PLAYERS of their group.
- In STAGE 2, the 8 PLAYERS will be randomly paired. PLAYERS will be rematched at every round.
- Then, just like in PHASE 2, each player has to select one among the 4 OPTIONS proposed by the 4 REFEREES.
- Just like in PHASE 2, the computer randomly determines which of the two OPTIONS chosen by the PLAYERS is played by the pair. That is, the computer designs a CHOOSER.
- Just like in PHASE 2, in the game, both PLAYERS have to choose simultaneously, whether TO BID or NOT TO BID.
- The monetary consequences for the players of their decision are exactly the same as in PHASE 2.

REFEREES' PAYOFF

The REFEREES' payoffs depend on

1. the OPTION they offer,
2. how many REFEREES in their group offer the same OPTION
3. how many CHOOSERS choose the same OPTION
4. Players' actions in the game.

We shall make this clearer with some examples.

CASE 1

- First, suppose that the REFEREE offered an OPTION with payoffs (b1, b2) and that only one CHOOSER has chosen this option.
- The payoff of each REFEREE depends on the positive VALUE randomly generated by the computer and that each REFEREE (and only her) knows, and, in addition, on the sum of the payoffs b1+b2 in the following way:

- if both players bid, the REFEREE wins the difference between his VALUE and the sum of the payoffs; that is, V-(b1+b2);
- if one player bids and the other does not, the REFEREE wins ONE FOURTH of the difference between his VALUE and the sum of the payoffs; that is, $\dfrac{V-(b1+b2)}{4}$.
- if nobody bids, the REFEREE does not win anything.

In this case, the PAYOFF MATRIX for the REFEREE would be as follows:

BID	NO	YES
NO	0	(V-(b1+b2))/4
YES	(V-(b1+b2))/4	V-(b1+b2)

CASE 2

- Suppose now that more than one CHOOSER chose the option that the REFEREE offered. Moreover, suppose moreover that this REFEREE is the only one that picked this OPTION.
- In this case, the REFEREE gets the sum of the payoffs obtained with each couple that chose her OPTION.
- The payoff with each couple will be determined as in CASE 1, taking into account if they bid, if only one bids or nobody bids.

CASE 3

- Suppose now that one or more CHOOSERS chose an option that the REFEREE offered. Moreover, suppose that more than one REFEREE picked the same OPTION. In this case, every single REFEREE that chose the same OPTION gets a payoff with the same structure as in CASE 2, but now, sharing this payoff with the other REFEREES that picked the same option.

CASE 4

- Suppose now that no couple chose the option that the REFEREE offered. In this case, her payoff for this round will be 0.

Figure C1. Phase 1: sample user interface (Round 12, Player 1)

Figure C2. Phase 2, Stage 1: sample user interface (Round 1, Player 1)

Figure C3. Phase 2, Stage 2: sample user interface (Round 1, Player 1)

Figure C4. Phase 3, Stage 1: sample user interface (Round 1, Principals)