Stated and revealed inequality aversion in three subject pools

Benjamin Beranek[1], Robin Cubitt[1], Simon Gächter[1,2]

[1] *University of Nottingham*[*], [2] *CESifo, IZA*

25 March 2015

Abstract. This paper reports data from three subject pools (n=717 subjects) using techniques based on those of Loewenstein, et al. (1989) and Blanco, et al. (2011) to obtain parameters, respectively, of stated and revealed inequality aversion. We provide a replication opportunity for those papers, with two innovations: (i) a design which allows stated and revealed preferences to be compared at the individual level; (ii) assessment of robustness of findings across subjects from a UK university, a Turkish university and Amazon Mechanical Turk. Our findings on stated aversion to inequality are qualitatively similar to those of Loewenstein, et al. in each of our subject pools, whereas there are notable differences between some of our findings on revealed preference and those of Blanco, et al. We find that revealed advantageous inequality aversion is often stronger than revealed dis-advantageous inequality aversion. In most subject pools, we find some (weak) correlation between corresponding parameters of stated and revealed inequality aversion.

Keywords: inequality aversion, replication, revealed and stated preferences, robustness across subject pools, MTurk.

JEL Classification Codes: C90

Acknowledgements

This paper is part of and funded by the research project ERC-AdG 295707 COOPERATION. We also acknowledge support from the ESRC funded Network for Integrated Behavioural Science (NIBS, ES/K002201/1). We received helpful comments from the editor, Nikos Nikiforakis, three referees, and from Abigail Barr, Dirk Engelmann, Lucas Molleman, Pawel Smietanka, Klaus Schmidt, Jonathan Schulz, Till Weber, Ori Weisel, and participants at the ESA conference in Prague and a seminar at the University of Nottingham. We are grateful to Mariana Blanco, Dirk Engelmann and Hans Normann for providing us with the data, zTree code and instructions from Blanco, et al. (2011).

[*] Centre for Decision Research and Experimental Economics, School of Economics, Sir Clive Granger Building, University Park, Nottingham NG7 2RD.
Corresponding author: Simon Gächter; email: simon.gaechter@nottingham.ac.uk.

1. Introduction

Inequality aversion, the dislike of unequal outcomes, has become established as one of the core postulates of behavioural economics. Although discussion of equity concerns is by no means new (e.g., Adams 1965; Selten 1978), the recent literature took off with publication of formal models of inequality averse preferences by Bolton (1991), Bolton and Ockenfels (2000) and Fehr and Schmidt (1999), with the latter paper providing the most widely applied model. In this paper, we are concerned with the *range of empirically relevant* parameters of inequality aversion. We concentrate on aversion to inequality in bilateral monetary comparisons, as in Fehr and Schmidt's theory.

To our knowledge, Loewenstein, et al. (1989) is the first paper to provide systematic evidence on this. They presented their subjects with (hypothetical) life-like scenarios that involved distributions of money between the subject and a comparator and asked the subjects to rate their satisfaction with those distributions. We refer to these measurements as *stated preferences* because they do not involve choices, but un-incentivized assessments of satisfaction. On the basis of these ratings, Loewenstein, et al. estimated 'social utility functions' and found that subjects dislike inequality when it is to their advantage and when it is to their disadvantage. However, aversion to disadvantageous inequality was considerably stronger than aversion to advantageous inequality. Fehr and Schmidt use this observation (pp. 821, 823-4) to justify their distinctive assumption that disadvantageous inequality aversion (measured in their model by a parameter called α) is at least as strong as advantageous inequality aversion (measured by a parameter called β). A key subsequent step taken by Blanco, et al. (2011) was to provide *individual-level* measures of α and β, respectively, by using subjects' choices in two particular games. Thus, importantly, their measures of the parameters reflect *revealed preferences*, in the traditional sense.

In this paper, we replicate the Loewenstein, et al. experiments using updated versions of their scenarios to elicit social utility functions; and we replicate Blanco, et al.'s measurement of α and β using their games and procedures.[1] In the latter case, like Blanco, et al., we will observe the joint distribution of α and β and so be able to reassess the extent to which elicited values are consistent with Fehr and Schmidt's assumption that $\alpha \geq \beta$. However, our most novel contribution is that we link stated and revealed inequality aversion at the individual level: for each subject, our experimental design yields parameters of stated disadvantageous and advantageous inequality aversion obtained with methods akin to those of Loewenstein, et al., and values of α and β revealed by choices using Blanco, et al.'s methods. As they refer to the same inequalities but are obtained with different methods, we use *a* and *b* to denote the stated preference analogues of α and β, respectively.

If inequality aversion is a general sentiment triggered across different situations, then stated and revealed measures should be positively correlated across individuals. If they are, measurements of stated and revealed preferences cross-validate each other. If they are not, this would call into question how strongly the findings of Loewenstein, et al. could support the modelling assumptions of Fehr and Schmidt.

We also investigate the association between inequality aversion and proneness to guilt, which we measure using the GASP (guilt and shame proneness) scale of Cohen, et al. (2011). This is particularly relevant to aversion to advantageous inequality, the parameter of which (here *b* or β) is often referred to as the "guilt" parameter (e.g., Blanco, et al., p. 322).

Our data are from three subject pools, two of them drawn from the student bodies of the University of Nottingham (UK) and Izmir University of Economics (Turkey) and the third

[1] We are aware of three further studies, by Dannenberg, et al. (2007), Dannenberg, et al. (2012), and Teyssier (2012), that elicit revealed preferences using methods akin to Blanco, et al. but with research questions distinct from ours. We are not aware of previous replications of Loewenstein, et al.'s seminal work.

from the American online workforce of Amazon Mechanical Turk (MTurk).[2] Across all subject pools, 717 people participated in our experiments. Apart from differences pertaining to the subject pools, the experimental procedures were essentially uniform.

Our main results are as follows. Notwithstanding some differences in intensity, stated inequality averse preferences are qualitatively similar in all three subject pools in that, like Loewenstein, et al., we find that $a \geq b$ and $b \geq 0$ for most subjects. This provides strong support for the findings of Loewenstein, et al. that inspired Fehr and Schmidt's theory. However, the support for some aspects of that theory itself is weaker, as we find violations at the individual and the median level in all subject pools of the assumption that $\alpha \geq \beta$. We find only weak positive correlation between a and α. Correlation between b and β is significantly positive and exceeds that for a and α in all subject pools. We find females are more averse than males to advantageous inequality and that there is an association between inequality aversion and the GASP measure of proneness to guilt and shame. Although there are some differences between our findings from different subject pools, they are mostly not important for central tendencies of parameters, especially once other factors are controlled for.

2. Methods

For brevity, we focus in this section on the main features of our experimental designs, relegating technical and procedural details, instructions and scenario texts to the online supplementary materials. Each subject completed all of the game tasks described in this section, a selection of scenario tasks, the GASP task and some other tasks with no feedback until all tasks had been completed. We describe the games here in terms of "points", as we did to subjects. At the end of the experiment, points from one game were converted to cash.

[2] MTurk is an online labor market. Horton, et al. (2011) discusses its usefulness for experimental economics. A more sceptical view is expressed by Chandler, et al. (2014).

The core of this study is the two-person version of the Fehr and Schmidt (1999) model of inequality aversion:

$$U_i = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}, j \neq i,\ \alpha_i \geq \beta_i; 1 > \beta_i \geq 0. \qquad (1)$$

In this functional form, U_i denotes person *i*'s utility, x_i person *i*'s monetary payoff and x_j the other person's monetary payoff. The parameter α_i governs *i*'s disutility from disadvantageous inequality, i.e. from x_i falling short of x_j; and the parameter β_i governs *i*'s disutility from advantageous inequality, i.e. from x_i exceeding x_j. A core assumption of Fehr and Schmidt is that advantageous inequality has less of a negative impact on overall utility than disadvantageous inequality of the same magnitude, i.e. $\alpha_i \geq \beta_i$. A central goal of our paper is to provide fresh estimates of the joint distribution of α_i and β_i.

A subject's parameter β_i of advantageous inequality aversion is elicited using a *Modified Dictator game* that Blanco, et al. introduced and which we implement in the same way. The dictator has to make 21 decisions, each a choice between the distribution (20 points for self, 0 points for other) and an equal distribution (*x* points for self, *x* points for other), where "other" refers to a passive player. The equal distributions increased in increments of 1 point from (0, 0) to (20, 20) in the obvious notation. As explained by Blanco, et al. (p. 325-326), the dictator's β_i parameter is theoretically determined by the equal distribution ($\tilde{x}_i$, $\tilde{x}_i$) which he regards as good as the distribution (20, 0). From equation (1), $U_i(20, 0) = U_i(\tilde{x}_i, \tilde{x}_i)$ if, and only if, $20 - 20\beta_i = \tilde{x}_i$. Thus,

$$\beta_i = 1 - \frac{\tilde{x}_i}{20}.$$

Following Blanco, et al., we assume that, as *x* rises in steps, subjects will switch (once) from choosing (20, 0) over (*x*, *x*) to making the opposite choice; and, given this, we approximate $\tilde{x}_i$ as the average of the highest *x* for which (20, 0) is chosen and the lowest *x* for which (*x*, *x*) is

chosen. $\beta_i = 1$ (resp. 0) is assigned to a subject who always (resp. never) chooses the equal option.

Following Blanco, et al. (p. 325) (and in line with a suggestion of Fehr and Schmidt) behavior in the *Ultimatum game* of Güth, et al. (1982) can be used to elicit the parameter α_i of revealed disadvantageous inequality aversion. One player (the proposer) proposes to the other player (the responder) an allocation of a fixed sum (here of 20 points). Then, the responder chooses between accepting the proposal and rejecting it. In the former case, the proposal is implemented; but, in the latter case, both participants receive 0 points. All subjects make decisions in both roles, using the strategy method for the responder's decision, so as to provide a response to all distributions that might be proposed.

A subject's strategy in the role of responder yields an estimate of their α_i parameter. This is determined theoretically by the proposal $\tilde{s}_i$ at which the responder is just indifferent between accepting and rejecting. From (1), $U_i(\tilde{s}_i, 20 - \tilde{s}_i) = \tilde{s}_i - \alpha_i(20 - \tilde{s}_i - \tilde{s}_i) = 0$ determines the point of indifference, and thus,

$$\alpha_i = \frac{\tilde{s}_i}{2(10-\tilde{s}_i)}.$$

Following Blanco, et al., when there is no more than one switch-point in the responder's strategy, we approximate $\tilde{s}_i$ with the average of the lowest accepted offer and the highest rejected offer. Subjects who do not reject any offers are assigned $\alpha_i = 0$ and those who reject all offers less favorable to them than an equal split are assigned $\alpha_i = 4.5$ (Blanco, et al., p. 325).

Our elicitation of α_i and β_i relies on subject *i* "switching" no more than once in the relevant sequence of choices for each parameter. Following Blanco, et al., any subject who

switched multiple times in either of these sequences, and is thus not *well behaved*, is excluded from the data reported in Section 3.

We turn now to the elicitation of *stated* preferences. In Studies 1 & 2 of Loewenstein, et al. (1989), participants read various scenarios describing a range of possible distributions of outcomes, with the subject in the role of one of the affected parties in a bilateral dispute. Subjects ranked their satisfaction of outcomes on an 11-point scale. In our study, we follow the design of Loewenstein, et al., but use modernized scenario tasks. This part of the design is 2×2×2, varying the nature of the issue disputed (distributing the proceeds of an invention or of a plot of land between two parties), the prior relationship between the two parties (positive or negative), and whether it is gains or losses which are to be distributed. Thus, in total there are eight different scenarios. For each one, the task is to rate 21 distributions of payoffs for the subject and the other person described in the scenario. Each subject was presented with four different scenarios. For each subject, the resulting 84 ratings are used to estimate a "social utility" function of the same form as (1) (plus the addition of a constant) to obtain estimates of their stated advantageous and disadvantageous inequality aversion parameters (called a_i and b_i). The estimation used OLS, with the subject's stated satisfaction as the dependent variable. By construction, this procedure produces a value of a_i and a value of b_i for each subject but, as explained above, subjects with non-well-behaved *revealed* preferences are excluded from Section 3. This guarantees that the revealed and stated preferences reported are drawn from the same set of subjects.

Though we may expect positive rank correlation across individuals between a_i and α_i (resp. between b_i and β_i) if inequality aversion is a general sentiment across domains, the presence of such correlation is certainly not built in to the design. The scenario tasks and the games are quite different from one another: for example, only the latter are incentivised, the

response modes are not the same, and the contexts described in the scenarios differ from those posed by the game instructions.

We are interested in the generalizability across subject pools of the findings on stated and revealed inequality aversion. Replication in a culturally different society and outside the university environment is important because there is mounting evidence that student subjects from European or North American universities often are quite special, when compared to others (Henrich, et al. 2010; Barr, et al. 2009; Herrmann, et al. 2008).

We conducted sessions at the University of Nottingham (n=104 students, all British); at Izmir University of Economics (n=206 students, all Turkish); and on the MTurk platform (n= 407 adult residents of the US). Culturally, there is significant distance between the UK and Turkey; and the MTurk sample differs from both university samples, especially by having greater variety of ages and education levels, and by being American.

In both university samples, the experiment was programmed in zTree (Fischbacher 2007) using, in the case of the games, zTree code and instructions provided by Blanco, et al.. In Nottingham, recruitment was done using ORSEE (Greiner 2004). In Izmir, recruiting required approaching students on campus. The MTurk experiments were conducted using the online survey software Qualtrics and the MTurk platform. In all cases, the experiments were followed by post-experimental questionnaires, which elicited socio-demographic information, as well as the measure of guilt and shame proneness (Cohen, et al. 2011).

3. Results

3.1. Stated inequality aversion

Our first result concerns the a_i and b_i parameters of stated inequality aversion derived from the scenarios. We constructed the social utility curves of Figure 1A by averaging the a_i

and b_i parameters across each subject pool (and normalizing the constant to zero). Fig. 1A shows averaged utility as a function of the difference between Own and Others payoff, when one's own payoff is zero. In all three of our subject pools, averaged utility is positively sloped in the region of disadvantageous inequality (to the left of zero on the horizontal axis) and negatively sloped in the region of advantageous inequality. Also, in all three of our subject pools and in line with Loewenstein, et al., the slope of averaged utility is greater in absolute value in the former region than in the latter, implying that disadvantageous inequality had a greater negative impact on satisfaction ratings than the corresponding advantageous inequality.

Fig. 1A also shows that, especially in the region of advantageous inequality, the averaged social utility curves of the Nottingham and Izmir subject pools are quite similar to one another and to corresponding curves from the Loewenstein, et al. findings (see online materials). In contrast, the averaged social utility curves from the MTurk sample show more pronounced aversion to both forms of inequality. Kruskal-Wallis tests confirm that there are statistically significant differences between our subject pools in b_i values ($\chi^2(2)=8.779$, $p=0.0124$) and, especially, in a_i values ($\chi^2(2)=23.858$, $p<0.001$), an issue to which we return in Section 3.4.

FIGURES 1A- 1D

Figs. 1B to 1D illustrate the joint a_i and b_i distributions for each of our subject pools. Recall that a key assumption of Fehr and Schmidt is that $\alpha_i \geq \beta_i$ which they justify referring to the Loewenstein, et al results. In our notation, the corresponding finding to that of Loewenstein, et al. would be tendency for a_i to exceed b_i. We find strong support for this, as $a_i \geq b_i$ for 87%, 77%, and 80% of participants in the Nottingham, Izmir, and MTurk subject pools, respectively. However, a non-trivial minority (35%, 32% and 24% respectively)

violate the condition $b_i \geq 0$, so displaying a stated preference *for* advantageous inequality. This finding is consistent with observation of the existence of spiteful preferences in related literature (e.g., Balafoutas et al., 2012; Iriberri and Rey-Biel, 2013).

Figures 1B–1D also report that a_i and b_i are positively correlated across individuals in each subject pool. In the pooled data, the corresponding Spearman rho is 0.3784 ($p<0.0001$).

3.2. Revealed inequality aversion

Table 1 shows the distribution of α_i (top panel) and β_i (lower panel) for each subject pool, using the categories of Blanco, et al.'s Table 2. We compare our observed distributions to Blanco, et al.'s interpretation (p. 326) of the distribution which Fehr and Schmidt deem plausible and to the distributions which Blanco, et al. themselves observe. Table 1 reports the relevant Chi^2-tests in each case, as well as the mean, median and standard deviation of each parameter in each subject pool.

TABLE 1

The upper panel of Table 1 reveals that, in all three of our subject pools, values of α_i in the range of $\alpha_i < 0.4$ are substantially more frequent than in the Blanco, et al. data (between 46% and 59% of our subjects have an $\alpha_i < 0.4$, compared to 31% in Blanco, et al.). Values of $\alpha_i \geq 4.5$ are, with the exception of Nottingham, also more frequent in our subject pools than in the Blanco, et al. data (7%, 24%, and 17%, in Nottingham, Izmir, and MTurk, respectively, compared to 13% in Blanco, et al.). Chi^2-tests confirm that all three of our subject pools differ significantly (at p=0.03 or lower) in respect of α_i from both the Fehr and Schmidt and the Blanco, et al. distributions of this parameter.

In contrast, our distributions of β_i values appear similar to Blanco, et al.'s findings. Chi^2-tests find no significant difference from the distribution of β_i values reported by Blanco, et al.

in our Nottingham and MTurk samples (p=0.26 and 0.21, respectively), and only a weakly significant difference in the Izmir sample (p=0.08). Comparing our distributions to those assumed by Fehr and Schmidt, using Chi^2-tests, reveals significantly different distributions in the Izmir and MTurk sample (p<0.01), but an insignificant difference between the Fehr-Schmidt distribution and that of our Nottingham sample (p=0.15).

Our next result concerns Fehr and Schmidt's assumption that $\alpha_i \geq \beta_i$. A first, aggregate level, take is provided by comparing the means and medians documented in Table 1. We find that the mean value of α_i is indeed larger than the mean β_i in all our subject pools (as in Blanco, et al.). However, the median α_i is *lower* than the median β_i in all our subject pools (unlike in Blanco, et al.).

Table 1 also shows notable variation in the percent of 'well-behaved' participants (as defined above) in each subject pool. In the Blanco, et al. subject pool, 85% of participants were well-behaved. Our Nottingham and MTurk subject pools displayed similar percentages of well-behaved participants (82% and 90% respectively), but only 45% of our Izmir sample met the criteria of well-behavedness.

The four panels of Fig. 2 give the joint (α_i, β_i) distributions for the Blanco, et al. subject pool and for each of our subject pools. As was foreshadowed in the medians, we see many violations of the assumption that $\alpha_i \geq \beta_i$ in our subject pools. Whereas Blanco, et al. reported 38% of their participants violating this assumption, we find 55%, 59%, and 51% of participants violating it in Nottingham, Izmir, and MTurk, respectively. Like Blanco, et al., we also find that α_i and β_i are uncorrelated in Nottingham and Izmir; in the MTurk sample the correlation between α_i and β_i is slightly (but significantly) negative. In the pooled data, the correlation is very slightly negative (rho = - 0.089; p=0.038).

FIGURE 2

3.3 Relationship between stated and revealed preferences

Fig. 3 shows the joint distribution of a_i (stated) and α_i(revealed) parameters of disadvantageous inequality aversion for each subject pool, with the associated Spearman's rho and its significance level. Surprisingly, there is no significant correlation between a_i and α_i in the Izmir pool; and, though the correlation is statistically significant in the other two pools, it is only rather weakly positive, especially in the MTurk sample. In the pooled data, the correlation is slightly positive (rho = 0.132; p=0.002).

FIGURE 3

The corresponding materials for the joint distribution of b_i (stated) and β_i (revealed) parameters of advantageous inequality aversion are shown in Fig. 4. For these parameters, the correlation is positive and statistically significant in all three subject pools. The degree of correlation is still quite modest, but higher in each subject pool than for a_i and α_i. In the pooled data, the correlation is moderately positive (rho = 0.2785; p<0.001).

FIGURE 4

3.4 The role of socio-demographics and guilt proneness for inequality aversion

Finally, we pool the data from all three subject pools and separately regress our four measures of inequality aversion (a_i, b_i, α_i, β_i) on three standard socio-demographic variables – namely age, a female dummy, and a dummy for having some post-secondary education – and on a dummy for having studied Economics or Business, the GASP scale, and on dummies for Izmir and MTurk (the Nottingham subject pool being the omitted category). Across all our subject pools there is considerable age variation (18-75 years), largely due to the MTurk population. Between 41% and 46% of participants were females, across the three subject pools. All our university student subjects and 82% of the MTurk workers have post-

secondary education status, with 5% having studied Economics and 12% Business. Table 2 records the results of the regressions.

TABLE 2

Age is a significant predictor of stated inequality aversion of both forms, but not of the corresponding revealed preference parameters. Compared with males, females state slightly stronger aversion to disadvantageous inequality aversion, but show significantly higher estimates for stated and revealed advantageous inequality aversion. This result is consistent with the experimental evidence that females give more than men in dictator games (e.g., Eckel and Grossman 1998; Engel 2011). Participants who had studied either Economics or Business showed marginally significantly lower aversion to advantageous inequality (for Economics in revealed preference but for Business in stated preference). Finally, with the exception of α_i, all inequality aversion parameters are highly significantly positively correlated with GASP (higher scores indicate a greater proneness towards guilt and shame). The remaining subject pool difference that stands out once all these factors are controlled for is that Izmir subjects have significantly lower values of a_i.

4. Discussion and conclusion

In terms of replication, our main results are as follows. The qualitative findings of Loewenstein, et al. appear rather robust in that the central tendencies of our *stated* preference data, in each subject pool, support the hypothesis of aversion to advantageous and disadvantageous inequality, with the latter the more keenly felt. Thus, these findings reinforce one of the main ingredients of Fehr and Schmidt's motivation for their model. In contrast, our findings on *revealed* preferences conform less closely to the assumptions of

Fehr and Schmidt's model and to the findings of Blanco, et al., whose revealed preference techniques we use. We find widespread violation of Fehr and Schmidt's assumption that $\alpha_i \geq \beta_i$. Although our results on the distribution of the parameter β_i of aversion to advantageous inequality are similar to the corresponding findings of Blanco, et al., our distributions of the parameter α_i of disadvantageous inequality aversion that differ markedly from that observed by Blanco, et al. Compared with them, we find a notably larger proportion of *low* values of the parameter in all of our subject pools; and also a larger proportion of *high* values of the parameter, in particular in our Izmir sample.

Below, we comment further on two of our most striking findings - weak correlation between stated and revealed preferences and the frequent violation of Fehr and Schmidt's assumption that $\alpha_i \geq \beta_i$, just mentioned – and on differences between our subject pools.

We observe statistically significant positive rank correlations across individuals between (stated) b_i and (revealed) β_i parameters of advantageous inequality aversion in all three of our subject pools; and between (stated) a_i and (revealed) α_i parameters of disadvantageous inequality aversion in two of those pools.[3] But, in all six cases, Spearman's rho was below 0.32, suggesting only a weak relationship. We can think of three possible reactions to this.

One perspective (provided by a referee) is that difference between stated and revealed preferences is an indication of "hypothetical bias" in the former, arising perhaps because subjects do not take un-incentivized tasks seriously or use them to indulge in cheap talk. A second perspective (provided by another referee) is that the difference between stated and revealed preferences, combined with greater conformity of the former to theoretical predictions, indicates that the scenario tasks "work" better, perhaps because subjects find

[3] It is interesting that the correlation is stronger between the "pro-social" parameters β_i and b_i. Studies which report correlation across games between different measures of revealed pro-sociality include (e.g., Yamagishi, et al. (2013), Dariel and Nikiforakis (2014), Peysakhovich, et al. (2014)).

them more recognizable or accessible than the stripped-down lab games. A third perspective is that the scenario tasks and the laboratory games both "work", but they measure slightly different things - in one case, an attitude and, in the other case, willingness to take a certain kind of action. These are correlated because there is an underlying propensity to act on ones attitudes. But, the correlation need not be strong, for example if the propensity to act on ones attitudes is itself a trait whose strength varies across individuals.

To elaborate, the satisfaction ratings of the scenario tasks may indicate subjects' happiness with (or feelings about) different outcomes, whereas the Modified Dictator and Ultimatum games indicate subjects' willingness to sacrifice monetary payoffs in order to change the payoff of the other player in the game. This perspective chimes with the discussion of Blanco, et al. (Section 7) of their finding that the Fehr-Schmidt model, taken with parameter values elicited with their revealed preference methods, predicts the play of games other than those used in the elicitation less successfully at the individual level than at the aggregate level. They point out that willingness to give up money in order to change the other players' payoff may be sensitive to the nature of the game, as well as to the type of inequality faced.

These considerations are also relevant to our findings about the relative strength of aversion to advantageous and disadvantageous inequality. Even if adherence to some ethical codes might induce the opposite attitude, we would expect most subjects to be happier on receiving the larger part of some given unequal allocation between two people than on when receiving the smaller part. If the satisfaction ratings of our scenario tasks are indicators of happiness, in this sense, then our stated preference findings strongly support this expectation. In contrast, our finding that a majority of subjects violate the assumption that $\alpha_i \geq \beta_i$ is a matter of revealed preferences. Viewed more narrowly, it is a matter of the trade-offs that subjects are willing to make in two particular games.

A subject assigned a low value of α_i is one who is reluctant to leave positive offers on the table when playing as respondent in the Ultimatum game. We report more instances of this than most previous studies, but reluctance to leave money on the table is not completely counter-intuitive behavior, even for a subject who feels unhappy about getting less than the proposer. And, of course, *homo economicus* has $\alpha_i = 0$.

In the Modified Dictator game with which β_i is elicited, our findings are comparable with those of Blanco, et al.. Mean values of around 0.5 seem quite high (especially relative to *homo economicus*), but the discussion of Blanco, et al. (p. 333) suggests a possible reason for this shared finding. The active player may feel responsible for the passive player in the Modified Dictator game; and looking out for that player's interests would tend to boost the elicited value of β_i, even for a subject who would not put much weight on the payoff of another in different circumstances.

These arguments suggest that, taken on its own, a finding that some individual violates $\alpha_i \geq \beta_i$ may not be all that surprising, when one keeps in mind that the condition is on revealed preference. Nevertheless, we find more frequent violations than Blanco, et al. had, and this was contrary to our expectations. Further studies would be useful, especially in non-standard subject pools.

That said, the similarities between our findings from distinct subject pools are arguably more striking than the differences, with two exceptions each of which relates to revealed preference. The first is the much greater incidence of non-well-behaved responses to the revealed preference tasks in Izmir than in the other two subject pools. The second is greater incidence among well-behaved subjects of extreme values (high and low) of α_i among the Izmir subject pool, as compared with Nottingham and MTurk. One possible interpretation of these findings is a lower level of understanding of the relevant tasks in some subject pools.

But, we cannot rule out some more fundamental, society-related subject-pool differences (a possibility suggested by Herrmann, et al. 2008).

There is nothing inherently puzzling about one society displaying more extreme values of revealed aversion to disadvantageous inequality than another, especially as this aversion is inferred from the subject's strategy as responder in the Ultimatum game. It may be that, in some societies, there is a strong motivation not to leave money on the table, but this can be over-ridden by a sense of insult and, if it is, then the opposite reaction is also powerful. As Blanco, et al. (Section 7) notes, the Fehr-Schmidt model can be re-interpreted as an indirect reduced-form for reciprocal motivations. Such motivations could affect the aversion to disadvantageous inequality that we infer from the responder's strategy in the Ultimatum game. Thus, a possible explanation of differences between subject pools in this parameter is that they differ either in the strength of their reciprocity or in the consistency across individuals of how they balance reciprocal concerns with pure aversion to inequality.

The interpretation of the Fehr-Schmidt model as a reduced-form for reciprocal motivations is also relevant to points discussed earlier. If the mapping between material inequality and reciprocity is sensitive to context, that might contribute to the weak association which we find between stated and revealed aversion to disadvantageous inequality. To the extent that positive and negative reciprocity are distinct motivations (as is suggested by existing evidence from related ultimatum and dictator games, e.g., Yamagishi, et al. 2012; Peysakhovich, et al. 2014), this perspective would also help to explain why positive and negative inequality aversion, as revealed in the Blanco, et al. tasks are not strongly positively correlated and weaken any expectation that one will always be stronger than the other.

References

Adams, J. S. (1965). *Inequity in social exchange*. New York: Academic Press.

Balafoutas, L., Kerschbamer, R., & Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of Economic Behavior & Organization*, *83*, 125-135.

Barr, A., Wallace, C., Ensminger, J., Henrich, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Lesorogol, C., Marlowe, F., Mcelreath, R., Tracer, D., & Ziker, J. (2009). *Homo aequalis: A cross-society experimental analysis of three bargaining games*. Oxford University Department of Economics Discussion Paper 422.

Blanco, M., Engelmann, D., & Normann, H. T. (2011). A within-subject analysis of other-regarding preferences. *Games and Economic Behavior*, *72*, 321-338.

Bolton, G. E. (1991). A comparative model of bargaining - theory and evidence. *American Economic Review*, *81*, 1096-1136.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, *90*, 166-93.

Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among amazon mechanical turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods*, *46*, 112-130.

Cohen, T. R., Wolf, S. T., Panter, A. T., & Insko, C. A. (2011). Introducing the GASP scale: A new measure of guilt and shame proneness. *Journal of Personality and Social Psychology*, *100*, 947-966.

Dannenberg, A., Riechmann, T., Sturm, B., & Vogt, C. (2007). *Inequity aversion and individual behavior in public good games: An experimental investigation*. ZEW Discussion Paper 07-034.

Dannenberg, A., Riechmann, T., Sturm, B., & Vogt, C. (2012). Inequality aversion and the house money effect. *Experimental Economics*, *15*, 460-484.

Dariel, A., & Nikiforakis, N. (2014). Cooperators and reciprocators: A within-subject analysis of pro-social behavior. *Economics Letters*, *122*, 163-166.

Eckel, C., & Grossman, P. (1998). Are women less selfish than men?: Evidence from dictator games. *The Economic Journal*, *108*, 726-735.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*, 583-610.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*, 817-68.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for readymade economic experiments. *Experimental Economics*, *10*, 171-178.

Greiner, B. (2004). An online recruitment system for economic experiments. In *Forschung und wissenschaftliches Rechnen GWDG Bericht 63*, ed. K. Kremer, & V. Macho. Göttingen: Gesellschaft für Wissenschaftliche Datenverarbeitung.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, *3*, 367-88.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61-83.

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, *319*, 1362-1367.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *14*, 399-425.

Iriberri, N., & Rey-Biel, P. (2013). Elicited beliefs and social information in modified dictator games: What do dictators believe other dictators do? *Quantitative Economics*, *4*, 515-547.

Loewenstein, G., Thompson, L., & Bazerman, M. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, *57*, 426-441.

Peysakhovich, A., Nowak, M. A., & Rand, D. G. (2014). Humans display a 'cooperative phenotype' that is domain general and temporally stable. *Nature Communications*, *5*, doi:10.1038/ncomms5939.

Selten, R. (1978). The equity principle in economic behavior. In *Decision theory and social ethics*, ed. H. Gottinger, & W. Leinfellner. Dordrecht: Reidel.

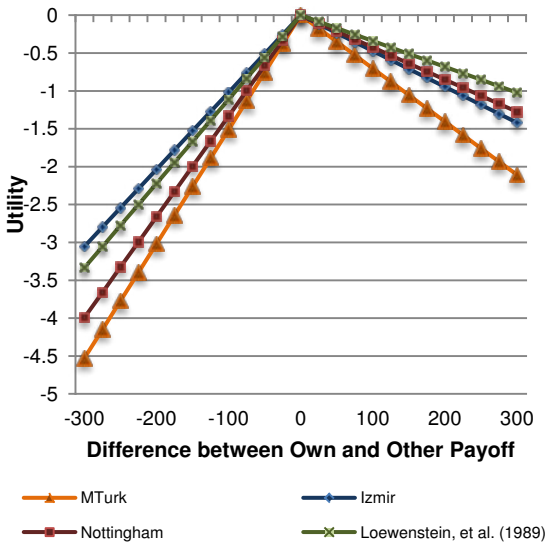
Teyssier, S. (2012). Inequity and risk aversion in sequential public good games. *Public Choice*, *151*, 91-119.

Yamagishi, T., Horita, Y., Mifune, N., Hashimoto, H., Li, Y., Shinada, M., Miura, A., Inukai, K., Takagishi, H., & Simunovic, D. (2012). Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity. *Proceedings of the National Academy of Sciences*, *109*, 20364-20368.

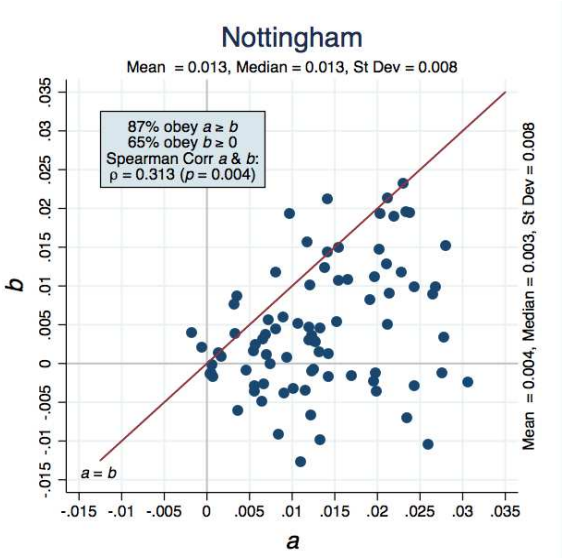
Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., Miura, A., Inukai, K., Tanida, S., Kiyonari, T., Takagishi, H., & Simunovic, D. (2013). Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes*, *120*, 260-271.

Figure 1. Stated Preferences on Aggregate and Individual Levels. Fig. 1A. Utility (satisfaction ratings) as a function of difference between own and other payoff in the scenario tasks. **Figs. 1B to 1D. Joint *a* and *b* distributions per subject pool.** Each dot represents a participant's *a* and *b* parameters as calculated from their stated preferences in the updated Loewenstein, et al (1989) scenario tasks. (The corresponding individual level data for Loewenstein, et. al. are not available.) Observations to the left of the $a = b$ line have $a < b$.

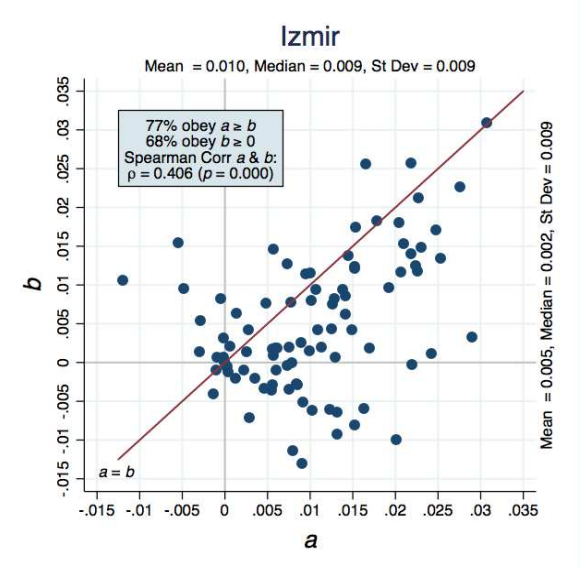
A



B



C



D

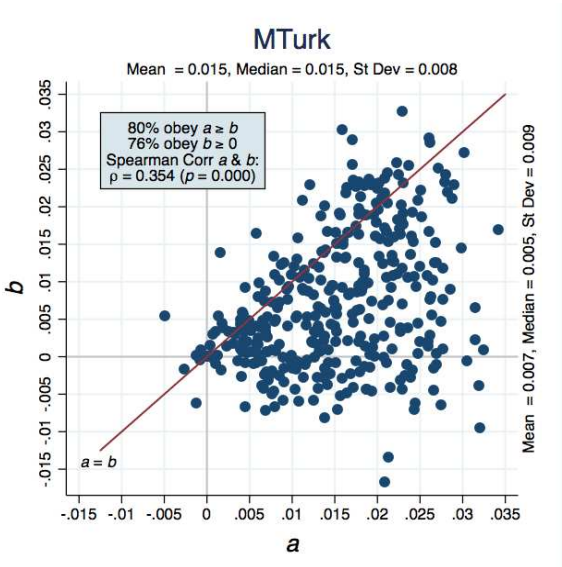


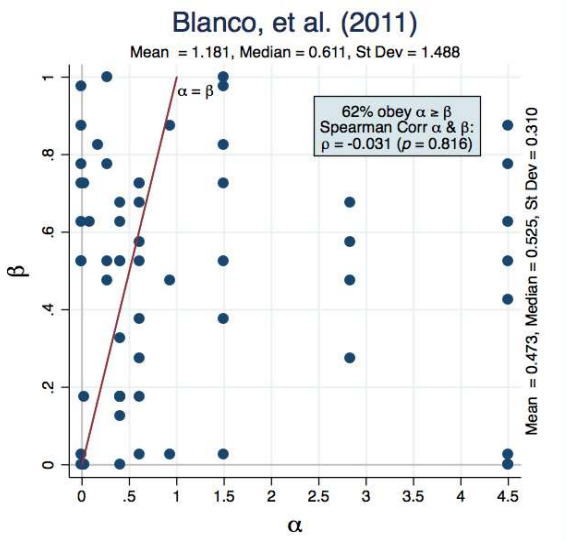
Table 1. Distribution of α and β values. BEN refers to the Blanco, et al. (2011) observed distribution and F&S to Fehr & Schmidt. The data in these two columns and the row classifications are reproduced from Blanco, et al. (p. 325). Percent Well Behaved includes participants who had at most one switching point in the Ultimatum Game and at most one switching point in the Modified Dictator game. Only these participants are included in the analysis of this paper; all others are excluded.

α	F&S	BEN	Nottingham	Izmir	MTurk
$\alpha < 0.4$	30%	31%	54%	59%	46%
$0.4 \leq \alpha < 0.92$	30%	33%	18%	12%	17%
$0.92 \leq \alpha < 4.5$	30%	23%	21%	5%	20%
$4.5 \leq \alpha$	10%	13%	7%	24%	17%
Mean		1.181	0.754	1.227	1.218
Median		0.611	0.269	0.026	0.410
Std. Dev.		1.488	1.198	1.884	1.670
Chi^2 test to F&S	$\chi 2\ (3)$	1.790	11.226	37.751	17.211
	p value	0.618	0.011	0.000	0.001
Chi^2 test to BEN	$\chi 2\ (3)$		9.014	24.933	9.699
	p value		0.029	0.000	0.021

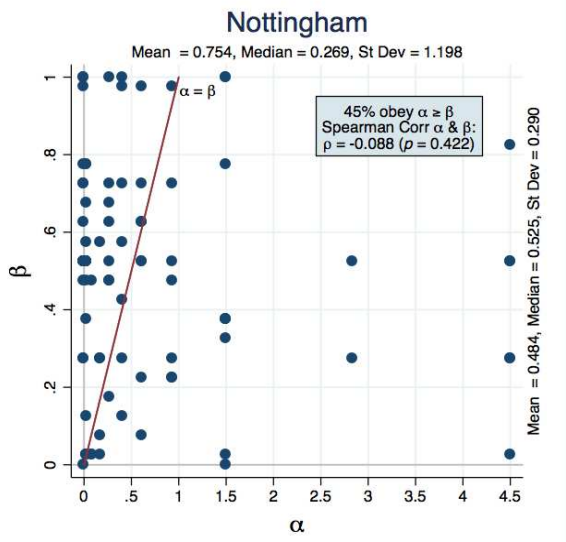
β	F&S	BEN	Nottingham	Izmir	MTurk
$\beta < 0.235$	30%	29%	21%	16%	20%
$0.235 \leq \beta < 0.5$	30%	15%	25%	11%	19%
$0.5 \leq \beta$	40%	56%	54%	73%	61%
Mean		0.473	0.484	0.589	0.512
Median		0.525	0.525	0.575	0.525
Std. Dev.		0.310	0.290	0.315	0.302
Chi^2 test to F&S	$\chi 2\ (2)$	8.51	3.816	21.517	14.491
	p value	0.014	0.148	0.000	0.001
Chi^2 test to BEN	$\chi 2\ (2)$		2.729	5.033	3.109
	p value		0.256	0.081	0.211
Total Sample Size		72	104	206	407
Percent Well Behaved		85%	82%	45%	91%

Figure 2. Joint α and β Distributions. Each dot represents a participant's α and β parameters as calculated from their revealed preferences in the Blanco, et al (2011) games. Observations to the left of the $\alpha = \beta$ line have $\alpha < \beta$ which violates the Fehr and Schmidt (1999) assumption.

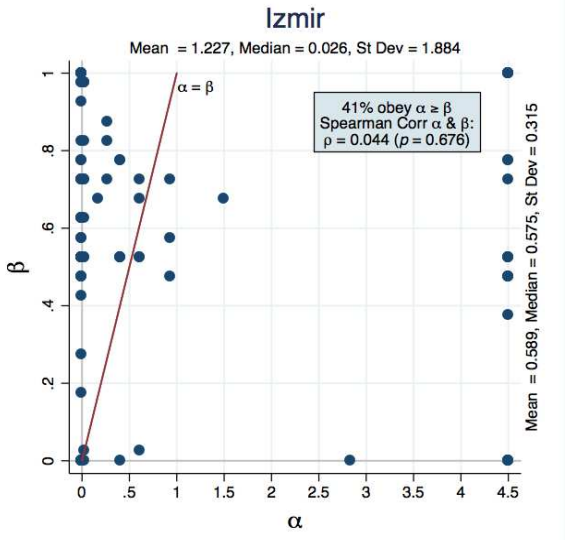
A



B



C



D

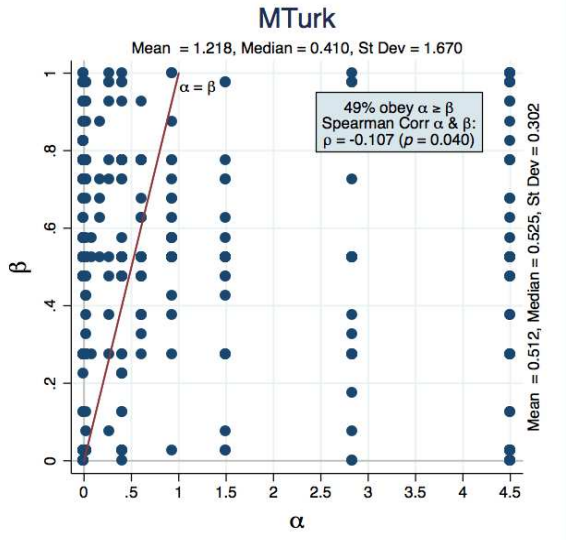
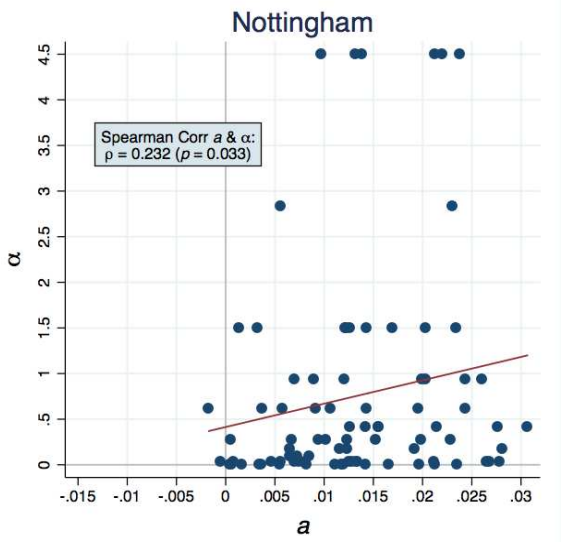
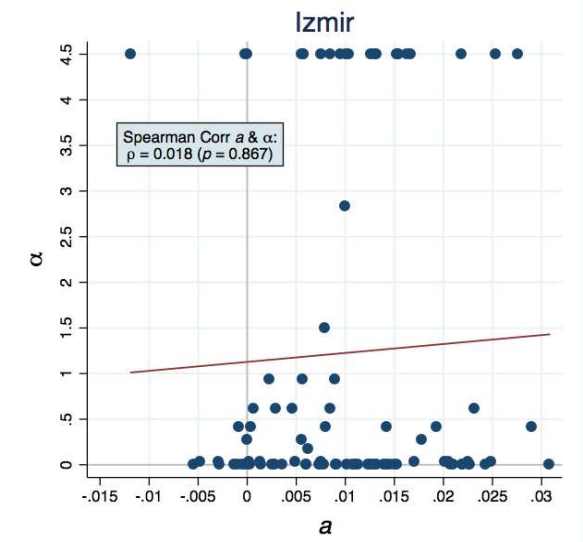


Figure 3. Joint a and α Distributions. Each dot represents a participant's a as calculated from their stated preferences and α parameters as calculated from their revealed preferences. The line results from the linear regression of α on a.

A



B



C

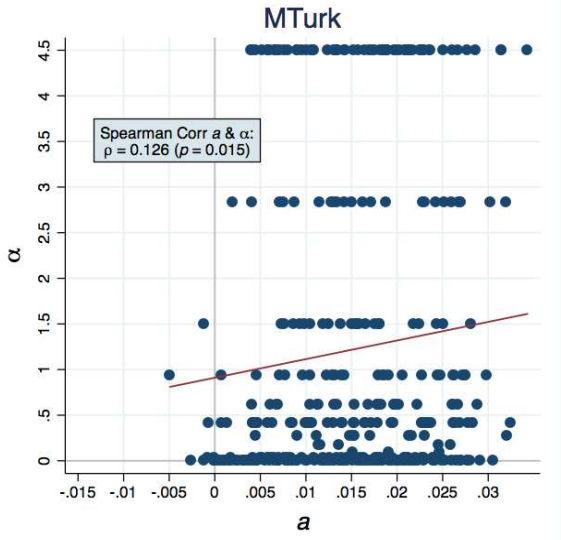
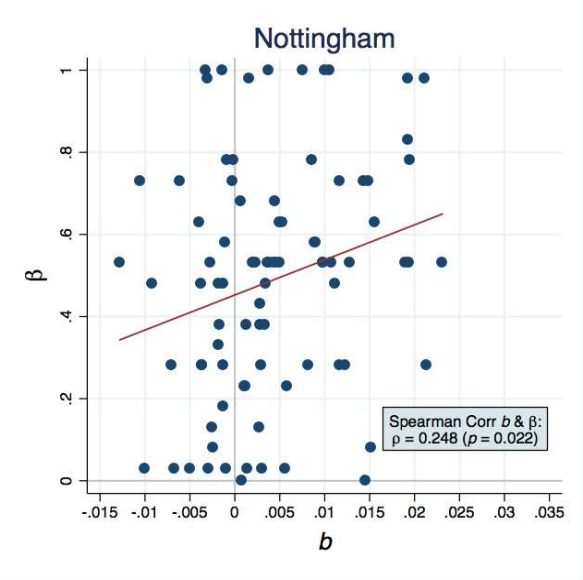
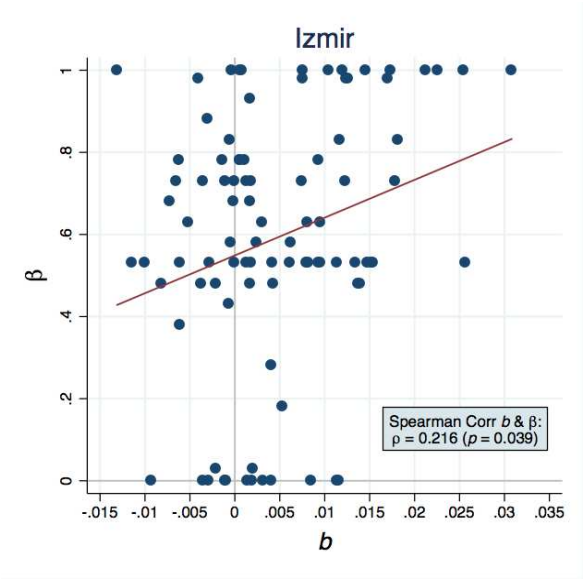


Figure 4. Joint b and β Distributions. Each dot represents a participant's b as calculated from their stated preferences and β parameters as calculated from their revealed preferences. The line results from the linear regression of β on b.

A



B



C

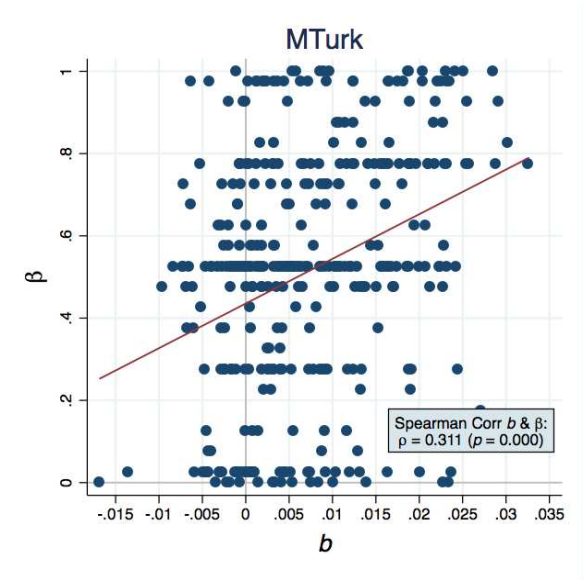


Table 2. OLS regression analysis of demographic and psychological determinants of stated and revealed parameters of inequality aversion.

	a	*b*	α	β
Age	0.0002***	0.0002***	0.0130	0.0016
	(0.0000)	(0.0000)	(0.0086)	(0.0015)
Female	0.0014*	0.0045***	0.1006	0.0674***
	(0.0007)	(0.0007)	(0.1491)	(0.0267)
Post Secondary	-0.0020*	-0.0016	0.0670	0.0269
	(0.0011)	(0.0011)	(0.2308)	(0.0413)
Economics	-0.0021	-0.0017	-0.3219	-0.0939*
	(0.0015)	(0.0015)	(0.3098)	(0.0554)
Business	0.0002	-0.0019*	0.1591	-0.0505
	(0.0010)	(0.0010)	(0.2115)	(0.0378)
GASP	0.0013***	0.0016***	0.0329	0.0572***
	(0.0004)	(0.0004)	(0.0873)	(0.0156)
Izmir	-0.0041***	-0.0003	0.4332*	0.0753*
	(0.0012)	(0.0013)	(0.2540)	(0.0454)
MTurk	-0.0015	-0.0003	0.2382	-0.0239
	(0.0011)	(0.0012)	(0.2362)	(0.0423)
Constant	0.0060***	-0.0060***	0.2695	0.1571*
	(0.0023)	(0.0023)	(0.4690)	(0.0839)
N	546	546	546	546
F (8, 537)	8.88***	13.53***	1.46	5.33***
Adjusted R^2	0.1037	0.1553	0.0067	0.0597

GASP denotes the score from the guilt and shame proneness scale by Cohen, et al. (2011). $* p < 0.10, ** p < 0.05, *** p < 0.01$

Online Supplementary Materials for:
Stated and revealed inequality aversion in three subject pools

Benjamin Beranek, Robin Cubitt, Simon Gächter
University of Nottingham

Contents

I. Section SM.1. Further Details on Procedures

A. Recruitment

We used ORSEE (Greiner (2004)) to recruit subjects in our Nottingham study. In the UK, students do not typically attend the University closest to their childhood home. By restricting our Nottingham sample to UK citizens, we exclude University of Nottingham students from other countries. In Izmir, recruitment was done by approaching students on campus. This occurred in three primary ways: soliciting volunteers from the end of lectures, contacting participants in the school cafeteria and other social places, and via posters advertising the sessions. Recruitment for the MTurk subjects was done through the creation of two separate 200 subject MTurk HITS.

The way that we structure our MTurk HITS is such that subjects find our HIT on the MTurk platform, click through a link to Qualtrics where they complete the experiment via the Qualtrics survey platform, receive a completion code upon finishing, and then return to MTurk where they input the completion code. Seven participants completed the entirety of the Qualtrics survey, but returned to MTurk after the first 200 participants from their respective MTurk HIT had already entered their completion codes and so were unable to be compensated for their efforts. These subjects contacted the researchers via email and we were able to process their payments via a follow up task.

In the cases where there was an even number of these special participants they were paired together with one other participant from their session and payment was calculated according to their and their random partners' choices. In the instance of the 407th subject, we randomly selected one of the first 406 and matched that randomly selected subject with the 407th subject calculating payments as if they had actually been matched together. We paid the 407th subject according to what they would have received

had they actually been matched with the randomly selected subject. We did not double pay that randomly selected subject from the original 406 subjects.

Some MTurk participants started the experiment, but dropped out before completing the entirety of it. We were able to observe the decisions of all MTurk participants – both those who completed the task and those who did not – via the Qualtrics software. 430 subjects clicked through the link from MTurk and entered the Qualtrics survey. 407 of the 430 (94.65%) subjects completed the entirety of the Qualtrics survey and are included in our analysis. 23 of the 430 (5.35%) subjects clicked through the link from MTurk, entered the Qualtrics survey, and did not complete it in its entirety. 12 of 430 (2.79%) [or 12 of 23 (52.17%) people who started but did not complete the Qualtrics survey] did not even begin the Qualtrics survey. The remaining 11 of 430 (2.56%) [or 11 of 23 (47.83%) people who started but did not complete the Qualtrics survey] all completed the Modified Dictator Game after which 2 more dropped out and only 9 of 430 (2.09%) [or 9 of 23 (39.13%) people who started but did not complete the Qualtrics survey] completed the Ultimatum Game. These 9 eventually dropped out and so we did not include them anywhere in our analysis aside from here.

B. Subject Pool Statistics

Table SM.1.1 includes subject pool details for subjects included in the analysis; that is, subjects who had well behaved preferences (as defined in the main text; see also below). The first row gives the proportion of female subjects. Post-Secondary refers to any higher education experience. By default all subjects in the Nottingham and Izmir subject pools meet this requirement while 82% of the MTurk subjects have had some amount of post-secondary education. It has been observed that studying economics or business can impact participants' preferences and for this reason it is included in our

analysis. The mean and standard deviation of age is given as well as the minimum and maximum values.

Table SM.1.1. Subject Pool Details.

	Nottingham	Izmir	MTurk
Female	46%	41%	46%
Post-Secondary	100%	100%	82%
Economics	15%	16%	1%
Business	7%	14%	15%
Age	20.21	21.51	32.94
	(1.91)	(1.71)	(10.12)
Minimum	18	18	19
Maximum	29	30	75
GASP	4.44	5.05	4.89
	(0.92)	(0.78)	(0.85)
Total Sample Size	104	206	407
Percent Well Behaved	82%	45%	91%

In the first three sections of the table, subject pool details are given only for subjects whose revealed preferences were well behaved. GASP denotes the score from the guilt and shame proneness scale by Cohen, et al. (2011). A GASP score = 1 (7) indicates low (high) levels of guilt and shame proneness. The fourth section includes information about all participants.

C. Order of Tasks

The order of tasks in our experimental sessions was the Blanco et al. (2011) Modified Dictator game and Ultimatum game plus two public goods games (greater details given below) followed by the scenario tasks. This order was chosen to preclude spillover effects from the scenario tasks into the games. We assumed that spill-overs from games to scenarios are less likely than from scenarios to games, but, admittedly, we have not tested this assumption. Each participant made decisions in all the roles of each game using the strategy method where necessary. A game was randomly selected for payment at the end of the experiment (the selection of each game had equal chance, as did each assignment of roles). Participants received no feedback until the end of the whole experiment. In the experiment,

we measured earnings in points which were exchanged into local currency at the end of each session.[1]

The public goods games are not analyzed in this paper, but, for sake of completeness, we report what we did. Subjects participated in two two-player one-shot public good games following the same procedures as Fischbacher et al. (Journal of Economic Psychology 33(4), 897-913, 2012). In the first public good game – known as the P-experiment in Fischbacher et al. (2012) – subjects' contribution preferences were elicited using a form of the strategy method. In the second public good game – known as the one-shot C-experiment – subjects' report their belief about their co-player's contribution and then their own contribution.

Subjects did not receive any feedback until after completing the scenario task. The feedback they received was twofold: (1) after all the games and scenario tasks were completed but before the questionnaire subjects learned which of the four games would be payoff relevant (in the Nottingham and Izmir experiments; in the MTurk experiment they received this information after completion of the MTurk HIT) and (2) only after completion of the questionnaire did participants see feedback and at this time they received feedback only on the payoff relevant game (again for the MTurk subjects, they received this information after the completion of the MTurk HIT).

D. Assigning Revealed Preferences Values

We also follow Blanco, et al. in assigning $\beta_i=1$ (resp. $\beta_i=0$) to subjects who always (resp. never) choose the equal distribution; and in excluding subjects with multiple switch points in the Modified Dictator Game from the analysis. We likewise follow Blanco et al. and assign

[1] In the UK, 1 point = 0.75 pence; in Turkey, 1 point = 1.5 Turkish Lira; and on Amazon Mechanical Turk, 1 point = 17.5 cents. Exchange rates were chosen to equalize purchasing power as much as possible in the student subject pools. Since MTurk is a naturally occurring work-place, a different payment structure (with a higher participation fee) was used, to conform to its conventions.

participants who do not reject any offers an $\alpha_i=0$ and participants who reject every offer below the equal split, an $\alpha_i=4.5$ (although in theory these participants could have $\alpha_i \geq 4.5$). Subjects with multiple switch points in the Ultimatum Game are excluded from the analysis in both the main paper and supplementary materials. In summary, subjects were excluded from the analysis by having:

- Multiple switch points in the Modified Dictator Game alone
- Multiple switch points in the Ultimatum Game alone
- Multiple switch points in both the Modified Dictator and Ultimatum Games

II. Section SM.2. Full Instructions of the Experiment

A. Background

Included here are the full instructions for the Nottingham subject pool. These instructions were based off of those by Loewenstein et al. (1989) for the scenarios and Blanco et al. (2011) for the Modified Dictator game and Ultimatum game. Blanco et al. (2011) kindly provided instructions and copies of the zTree files which were used in this experiment with only minor modifications. The text and all the materials including the zTree files for the Izmir subject pool were translated (both forward and reverse) into Turkish. In both Nottingham and Izmir, the experiments were conducted by native speakers and supervised by one of the authors [Beranek]. The Turkish version of scenarios and instructions, as well as all zTree files are available upon request.

The MTurk participants completed the experiment online using the online survey software Qualtrics via MTurk and the text and materials were Americanized. These materials are available upon request in the form of a PDF file.

B. Instructions

Economic Research Project

You are now taking part in an experiment. If you read the following instructions carefully, you can, depending on your and other participants' decisions, earn a considerable amount of money. It is therefore important that you take your time to understand the instructions. Please do not communicate with the other participants during the experiment. Should you have any questions, please ask us.

The experiment consists of four different sections. In each section you will be called to make one or more decisions. You will have to make your decisions without knowing other participants' decisions in the previous sections. Note further that the other participants will not know your decisions either.

Only one of the sections will be taken into account in determining your final payoff. This will be randomly determined as described below. Each section has the same probability of being selected. You should take your time to make your decision. All the information you provide will be treated anonymously.

The section that will be taken into account in determining your final payment will be selected as follows. **Participant Number 2** was randomly chosen at the very beginning of the experiment. This participant will draw a ball from a cloth bag **after** all participants have completed all sections. Each ball in the cloth bag has a different colour and each colour corresponds to a different section: yellow, blue, green, and red. The resulting colour and corresponding section will be used to calculate your payment.

The computer will randomly pair you with another participant in the room and will assign the roles. The matching and roles assignment will remain anonymous. You will not know which role you were playing until the end of the game.

Your earnings will be paid to you in cash at the end of the experiment at a rate of 1 point = 50 pence. Earnings will be confidential.

Yellow Section

In this section the situation is as follows:

> Person A is asked to choose between two possible distributions of money between her and Person B in twenty-one different decision problems. Person B knows that A has been called to make those decisions, and there is nothing he can do but accept them.
>
> The roles of Person A and Person B will be randomly determined at the end and will remain anonymous

Before making your decisions please read carefully the following paragraphs.

The decision problems will be presented in a chart. Each decision problem will look like the following:

Person A's Payoff	Person B's Payoff	Decision	Person A's Payoff	Person B's Payoff
20	0	Left Right	5	5

You will have to decide as Person A; hence if in this particular decision problem you choose left, you decide to keep the 20 points for yourself so Person B's payoff will be 0 points. Similarly, if you choose Right, you and the Person B will earn 5 points each.

You will need to choose one distribution (Left or Right) in each of the twenty-one rows you will have in the screen. If this is chosen as the payoff relevant section, the computer will randomly choose one of the twenty-one decisions. The outcome in the chosen decision will then determine your earnings.

The computer will randomly pair you with another participant in the room and will assign the roles. The matching and roles assignment will remain anonymous.

Please note that you will make all decisions as Person A but the computer might assign you Person B's role.

If you are assigned the role of A, you will earn the amount that you have chosen for Person A in the relevant situation and the person paired with you will earn the amount that you have chosen for Person B.

In the case that you are assigned the role of Person B, you will earn the amount that Person A whom you are paired with has chosen for Person B in the relevant situation.

Blue Section

In this section the situations is as follows:

> Person A is asked to choose one out of twenty-one possible distributions of money between her and Person B. Person B knows that A has been called to make these decision, and may either accept the distribution chosen by A, or reject it.
>
> In the case that Per*son B accepts A's proposed distribution, that will be implemented.* If B rejects the offer, both receive nothing.
>
> The roles of Person A and Person B will be randomly determined by the computer and will remain anonymous.

Before making your decision please read carefully the following paragraphs.

In the case that this section is selected to determine your earnings, the computer will randomly pair you with another participant in the room and will assign the roles. The matching and roles assignment will remain anonymous.

You will have to make decisions as if you were Person A and also as if you were Person B. In the latter case, you will have to decide whether you accept or reject each of A's possible twenty-one proposed distributions.

If you are assigned the role of Person A you will earn the payoff you chose for yourself if the Person B that you are paired with accepts your offer. Otherwise, you both will earn nothing.

If you are assigned the role of Person B, you will earn the payoff that the Person A that you are paired with chose for B, only if you had accepted that particular offer. Otherwise, you both earn nothing.

Green and Red Sections

[These sections were unrelated to this paper.]

Scenarios and Questionnaire

Scenarios

In this section of the project you will read two different scenarios. After reading each scenario, you will learn the outcome of the situation with a variety of payoffs for you and another party. Your task in this section is to rank your satisfaction with the various payoffs to yourself and the other party on a scale from very unsatisfied (-5) to very satisfied (5). Keep in mind that the order of the payoffs is randomly displayed, so you should be certain to rank your satisfaction of each outcome according to its corresponding payoffs listed to the left of the radio button input scale.

Questionnaire

While calculating your payoff, we would like to ask you to answer the following questionnaire.

Please answer each of the following questions as accurately as possible. Of course, your answers will be treated confidentially. Your honest answers will be of immense value for our scientific investigation. Thank you in advance for your cooperation.

III. Section SM.3. Determining Stated Preferences of Inequality Aversion.

A. Design of the Loewenstein et al. Experiment

The scenario tasks provide a near replication of the Loewenstein et al. (1989) (henceforth "LBT89") Study Two scenario tasks with a few exceptions noted below. Participants are asked to rate their satisfaction for outcomes to two scenario disputes. Not every participant faces the same scenarios; dispute type and relationship condition vary across the treatments. The disputes are regarding the gains or losses from disputes involving an invention and from the mutual ownership of a plot of land.

In the original LTB89 paper, the invention scenario regarded the development of cross-country water skis. We developed an alternative invention scenario regarding the development of a smartphone application which is identical in structure to the 1989 scenario, but we expect to be more readily comprehensible to our subjects.

The relationship condition is either a positive or a negative condition and is elaborated in the scenario descriptions. In the MTurk sample, we also included a third condition where there was no relationship manipulation; that is, the nature of the relationship was not mentioned. We refer to this condition as neutral.

We made small adaptations to the scenario text to reflect the individual characteristics of the subject pools – we have an Anglicized version for the Nottingham subject pool, an Americanized version for the MTurk subject pool, and a Turkish version for the Izmir subject pool. The text and all the materials of the Izmir subject pool were translated (both forward and reverse) into Turkish. In both Nottingham and Izmir, the experiments were conducted by native speakers. The scenario text for the Nottingham subject pool is included below. Complete scenario text for the other two subject pools are available upon request.

This was a 2x2x2 design and participants were randomly assigned to each dispute and relationship condition in such a way that they rated both gain and loss conditions for either (a) the invention dispute with a positive relationship condition and the plot dispute with a negative relationship condition or (b) the invention dispute with a negative relationship condition and the plot dispute with a positive relationship condition. For each of the four scenarios, the task is to rate 21 distributions of payoffs for the subject and another person described in the scenario. Each subject is presented with four (out of the eight) scenarios and therefore asked for a total of 84 ratings on a scale from -5 representing "very unsatisfied" to 5 representing "very satisfied."

The gain conditions are classified as 300, 500, and 600 received to self while the positive outcomes to the other player range from 0 to 900. The loss conditions are the same unit amounts expressed as amounts to pay and not profit. Following the procedures outlined by LTB89, the outcome pairs are randomly ordered to avoid automatic responding. The zTree screen shots from the invention dispute as presented to the Nottingham subject pool are included below in Figure SM.3.1.

We also included a neutral relationship condition in the American MTurk sample to see what extent the relationship frame impacted utility ratings. In those cases, no relationship information was given to the participants. In this case, this was a 2x3x2 design and participants were randomly assigned to dispute and relationship conditions.

Table SM3.1 summaries the sequences detailing the dispute and relationship conditions present in each as well as referencing the output from the OLS estimations which are included in the Scenario Estimates tab of the BCG_Data file.

Table SM.3.1. Summary Scenario Sequences and Resulting OLS Estimates

Dispute	Invention		Plot		Invention	Plot
Relationship	Positive	Negative	Positive	Negative	Neutral	Neutral
Gain (21 rankings)	Sequence 1A	Sequence 2A	Sequence 2C	Sequence 1C	Sequence 3A	Sequence 3C
Loss (21 rankings)	Sequence 1B	Sequence 2B	Sequence 2D	Sequence 1D	Sequence 3B	Sequence 3D
BCG_Data file – Scenario Estimates tab	Scen1PosDiff Scen1NegDiff	Scen2PosDiff Scen2NegDiff	Scen3PosDiff Scen3NegDiff	Scen4PosDiff Scen4NegDiff	Scen5PosDiff Scen5NegDiff	Scen6PosDiff Scen6NegDiff

- Subjects participated in one of three sequences. In each, they first read the Invention Dispute and then ranked their satisfaction with 21 gain distributions (either Sequence 1A, 2A, or 3A) and then 21 loss distributions (either Sequence 1B, 2B, or 3B).
- Next, they read the Plot Dispute and then ranked their satisfaction with 21 gain distributions (either Sequence 1C, 2C, or 3C) and then 21 loss distributions (either Sequence 1D, 2D, or 3D).
- The sequences varied according to relationship condition: Sequence 1 had a positive relationship frame for the invention dispute and a negative relationship frame for the plot dispute; Sequence 2 had a negative relationship frame for the invention dispute and a positive relationship frame for the plot dispute; and Sequence 3 had neutral relationship frames for both (only half the MTurk participants participated in sequence 3).
- We used each of these rankings as the dependent variable in an OLS estimation for the functional form below with the independent variables being own payoff and the difference between own and other payoff. Each of the sequences resulted in four different OLS parameter estimates (two for NegDiff and two for PosDiff) that can be found in the Scenario Estimates tab of the BCG_Data file which is available as a supplementary file in Excel format.
- The two NegDiff (PosDiff) estimates are averaged together between the scenarios in order to create the Stated a (Stated b) variables (see discussion in section SM.3.C – particularly page 26 – for further explanation and Table SM.3.2 for evidence supporting this procedure).

B. Scenario Texts

The scenarios were structured in the following way: first, the dispute is introduced; second, a relationship condition is introduced – positive or negative (or neutral in the MTurk subject pool); third, subjects rank their satisfactions with 21 gain distributions and 21 loss distributions. Participants do this entire sequence with two different scenarios for a total of 84 satisfaction ratings.

1. **Smartphone App Scenario (Updated 2012 version of the 1989 Patent Scenario) with Moderate Relationships, Anglicized**

    a. Dispute: "One day while eating lunch, a student who lives in your residence hall, Charlotte, mentioned to you an idea for a new Smartphone app: a classroom note application for your smartphone. It is similar to a normal word processing app except that you can record lectures, draw diagrams, and take photographs of PowerPoint slides all in real time. Charlotte thought of the idea several years ago, but had not done anything with it and had not been able to interest anyone in it. You find the idea of a classroom note Smartphone app exciting. You suggest to Charlotte that the two of you work together on the project. Over the next month you spend long hours together constructing a prototype of the classroom note app in the computer room. Since it was Charlotte's idea, you agree to pay the rent for the computer room space while you make the app. After extensively testing and refining the classroom note app at your university, you decide that you are ready to submit the app to the Smartphone app store. You complete the Smartphone app store submission, pay the registration fee, and send the app in for approval."

    b. Relationship:

        i. Moderate positive relationship: "Charlotte is a student in your residence hall. You like Charlotte a lot, and other people in the dorm also consider Charlotte to be very nice. Charlotte takes notes and picks up assignments for people who miss classes. Last week, Charlotte made all the arrangements for a small hall party and

offered her room to you and your out-of-town guest while she was away over the weekend. In short, Charlotte is kind, friendly, and dependable."

ii. Moderate negative relationship: "Charlotte is a student who lives in your residence hall. You have had several unpleasant personal experiences with Charlotte, and other people in the hall also consider Charlotte to be quite rude. Charlotte borrows notes and copies assignments, but does not say thank you and often fails to return items. Last week, Charlotte did not show up for an important intra-mural tournament game and insulted one of your friends. In short, Charlotte is selfish, irresponsible, and argumentative."

c. Outcome:

i. Gain: "Several weeks after you submitted the classroom note app to the Smartphone app store, you learn that your Smartphone app has not been approved because there are already similar apps that do the same thing. However, the app store has contacted the developer of one of these similar apps and she is interested in buying one of the innovative features incorporated in your design. You and Charlotte agree that the amount offered seems reasonable. The two of you negotiate how to split the profit."

ii. Loss: "Several weeks after you submitted the classroom note app to the Smartphone app store, you learn that your Smartphone app has not been approved because there are already similar apps that do the same thing. Nevertheless, you are responsible for paying for

the Smartphone app store registration fees. Both you and Charlotte receive copies of this bill and negotiate how to split the cost."

2. **Plot Scenario, with Moderate Relationships, Anglicized**
    a. Dispute: "You live adjacent to an empty plot separating you from your next-door neighbours to your left. No one knows who owns the plot, despite the fact that you and your next-door neighbours have lived there for more than 2 years. However, the local council recently informed you that the plot actually belongs to both you and your neighbours, but the percentage owned by each of you has to be negotiated."
    b. Relationship:
        i. Moderate positive relationship: "The Smiths are your neighbours. You like the Smiths a lot, and other neighbours consider the Smiths to be very kind as well. The Smiths always are available to help others. The Smiths are more than happy to water plants and take delivery of parcels when you're away. Last week, the Smiths loaned you some very expensive tools for a repair project and offered their guest bedroom for one of your out-of-town guests. In short, the Smiths are kind, friendly, and dependable."
        ii. Moderate negative relationship: "The Smiths are your neighbours. You have had several unpleasant experiences with the Smiths. Your other neighbours also consider the Smiths to be quite rude. The Smiths borrow things like tools and dishes, but they do not say thank you and often fail to return items. Last week, the Smiths blocked your driveway with their car and threatened to call the

police on a small party you were having. In short, the Smiths are selfish, irresponsible, and argumentative."

c. Outcome:

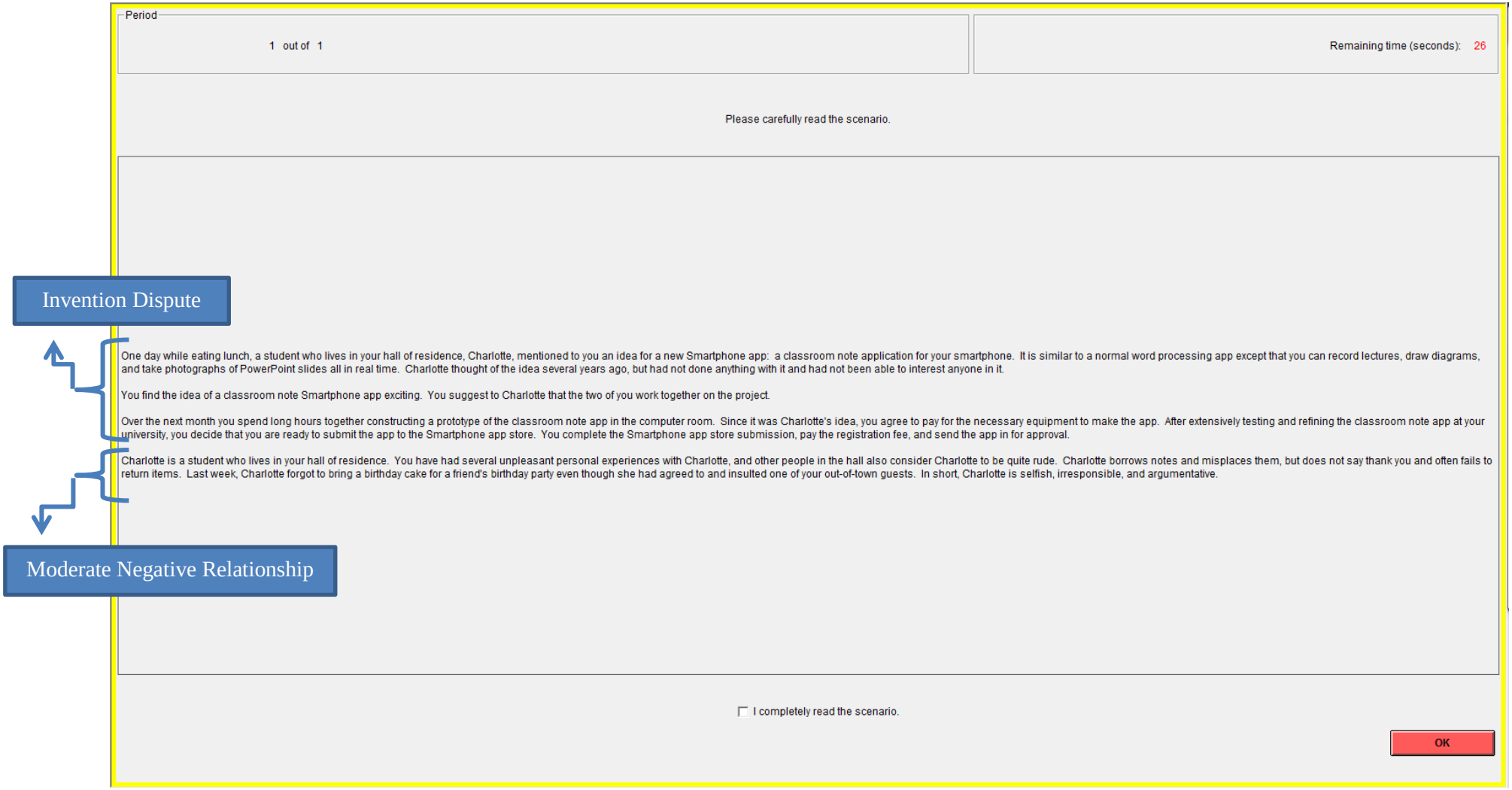
   i. Gain: "A third neighbour who lacks a garden has agreed to buy the plot. You and your neighbours would both be happy to have a garden between your houses. You and your neighbours need to decide how to split the profits."

   ii. Loss: "The plot is too small to sell. However, the local council has amassed taxes on the property that you and your neighbours must pay. You and your neighbours need to decide how to split the costs of the taxes."

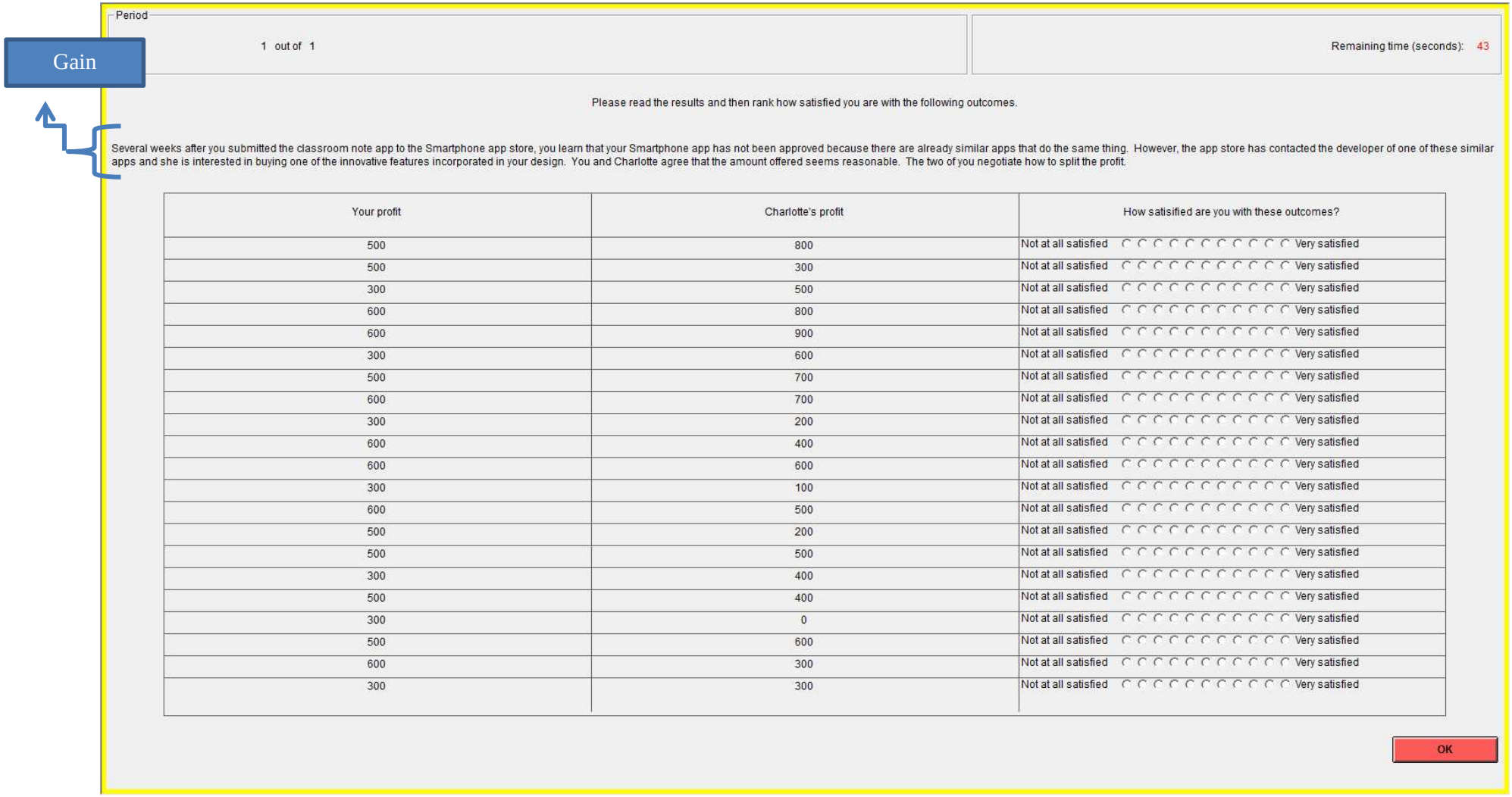
The following Fig. SM3.1 provides an example of the z-Tree screen shots subjects saw in Sequence 2. The screen shots of the other sequences are identical except for relevant differences in text.

Figure SM.3.1. Example Screen shots of zTree Scenario Decision Screens for Sequence 2.

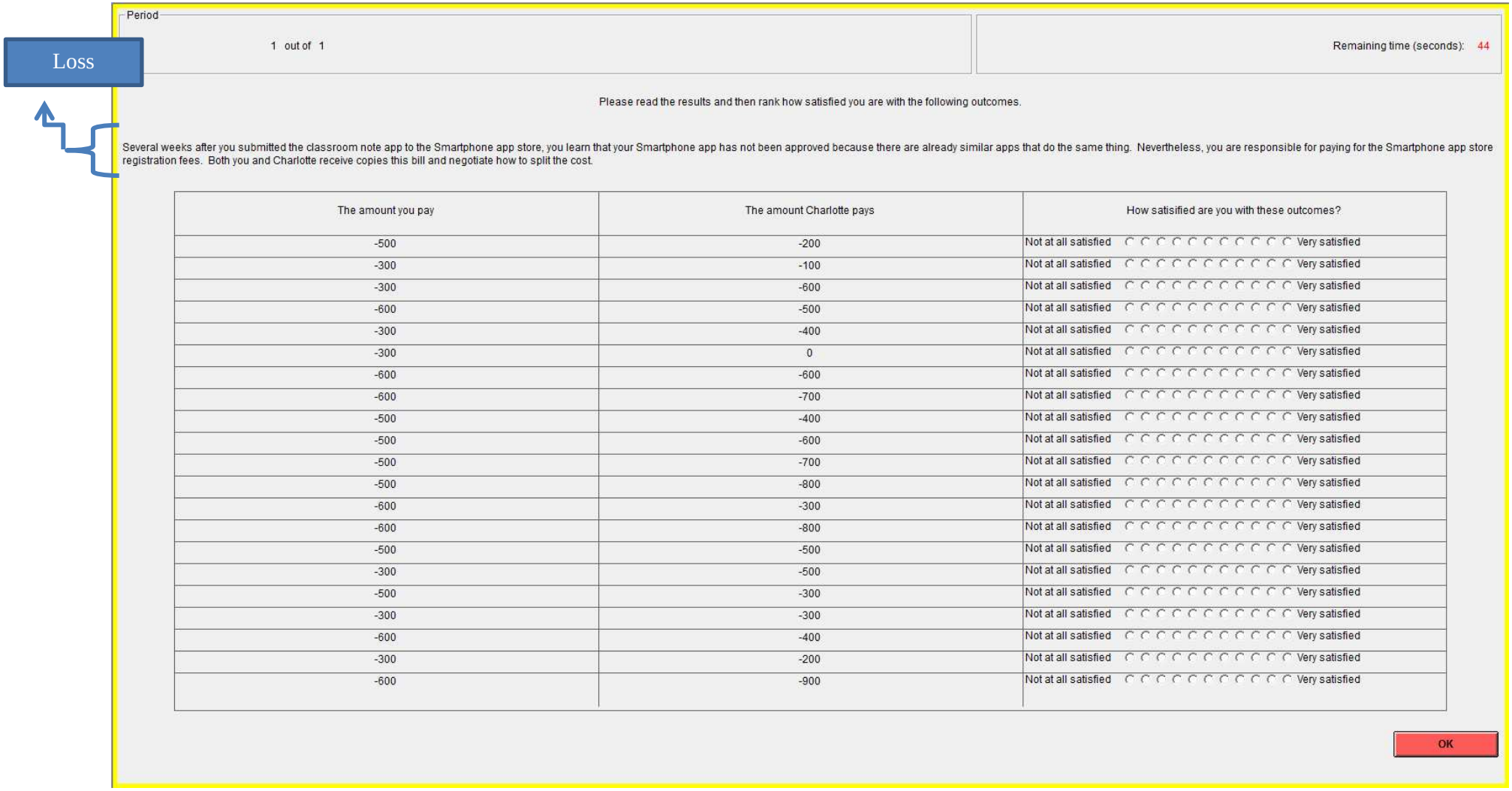
1. Smartphone App Scenario. a. Dispute and b.ii. Moderate negative relationship



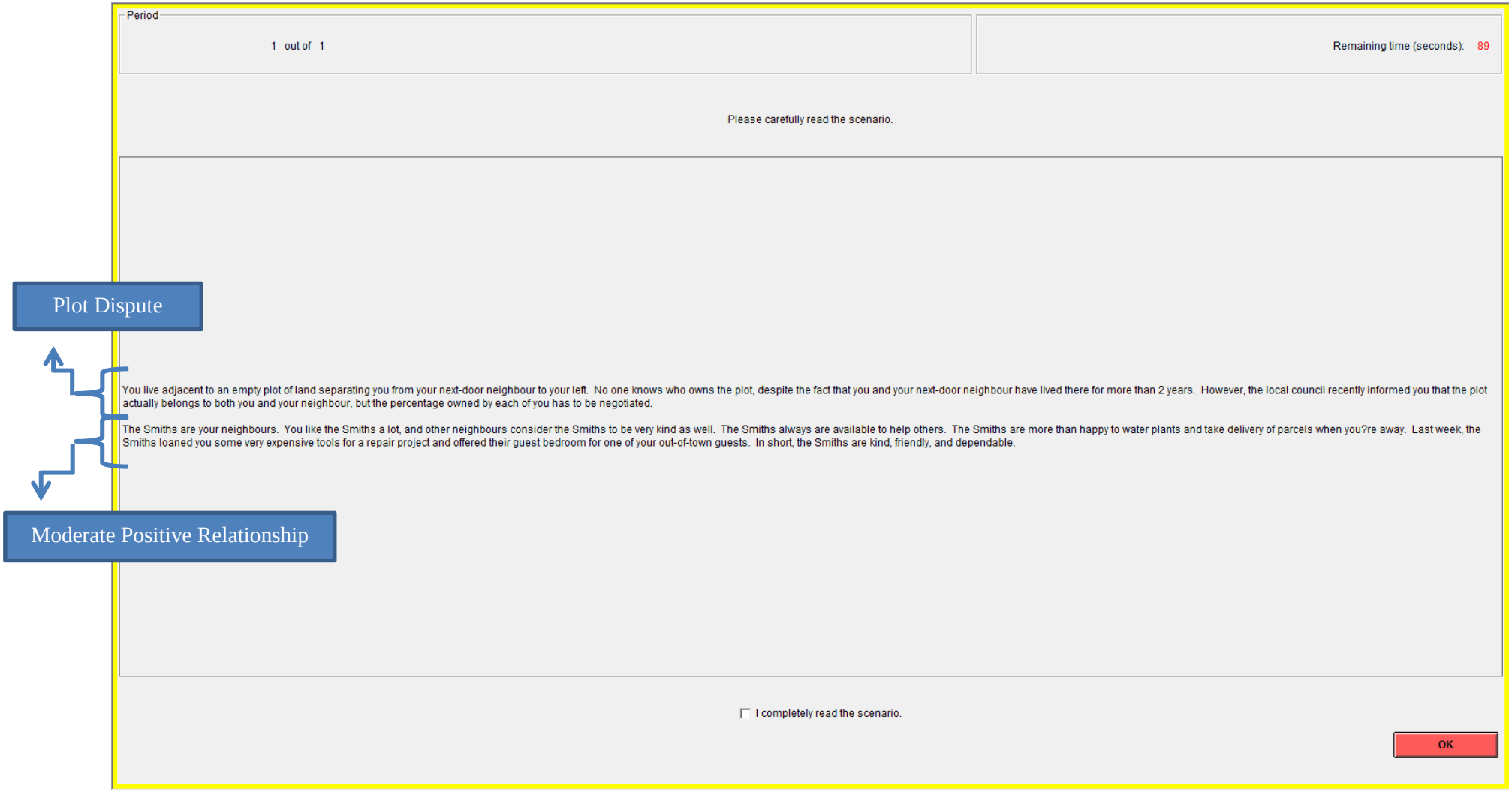
1. Smartphone App Scenario. c.i. Gain



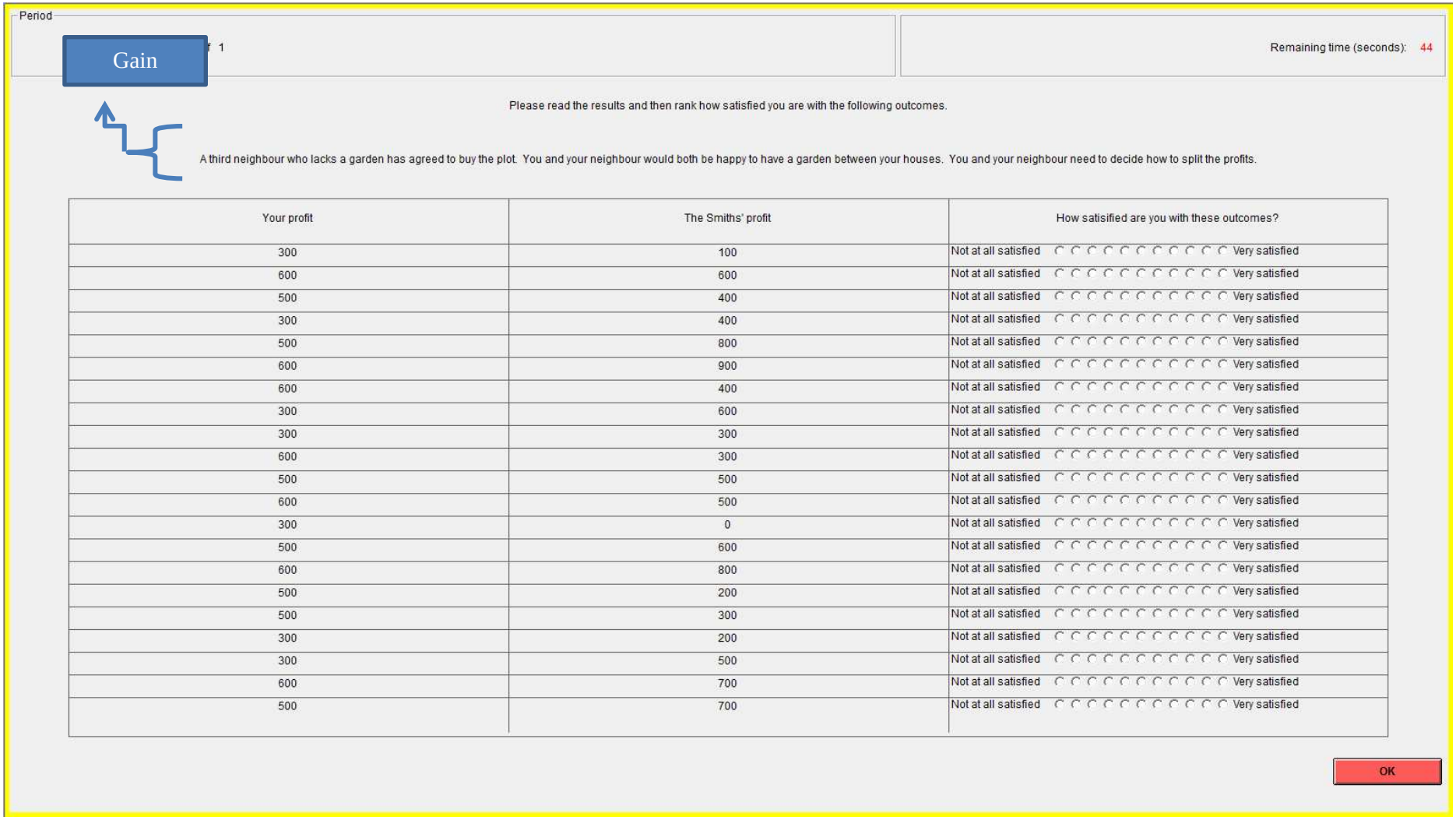
1. Smartphone App Scenario. c.i. Loss



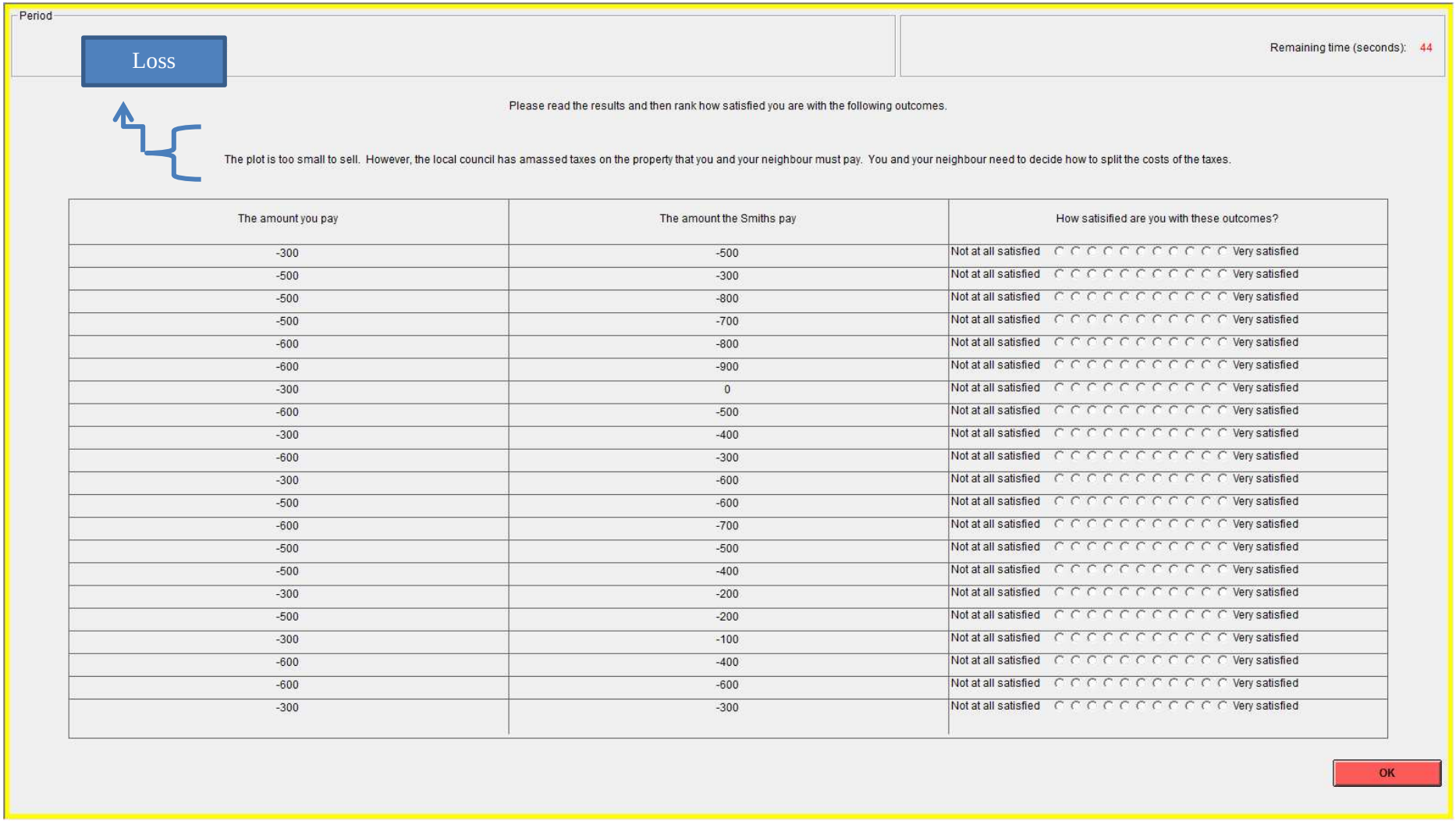
2. Plot Scenario. a. Dispute and b.i. Moderate positive relationship



2. Plot Scenario. c.i. Gain



2. Plot Scenario. c.ii. Loss



C. Estimation of a_i and b_i (the stated advantageous and disadvantageous inequality aversion parameters)

LTB89 specified their model according to three criteria: goodness of fit, simplicity, and flexibility. They examined five functional forms and the functional form that best satisfied their specifying criteria was one that included payoff for self and relative payoffs (positive and negative differences between own and other payoffs and their squared terms):

$$U_i = c_i + B_1 Self + B_2 NegDiff + B_3 NegDiff^2 + B_4 PosDiff + B_5 PosDiff^2$$

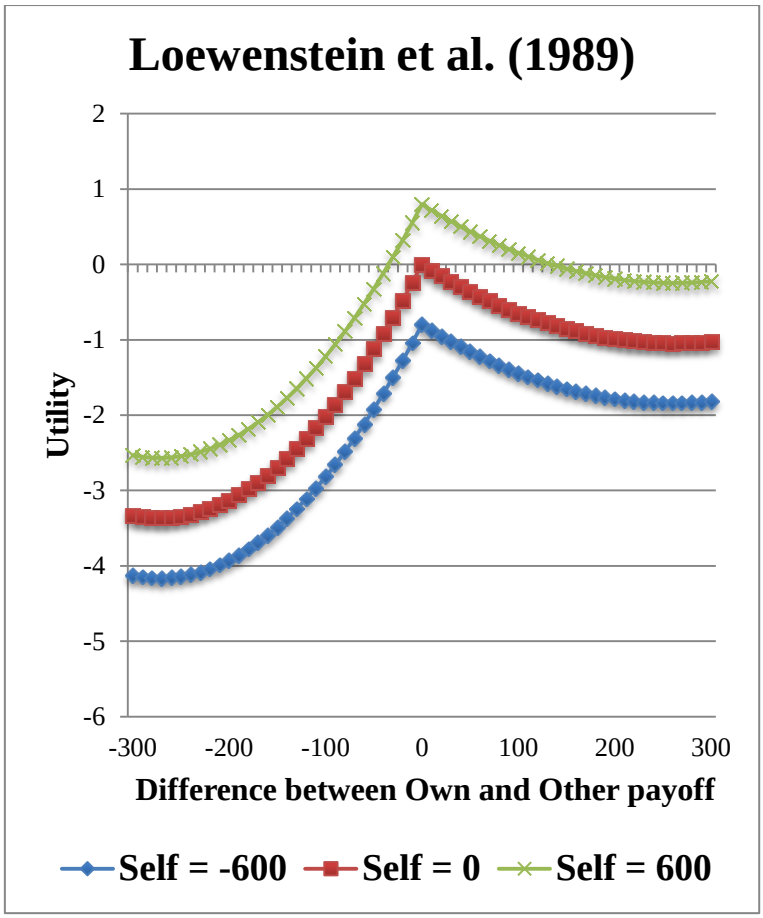
It should be noted that this functional form is quadratic which stands in contrast to the piece-wise linear nature of the Fehr-Schmidt model.

LBT89 used a subject's satisfaction ratings of each distribution as the dependent variable and then performed an OLS estimation to determine the parameters of the same model for each individual subject. Their study includes analysis both on an individual and on an aggregate level. LBT89 averaged together all of the individual level parameter estimates in order to make aggregate level figures and general statements.

For each of our subject pools, we followed LBT89's procedures and also tested each of the five functional forms suggested by LTB89. The same functional form (listed above) – as was the case in LBT89 – had the best goodness of fit as expressed by highest adjusted R^2 values for all of our subject pools.

One of the most familiar figures from the LTB89 paper is shown in Figure SM.3.2. The notable features of this figure are the tent like structure where highest utility is expressed when both payoffs are equal. Unequal payoffs lead to decreases in utility, but these decreases in utility are not equal. Utility is reduced more in the region of disadvantageous inequality than in the region of advantageous inequality. These are the features that Fehr and Schmidt (1999) cite as inspiration in the development of different aspects of their model.

Figure SM.3.2.The original quadratic LBT89 functional form expressed as a social utility curve emphasizing the importance of relative payoffs.



However, the Fehr and Schmidt parameters of inequality aversion are not directly comparable to the mean parameter estimates in the LBT89 model or the social utility curve shown here in Figure SM.3.2. The reason why direct comparison is not possible is because the functional form that LBT89 adopt is piece-wise quadratic whereas the model suggested by Fehr and Schmidt (1999) is piece-wise linear.

In order to make direct comparisons between the stated preferences as expressed in the scenarios and the revealed preferences elicited in the Blanco et al. (2011) games, we need to specify a piecewise linear functional form that emphasizes the importance of payoff differences. We can use the satisfaction ratings of each distribution as our dependent variable and then perform an OLS estimation to determine the parameters of a piece-wise linear model which is directly comparable to the Fehr and Schmidt (1999) model:

$$U_i = c_i + x_i - a_i \max\{x_j - x_i, 0\} - b_i \max\{x_i - x_j, 0\}, j \neq i$$

We estimate these parameter values for both of the scenarios each participant considers and then average the parameters together. The resulting averages are what we refer to as each individual's stated preferences of inequality aversion (a_i and b_i).

Some might find this averaging procedure questionable as both the dispute and relationship conditions vary between the two scenarios participants consider. A simple way to test this would be to look at the individual level correlation of the estimated parameters across the two scenarios each participant considered. Significantly high and positive Spearman rho values would give support to this technique. Table SM.3.2 shows the summary statistics and the correlations between the scenarios subjects considered. In all cases but one (that is, in 13 out of 14 cases), there is strong and significant positive correlation between the parameter estimates resulting from the satisfaction ratings in the two separate scenarios considered by each participant. In light of this, we follow LBT89 to average each individual's two parameter estimates to come up with our stated preferences of inequality aversion (a_i and b_i).

Figure SM.3.3 shows the estimated linear social utility curves for each subject pool. In these figures, the constants are normalized to zero. In these curves, utility is a function of the difference between Own and Other payoff at various levels of one's own payoff (-600, 0, 600). Figure 1a in the main text is constructed from these same parameter estimates in the condition when one's own payoff is zero; that is, Figure 1a is constructed from the middle series from each of the four subject pool graphs in Figure SM.3.3. Readers will notice that, aside from their linearity, the Figure SM.3.3 social utility curves appear similar to the one in Figure SM.3.2 (utility is highest for equal payoffs and disadvantageous inequality is disliked more than advantageous inequality).

Table SM.3.2. Summary statistics and correlations between Loewenstein et al. Dispute Conditions (Participants with Well Behaved Preferences)

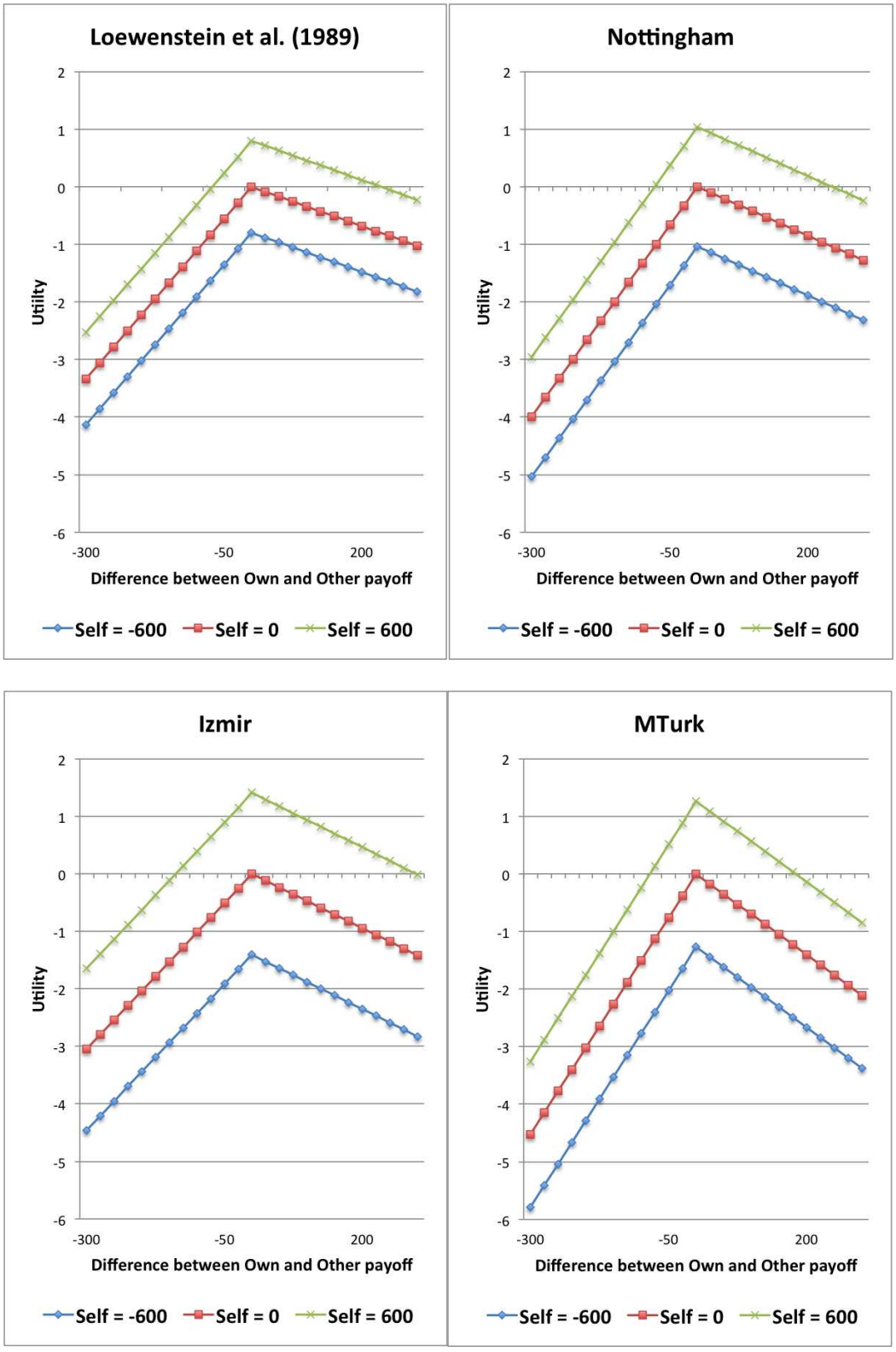
a (negdiff) Correlation between Scenarios

		Nottingham		**Izmir**		**MTurk**	
Relationship	Dispute Type	Observations	Mean	Observations	Mean	Observations	Mean
Positive	Invention	42	0.0102	38	0.0062	90	0.0113
Negative	Invention	43	0.0127	53	0.0103	93	0.0149
Positive	Plot	43	0.0148	53	0.0116	93	0.0155
Negative	Plot	42	0.0155	38	0.0126	90	0.0179
Neutral	Invention					189	0.0125
Neutral	Plot					189	0.0176
Spearman Correlation		ρ	p	ρ	p	ρ	P
Positive Invention, Negative Plot		0.6252	0.0000	0.4718	0.0028	0.4909	0.0000
Negative Invention, Positive Plot		0.6412	0.0000	0.5566	0.0000	0.4748	0.0000
Neutral Invention, Neutral Lot						0.6109	0.0000

b (posdiff) Correlation between Scenarios

		Nottingham		**Izmir**		**MTurk**	
Relationship	Dispute Type	Observations	Mean	Observations	Mean	Observations	Mean
Positive	Invention	42	0.00693	38	0.00826	90	0.00981
Negative	Invention	43	0.00050	53	0.00334	93	0.00363
Positive	Plot	43	0.00947	53	0.00964	93	0.01216
Negative	Plot	42	0.00010	38	-0.00343	90	-0.00136
Neutral	Invention					189	0.00824
Neutral	Plot					189	0.00744
Spearman Correlation		ρ	p	ρ	p	ρ	P
Positive Invention, Negative Plot		0.4105	0.0069	**0.2199**	**0.1845**	0.3069	0.0033
Negative Invention, Positive Plot		0.6283	0.0000	0.6309	0.0000	0.4337	0.0000
Neutral Invention, Neutral Plot						0.6256	0.0000

Figure SM.3.3. Linear Social Utility Curves.



LBT89 did not specify a piece-wise linear model and since their individual level data are not available we are unable to do so ourselves. We instead constructed a linear social

utility curve for LBT89 by separating the social utility curve into two components: one component in the domain of advantageous inequality and one component in the domain of disadvantageous inequality. We then fitted a line to the curve in each component. We used the slope of this fitted line as the average parameter estimate for the a parameter and the b parameter. As a check for the validity of this approach, we followed the same procedure for our Nottingham subject pool (constructing a and b values from the quadratic model that we estimated) and we report both the constructed and estimated the values for a and b in Table SM.3.3. The differences between our constructed and estimated Nottingham parameters are not big and therefore we deem this an appropriate approximation given the data limitations.

Table SM.3.3. Comparing the constructed parameters to estimated parameters

	a	b
LBT89 Constructed	0.01111	0.00341
Nottingham Constructed	0.01334	0.00558
Nottingham Estimated	0.01331	0.00426
Izmir Estimated	0.01018	0.00473
MTurk Estimated	0.01509	0.00703

To see the variation of the parameters of inequality aversion (including the variation in stated preferences) by populations, see the Kruskal-Wallis tests below in Table SM.4.1 and the Wilcoxon-Mann-Whitney tests in Table SM.4.2.

IV. Section SM.4. Supporting Analysis.

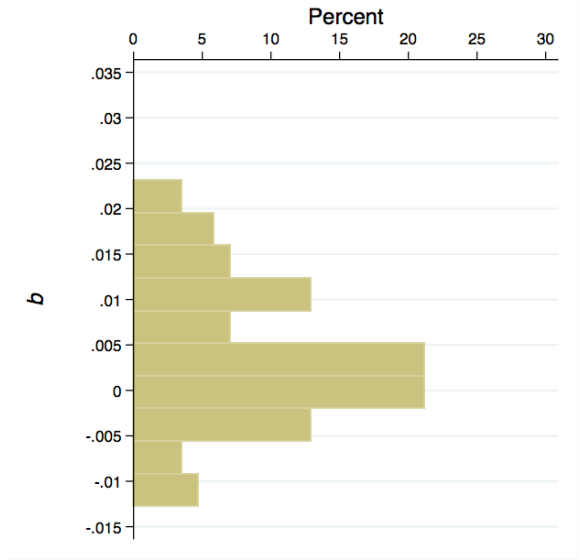
The data reported here are available as a supplementary file (BCG_Data.xlsx).

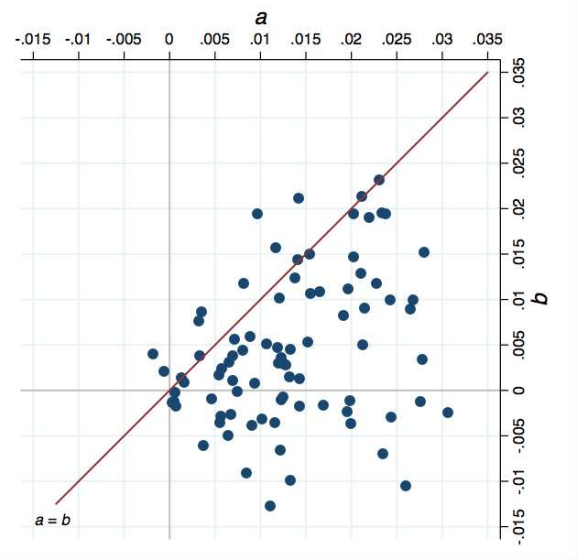
A. Supporting Analysis for Section 3.1. Stated Inequality Aversion

Figure SM.4.1.[2] Expanded Versions of Figures 1B-1D in the Main Text. Joint a and b distributions per subject pool. Each dot represents a participant's a and b parameters as calculated from their stated preferences in the updated Loewenstein, et al (1989) scenario tasks. Observations to the left of the $a = b$ line have $a < b$ which violates the Fehr and Schmidt (1999) assumption. The left and bottom panel are histograms of the b and a values, respectively.

[2] In order to document our elicited a and b values, as well as our α and β values, we show below expanded versions of Figures 1B-1D and Figure 2. In addition to the scatter plots of the figures in the main text, these expanded figures contain histograms of the distributions of the respective values. This exposition is inspired by Dannenberg et al. (2007).

Nottingham





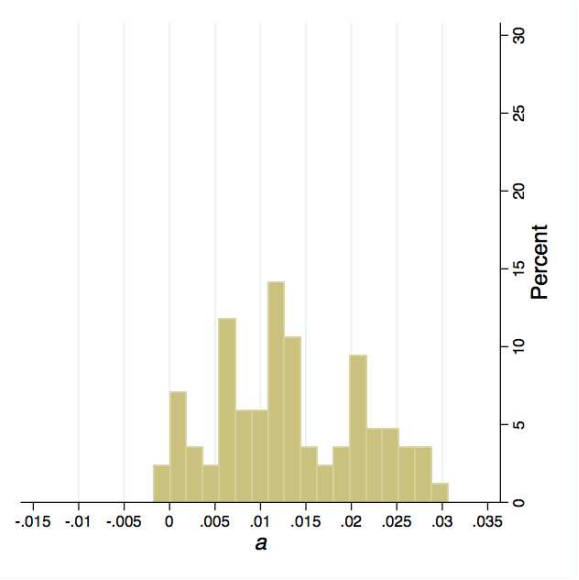
87% obey $a \geq b$

65% obey $b \geq 0$

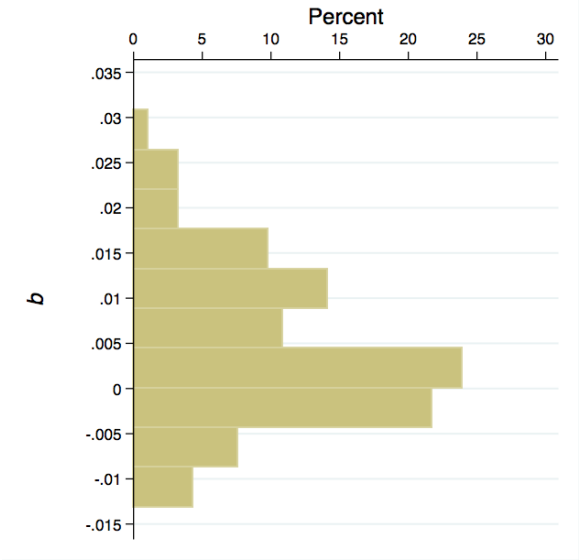
	a	*b*
Mean	0.013	0.004
Median	0.013	0.003
St. Dev	0.008	0.008

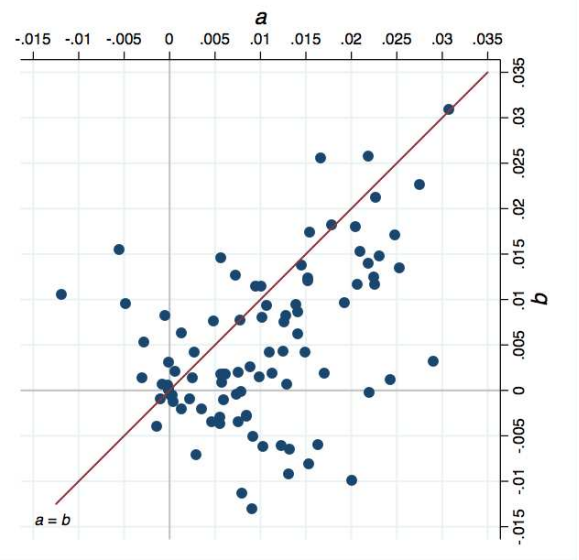
Spearman Correlation: a & b

$\rho = 0.313, p = 0.004$



Izmir





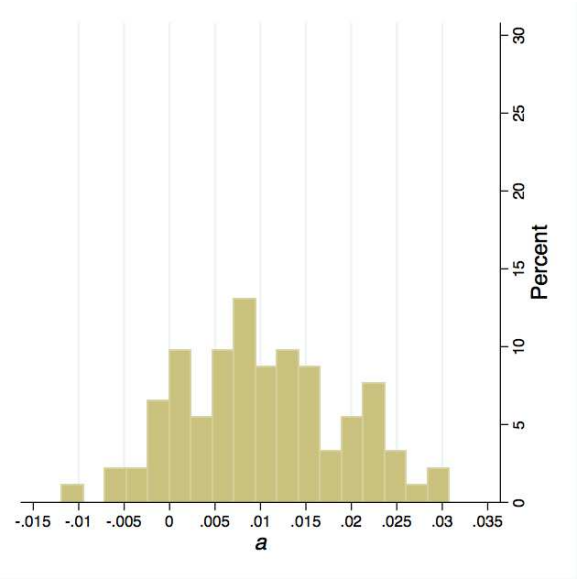
77% obey $a \geq b$

68% obey $b \geq 0$

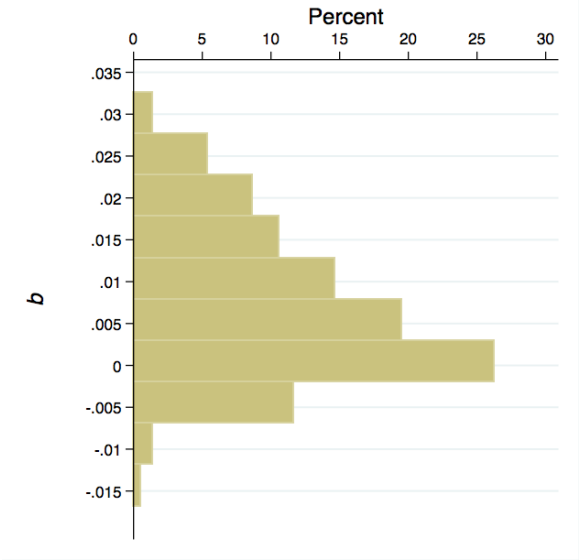
	a	*b*
Mean	0.010	0.005
Median	0.009	0.002
St. Dev	0.009	0.009

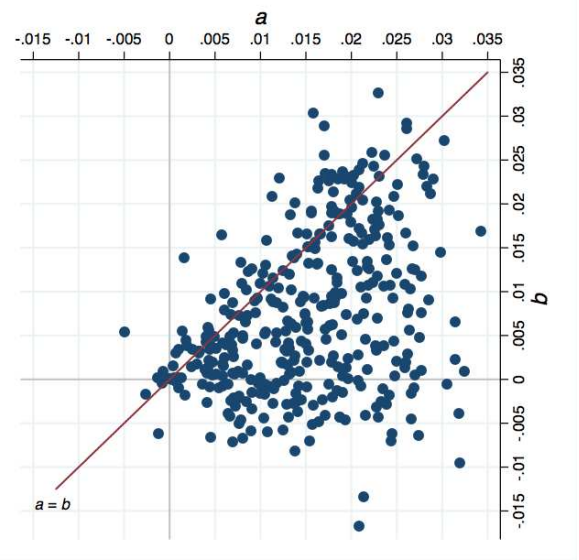
Spearman Correlation: a & b

$\rho = 0.406, p = 0.000$



MTurk





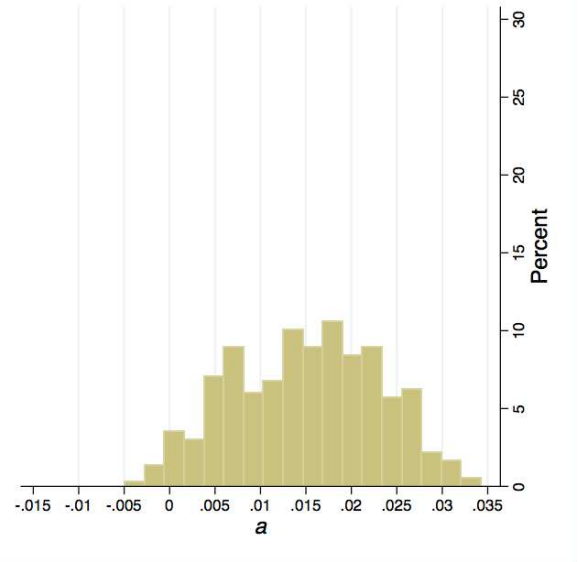
80% obey $a \geq b$

76% obey $b \geq 0$

	a	*b*
Mean	0.015	0.007
Median	0.015	0.005
St. Dev	0.008	0.009

Spearman Correlation: a & b

$\rho = 0.354, p = 0.000$



We list the results of Kruskal-Wallis tests in Table SM.4.1 indicating that for the most part the samples do not seem to originate from the same distribution beyond the categorization imposed above.

Table SM.4.1. Kruskal-Wallis Tests comparing parameters of stated inequality aversion by populations

	χ^2 (2)	p	χ^2 (2) with ties	p
a (negdiff)	23.858	0.0001	23.858	0.0001
b (posdiff)	8.779	0.0124	8.779	0.0124

In Table SM.4.2, we list the results of Wilcoxon-Mann-Whitney tests comparing the parameters of inequality aversion bilaterally between populations. We note that significant differences exists between all populations with regards to the stated a values. In regards to the stated b values, significant differences exist between MTurk and both Nottingham and Izmir populations.

Table SM.4.2. Wilcoxon-Mann-Whitney test comparing parameters of stated inequaliity aversion between populations

	Nottingham MTurk		Nottingham Izmir		Izmir MTurk	
	Z	p	Z	p	Z	p
a (negdiff)	**1.881**	**0.0600**	**-2.331**	**0.0197**	**4.762**	**0.0000**
b (posdiff)	**2.455**	**0.0141**	0.244	0.8075	**2.107**	**0.0351**

B. Supporting Analysis for Section 3.2 Revealed Inequality Aversion

The Chi^2 tests listed in Table 1 of the main text compare the distributions of each subject pool to the theoretical Fehr and Schmidt (1999) distribution and the observed Blanco et al. (2011) (BEN) distributions. Table SM.4.3 shows group Chi^2 tests with different combinations of groupings. The upper portion of the table involves comparisons to the theoretical Fehr and Schmidt (1999) distributions. The lower portion of the table involves comparisons to observed data. Various groupings are considered including combining the Blanco et al. (2011) observations with our Nottingham observations for a UK university group. For the most part group Chi^2 tests indicate that our groups are significantly different from one another. The exception is greater similarity between groups of observed β distributions as opposed to the comparison of observed β distributions to the Fehr and Schmidt (1999) theoretical distribution.

Table SM.4.3. Group Chi^2 Tests Comparing α and β Categories.

α Distributions		***β Distributions***	
Comparison to F&S Theoretical Distribution			
Groups: F&S, BEN, Nottingham, Izmir, Mturk			
$\chi 2$ (12)	53.982	$\chi 2$ (8)	27.562
p value	0.000	p value	0.001
Groups: F&S, Nottingham, Izmir, Mturk			
$\chi 2$ (9)	45.288	$\chi 2$ (6)	24.682
p value	0.000	p value	0.000
Groups: F&S, BEN & Nottingham Combined, Izmir, Mturk			
$\chi 2$ (9)	44.798	$\chi 2$ (6)	24.638
p value	0.000	p value	0.000
Comparison to Observations			
Groups: BEN, Nottingham, Izmir, Mturk			
$\chi 2$ (9)	34.484	$\chi 2$ (6)	11.566
p value	0.000	p value	0.072
Groups: Nottingham, Izmir, Mturk			
$\chi 2$ (6)	21.476	$\chi 2$ (4)	7.950
p value	0.002	p value	0.093
Groups: BEN & Nottingham Combined, Izmir, Mturk			
$\chi 2$ (6)	25.176	$\chi 2$ (4)	8.403
p value	0.000	p value	0.078

We list the results of Kruskal-Wallis tests in Table SM.4.4 indicating that there are weakly significant differences between our subject pools in β values; while, the α values do not seem to be originating from independent distributions. Note that we include the Blanco et al. (2011) data in the first instance and exclude it in the second.

Table SM.4.4. Kruskal-Wallis Tests comparing parameters of revealed inequality aversion by populations

With BEN	χ^2 (3)	p	χ^2 (3) with ties	p
α	**6.033**	**0.1100**	**6.215**	**0.1016**
β	7.359	0.0613	7.444	0.0590
Without BEN	χ^2 (2)	p	χ^2 (2) with ties	p
α	**3.308**	**0.1913**	**3.422**	**0.1807**
β	6.492	0.0389	6.575	0.0373

The results of Wilcoxon-Mann-Whitney tests comparing the parameters of inequality aversion bilaterally between populations are listed in Table SM.4.5. The α values between Nottingham and both Izmir and MTurk seem similar, but there are significant differences between Izmir and MTurk. In regards to β values, we note that Izmir is different than both Nottingham and MTurk (whereas Nottingham and MTurk are not significantly different from one another.

Table SM.4.5. Wilcoxon-Mann-Whitney test comparing parameters of revealed inequality aversion between populations

	Nottingham MTurk		Nottingham Izmir		Izmir MTurk	
	Z	p	Z	p	Z	p
α	0.591	0.5544	-1.432	0.1522	**1.739**	**0.0821**
β	0.860	0.3898	**2.411**	**0.0159**	**-2.195**	**0.0282**

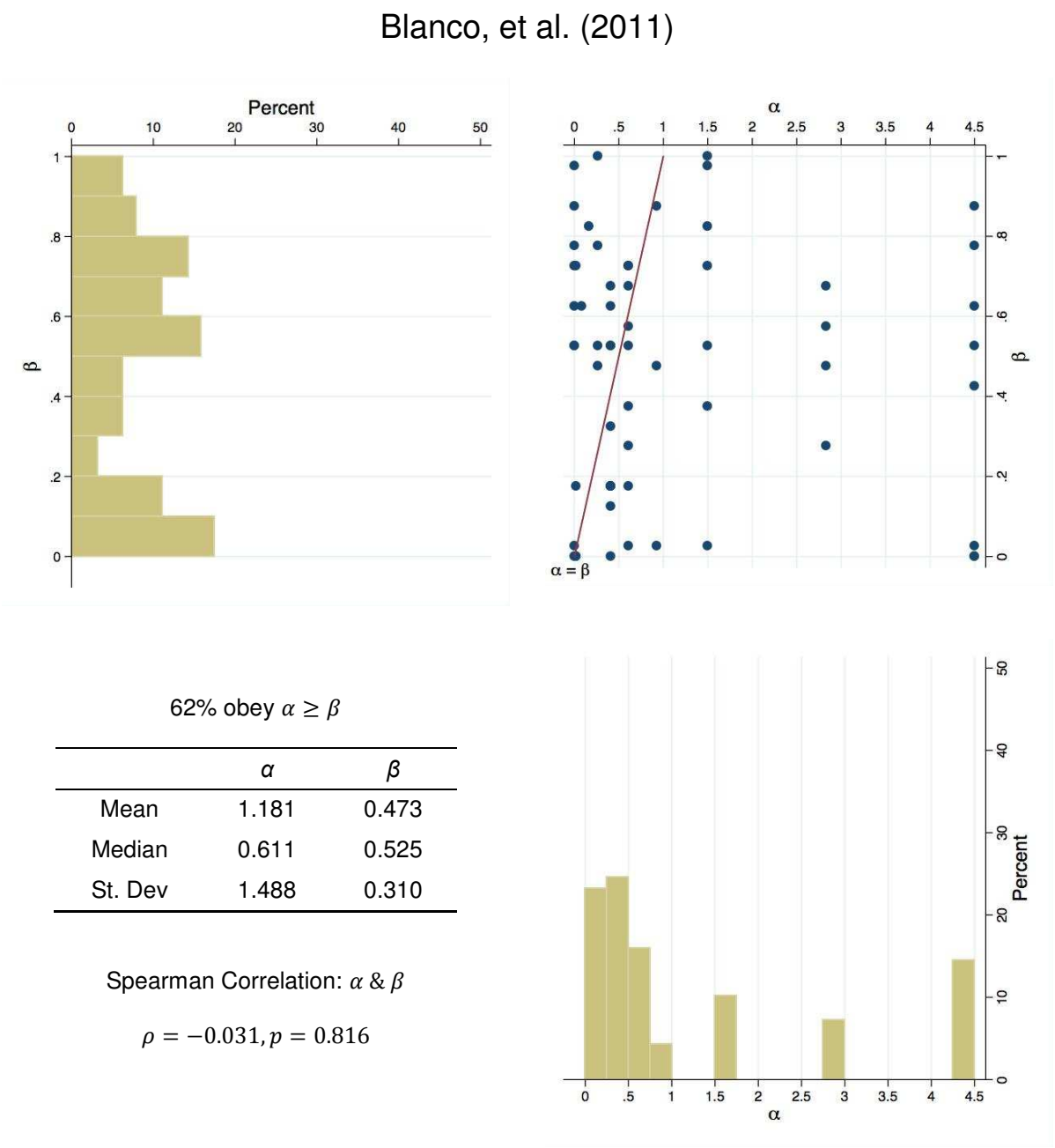
Finally, we note the significant variation in the percent of 'well-behaved' participants in each subject pool with only 45% of our Izmir sample meeting the criteria of well-behavedness as defined in the main text. Several referees requested information about the non well-behaved subjects in Izmir and here we report the proportions excluded for having multiple switch points in the various games in Table SM.4.6.

Table SM.4.6. Incidences of Multiple Switching in Izmir

		Modified Dictator Game		
		Multiple Switches	Single Switch	
Ultimatum Game	Multiple Switches	52	17	69
	Single Switch	45	92	137
		97	109	

Approximately 25% of subjects in Izmir have multiple switch points for both the MDG and the UG. About 22% have multiple switch points in just the MDG compared to just around 8% who have multiple switch point in just the UG. Again, 45% have well behaved preferences. There seem to be significantly more people who have multiple MDG switches as opposed to UG switches. This pattern is true in Izmir (elsewhere) where approximately 3x (5x) as many people have multiple switches in the MDG compared to the UG. There are also more people who report multiple switch points for both the MDG and the UG in Izmir (approximately 25% of subjects from that subject pool) versus elsewhere (approximately 2%).

Figure SM.4.2. Expanded Versions of Figure 2 in the Main Text. Joint α and β Distributions. Each dot represents a participant's α and β parameters as calculated from their revealed preferences in the Blanco, et al (2011) games. Observations to the left of the $\alpha = \beta$ line have $\alpha < \beta$ which violates the Fehr and Schmidt (1999) assumption. The left and bottom panel are histograms of the β and α values, respectively.

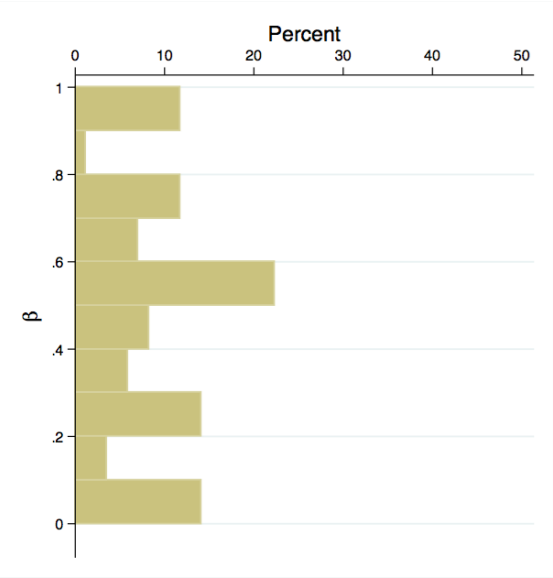


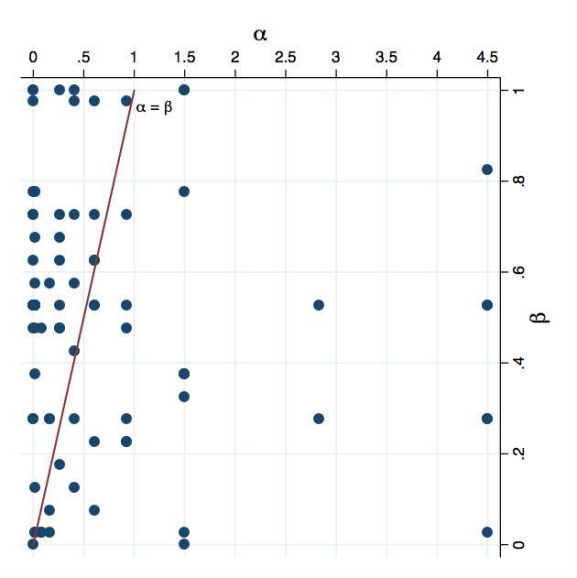
	α	β
Mean	1.181	0.473
Median	0.611	0.525
St. Dev	1.488	0.310

Spearman Correlation: α & β

$\rho = -0.031, p = 0.816$

Nottingham



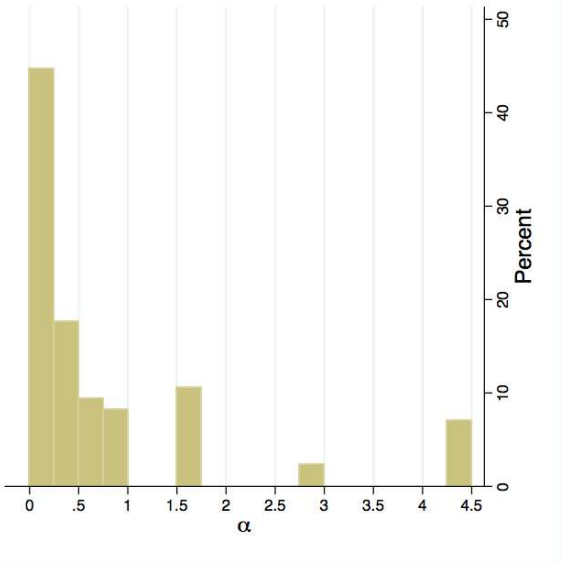


45% obey $\alpha \geq \beta$

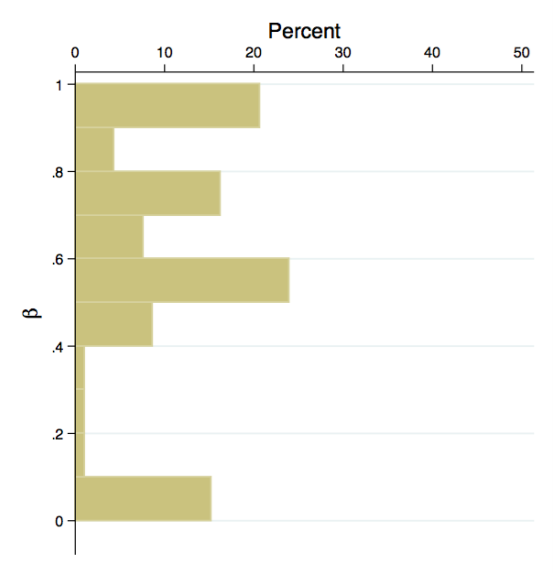
	α	β
Mean	0.754	0.484
Median	0.269	0.525
St. Dev	1.198	0.290

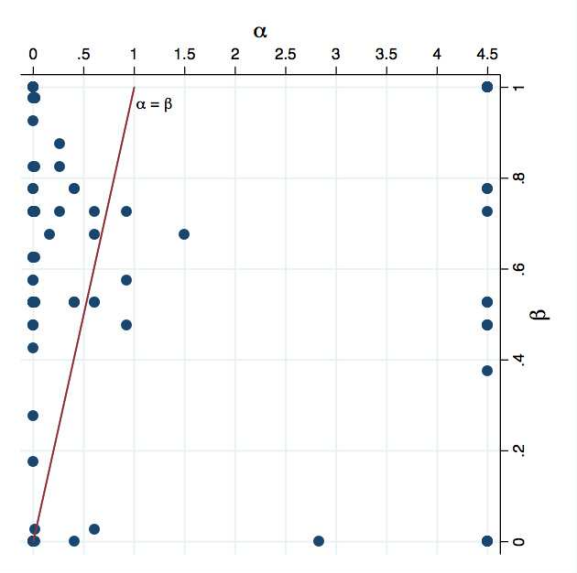
Spearman Correlation: α & β

$\rho = -0.088, p = 0.422$



Izmir



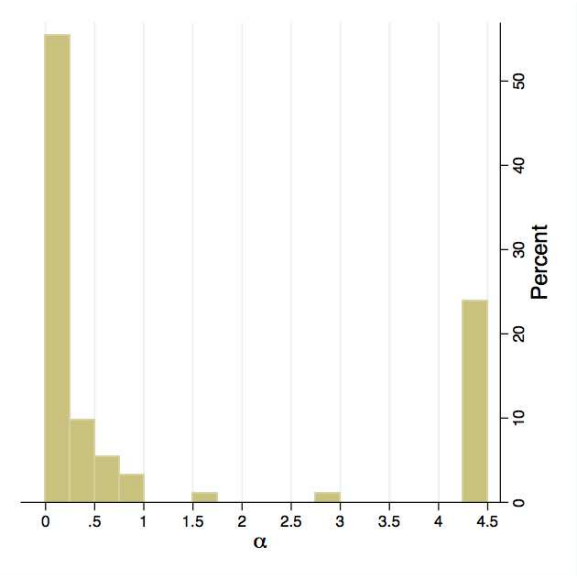


41% obey $\alpha \geq \beta$

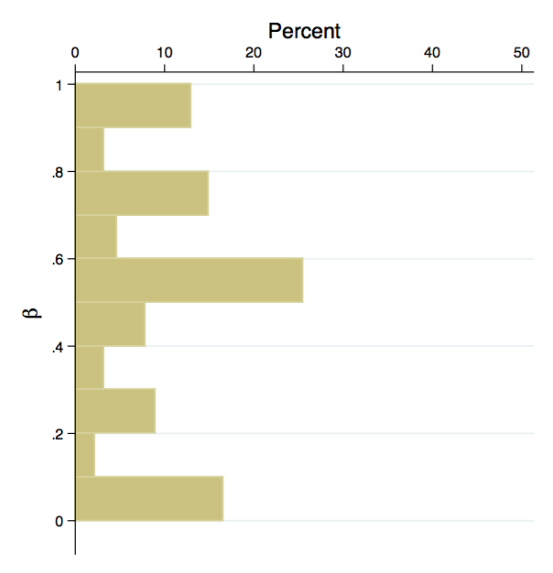
	α	β
Mean	1.227	0.589
Median	0.026	0.575
St. Dev	1.884	0.315

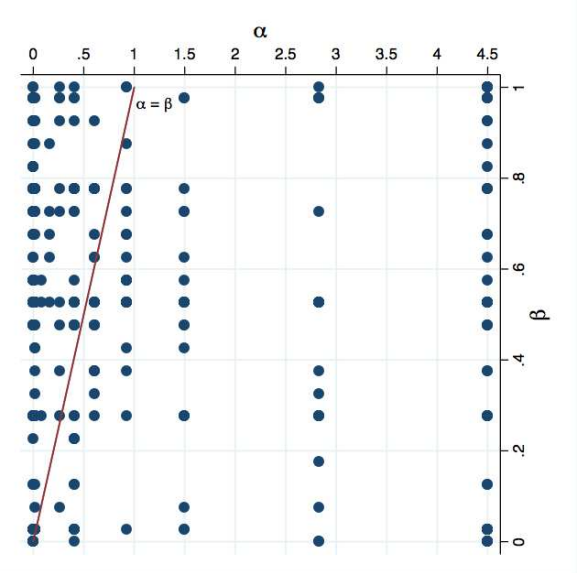
Spearman Correlation: α & β

$\rho = 0.044, p = 0.676$



MTurk





49% obey $\alpha \geq \beta$

	α	β
Mean	1.218	0.512
Median	0.410	0.525
St. Dev	1.670	0.302

Spearman Correlation: α & β

$\rho = -0.107, p = 0.040$