

Stated and revealed inequality aversion in three subject pools

Benjamin Beranek¹ · Robin Cubitt¹ ·
Simon Gächter^{1,2,3} 

Received: 31 October 2014 / Accepted: 8 April 2015 / Published online: 25 April 2015
© Economic Science Association 2015

Abstract This paper reports data from three subject pools ($n = 717$ subjects) using techniques based on those of Loewenstein et al. (J Personal Soc Psychol 57:426–441, 1989) and Blanco et al. (Games Econ Behav 72:321–338, 2011) to obtain parameters, respectively, of stated and revealed inequality aversion. We provide a replication opportunity for those papers, with two innovations: (1) a design which allows stated and revealed preferences to be compared at the individual level; (2) 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 C90

Electronic supplementary material The online version of this article (doi:[10.1007/s40881-015-0007-1](https://doi.org/10.1007/s40881-015-0007-1)) contains supplementary material, which is available to authorized users.

✉ Simon Gächter
simon.gaechter@nottingham.ac.uk

¹ Centre for Decision Research and Experimental Economics, School of Economics, University of Nottingham, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, UK

² CESifo, Munich, Germany

³ IZA, Bonn, Germany

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, 823–824) 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

¹ We are aware of three further studies, by Dannenberg et al. (2007, 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.

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) 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 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.

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

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

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).

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 Eq. (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 Sect. 3.

We turn now to the elicitation of *stated* preferences. In Studies 1 and 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 $2 \times 2 \times 2$, 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 Sect. 1. 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 into 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 Fig. 1a by averaging the a_i and b_i parameters across each subject pool (and normalizing the constant to zero). Figure 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.

Figure 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 Sect. 3.4.

Figure 1b–d 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; Iriberry and Rey-Biel 2013).

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

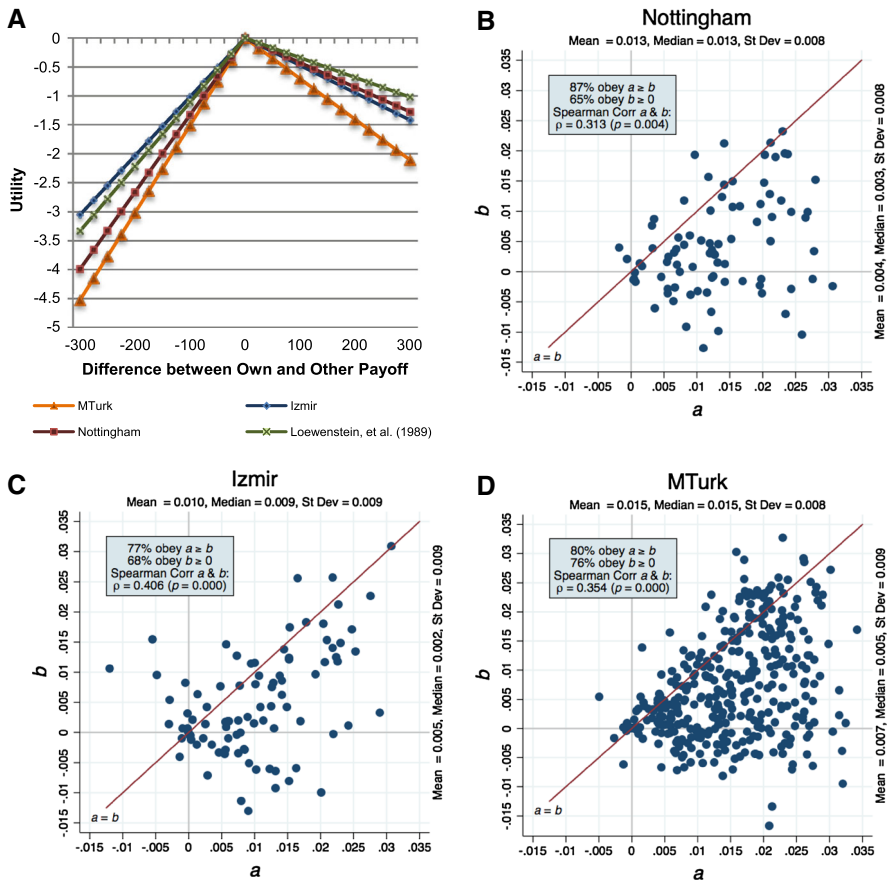


Fig. 1 Stated preferences on aggregate and individual levels. **a** Utility (satisfaction ratings) as a function of difference between own and other payoff in the scenario tasks. **b-d** Joint a and b distributions per subject pool. Each dot represents a participant's a and b parameters as calculated 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 $a < b$

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 squared-tests in each case, as well as the mean, median and standard deviation of each parameter in each subject pool.

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

Table 1 Distribution of α and β values

	F&S	BEN	Nottingham	Izmir	MTurk
α					
$\alpha < 0.4$	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 squared-test to F&S	χ^2 (3)	1.790	11.226	37.751	17.211
	<i>p</i> value	0.618	0.011	0.000	0.001
Chi squared-test to BEN	χ^2 (3)		9.014	24.933	9.699
	<i>p</i> value		0.029	0.000	0.021
β					
$\beta < 0.235$	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 squared-test to F&S	χ^2 (2)	8.51	3.816	21.517	14.491
	<i>p</i> value	0.014	0.148	0.000	0.001
Chi squared-test to BEN	χ^2 (2)		2.729	5.033	3.109
	<i>p</i> value		0.256	0.081	0.211
Total sample size		72	104	206	407
Percent well behaved (%)		85	82	45	91

BEN refers to the Blanco et al. (2011) observed distribution and F&S to Fehr and 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

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 squared-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 squared-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 squared-tests, reveals significantly different distributions in the Izmir and

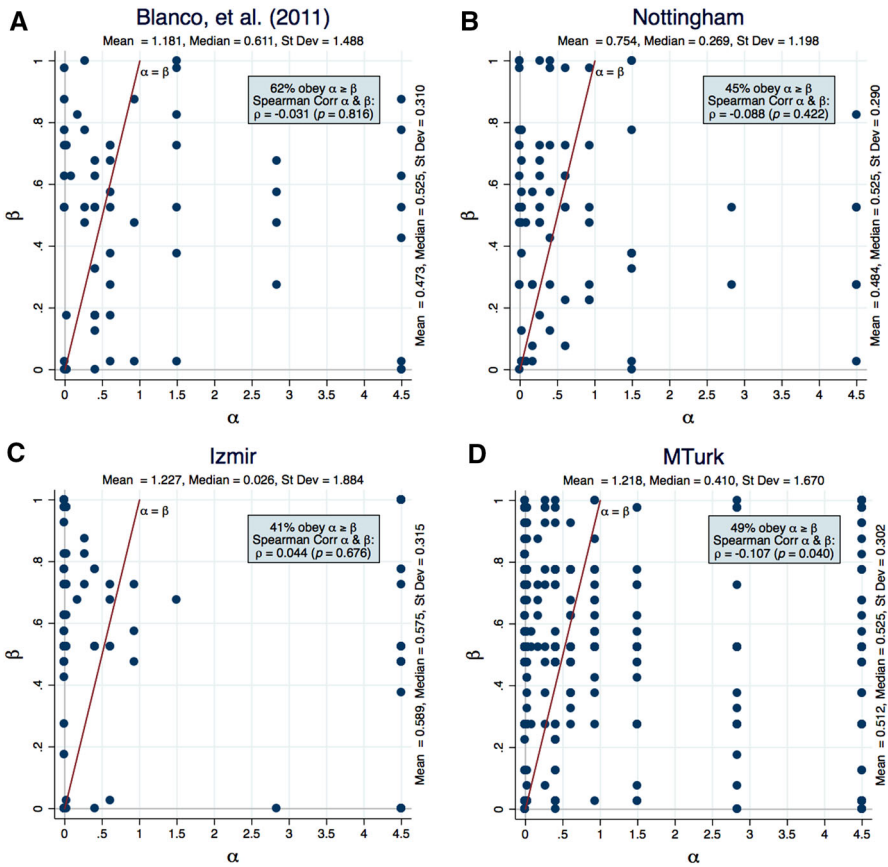


Fig. 2 Joint α and β distributions. Each dot represents a participant’s α and β parameters as calculated from their revealed preferences in the Blanco et al. (2011) games. Observations to the left of the $\alpha = \beta$ line have $\alpha < \beta$ which violates the Fehr and Schmidt (1999) assumption

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

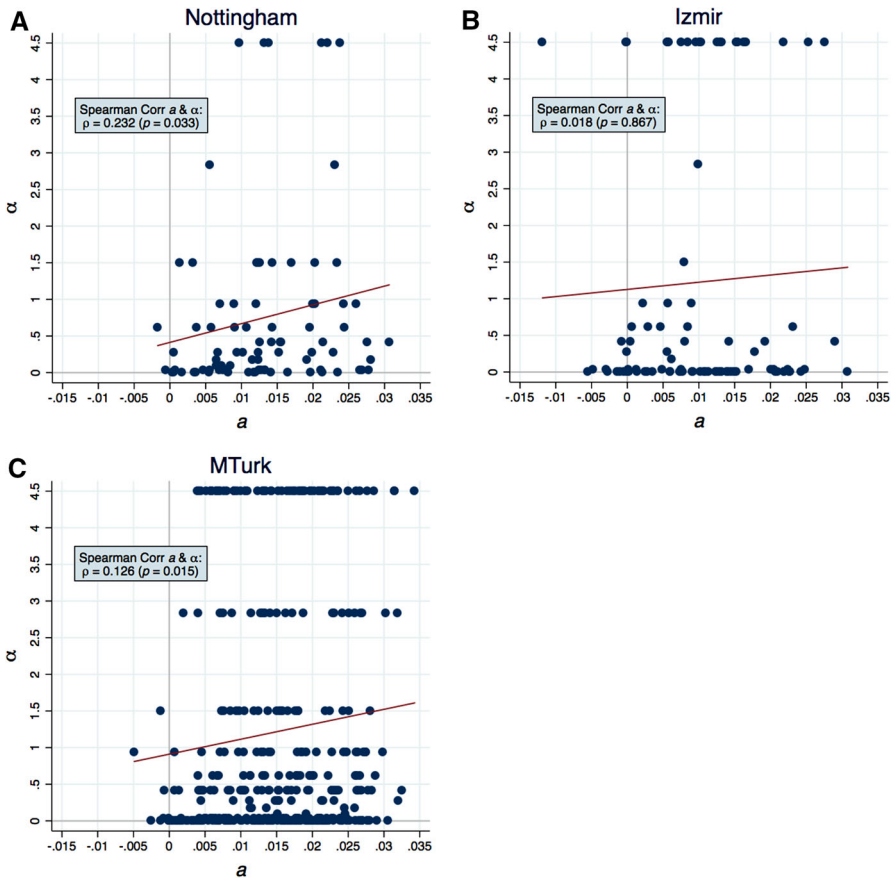


Fig. 3 Joint a and α distributions. Each dot represents a participant's a as calculated from their stated preferences and α parameters as calculated from their revealed preferences. The line results from the linear regression of α on a

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 ($\rho = -0.089$; $p = 0.038$).

3.3 Relationship between stated and revealed preferences

Figure 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,

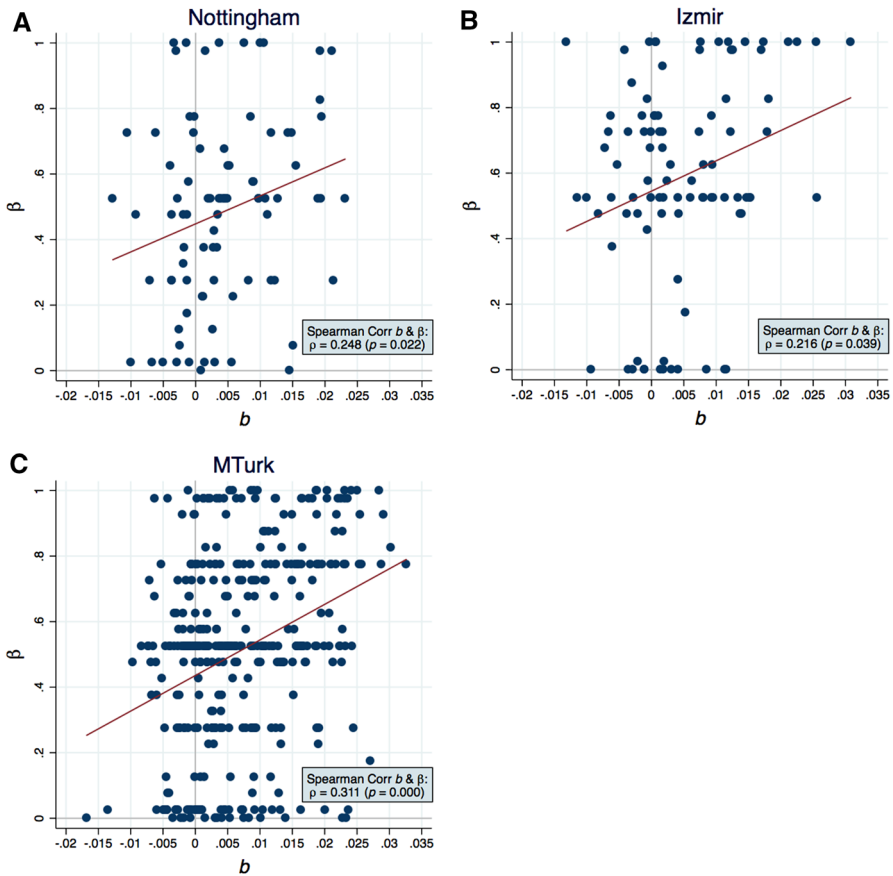


Fig. 4 Joint b and β distributions. Each dot represents a participant’s b as calculated from their stated preferences and β parameters as calculated from their revealed preferences. The line results from the linear regression of β on b

especially in the MTurk sample. In the pooled data, the correlation is slightly positive ($\rho = 0.132$; $p = 0.002$).

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 ($\rho = 0.2785$; $p < 0.001$).

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 ($a_i, b_i, \alpha_i, \beta_i$) on three standard socio-demographic

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

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

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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.

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 ρ 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 them more recognizable or accessible than the stripped-down lab games. A third perspective is that the scenario

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) and Peysakhovich et al. (2014)].

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.

Acknowledgments 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).

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 and 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, 166–193.
- 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, 817–868.
- 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 K. Kremer & V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen GWDG Bericht 63*. 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, 367–388.
- 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.
- Iriberrí, 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, 4939. doi:10.1038/ncomms5939.
- Selten, R. (1978). The equity principle in economic behavior. In H. Gottinger & W. Leinfellner (Eds.), *Decision theory and social ethics*. Dordrecht: Reidel.
- Teysier, 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., et al. (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., et al. (2013). Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes*, 120, 260–271.