CESifo WORKING PAPERS

9851
2022

July 2022

Meta-Analysis of Inequality Aversion Estimates

Salvatore Nunnari, Massimiliano Pozzi

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9851*

Meta-Analysis of Inequality Aversion Estimates

Abstract

We conduct an interdisciplinary meta-analysis to aggregate the knowledge from empirical estimates of inequality aversion reported from 1999 to 2022. In particular, we examine 85 estimates of disadvantageous inequality aversion (or envy) and advantageous inequality aversion (or guilt) from 26 articles in economics, psychology, neuroscience and computer science that structurally estimate the Fehr and Schmidt (1999) model of social preferences. Our meta-analysis supports the presence of inequality concerns: the mean envy coefficient is 0:426 with a 95% probability that the true value lies in the interval [0:240; 0:620]; the mean guilt coefficient is 0:290 with a 95% probability that the true value lies in the interval [0:212; 0:366]. Moreover, we observe high levels of heterogeneity, both across studies and across individuals, with estimated parameters sensitive to the experimental task and the subject population.

JEL-Codes: C900, C110, D630, D910.

Keywords: social preferences, inequality aversion, inequity aversion, envy, guilt, meta-analysis, multi-level random-effects model, Bayesian hierarchical model.

Salvatore Nunnari
Bocconi University
Department of Economics
Milan / Italy
salvatore.nunnari@unibocconi.it

Massimiliano Pozzi
Bocconi University
BELSS
Milan / Italy
pozzi.massimiliano@unibocconi.it

July 14, 2022
Nunnari acknowledges financial support from the European Research Council (Grant No. 852526).

1 Introduction

The standard economic model of choice assumes that individuals are only motivated by self-interest. In the last three decades, however, a large body of evidence from the experimental social sciences has showed that most people hold *other-regarding preferences*, that is, that they care about others' outcomes or whether others are treated fairly or not.

Models of decision-making augmented with other-regarding preferences have been successfully used to explain behavior which is commonly observed in laboratory experiments yet puzzling from the perspective of the standard economic model of choice. This includes responders' rejection of positive offers in ultimatum games (Güth, Schmittberger and Schwarze, 1982; Eckel and Grossman, 2001), proposers' positive offers in dictator games (Forsythe et al., 1994; Hoffman et al., 1994; Henrich et al., 2005), cooperation in the static prisoner's dilemma (Yamagishi and Kiyonari, 2000), positive contributions in the linear public good game (Ledyard, 1995), and positive amounts sent and returned in trust games (Berg, Dickhaut and McCabe, 1995; Burks, Carpenter and Verhoogen, 2003). Moreover, models of other-regarding preferences have been used to explain or predict behavior outside of the laboratory, with applications ranging from optimal climate policy (Azar and Sterner, 1996; Anthoff et al., 2009; Tol, 2010), industrial organization (Huck et al., 2001) and trade protection (Lü et al., 2012) to contract design (Fehr and Schmidt, 2004; Fehr et al., 2007, 2008) and redistributive policies (Epper et al., 2020).

The most cited and influential model of other-regarding preferences is the model of *inequity aversion* proposed by Fehr and Schmidt (1999) (FS henceforth).[1] In the simplest two-players version of this model, the utility agent i derives from outcome x is

$$U_i(x) = x_i - \alpha_i \max[x_j - x_i, 0] - \beta_i \max[x_i - x_j, 0],\ j \neq i.$$

The agent's utility does not depend only on her own payoff, x_i, but also on the compari-

[1] As of 26 April 2022, FS has 13,895 citations on Google Scholar and 4,889 citations on Web of Science.

son with the other agent's payoff, x_j. Assuming that $\alpha \geq \beta \geq 0$ (as in FS), this can be interpreted as a model of inequity aversion, since differences in payoffs cause disutility for agent i. Moreover disadvantageous inequality is assumed to be more painful than advantageous inequality. This parsimonious utility specification is able to explain many of the above mentioned "anomalies" while keeping the model simple and tractable at the same time.

Despite all the work social scientists have done in the past 20 years to give the model an axiomatic foundation and to test it in the laboratory, there is still no consensus on what are plausible values of α and β or on what is the distribution of these two preference parameters in relevant populations. In their original paper, FS calibrate a distribution of parameters to match the behavior observed in previous ultimatum game experiments (e.g., Roth and Erev 1995). This distribution assumes that α can take four different values in the population — 0, 0.5, 1 and 4 — with calibrated shares of, respectively, 30%, 30%, 30% and 10%; on the other hand, β was assumed to take three different values — 0, 0.25 and 0.6 — with calibrated shares of, respectively, 30%, 30% and 40%. More recently, Blanco, Engelmann and Normann (2011) estimated the coefficients at the individual level using ultimatum and dictator games and reported average estimates of 1.18 for α and 0.47 for β. The distributions in FS and in Blanco, Engelmann and Normann (2011) have been used as benchmark in theoretical work with inequity averse agents to deliver counterfactuals and policy recommendations (see, e.g., Fehr and Schmidt 2004, Fehr, Klein and Schmidt 2007, Fehr, Kremhelmer and Schmidt 2008, Normann and Rau 2015, and Vogt 2016).

In this paper, we aggregate the knowledge from empirical estimates of inequality aversion accumulated in over 20 years of research with the method of meta-analysis, that is, "the statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings" (Glass, 1976). In a meta analysis, studies are selected using a precise inclusion criteria; then, the information contained in these studies is codified and summarized to explain both regularities and variation across studies.[2]

[2]Thus, meta-analysis differs from narrative reviews that give, instead, a descriptive overview of a research topic, presenting the historical trajectory and the key findings in the literature. While providing a useful

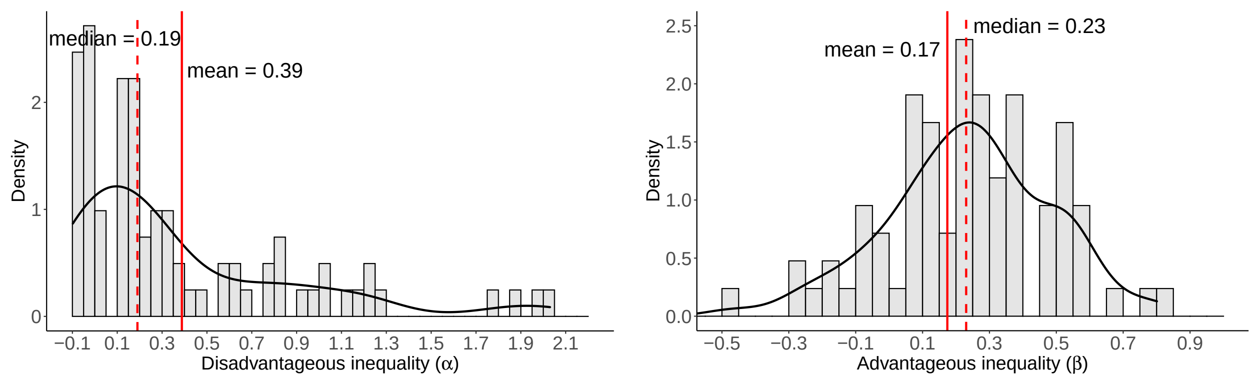


Figure 1: Distribution of Disadvantageous (α) and Advantageous Inequality (β) Coefficients. Notes: Bins for histograms are 0.05 wide; the Gaussian kernel density (solid black line) uses the Silverman's rule of thumb for bandwidth selection; in the panel for β, the horizontal axis is truncated at -0.5 for better visual rendering but the kernel density uses all available estimates.

In particular, we collect 85 estimates of disadvantageous inequality aversion (or envy) and advantageous inequality aversion (or guilt) from 26 articles in economics, psychology, neuroscience and computer science that structurally estimate the FS model of social preferences and we tackle three research questions. First, given the accumulated knowledge, what is the best estimate of α and β? Second, how do α and β vary depending on the characteristics of a study (e.g., the subject population, the experimental task, and the estimation methodology)? Third, is there evidence of selective reporting or publication bias?

In order to answer the first question, we initially conduct a non-parametric analysis. Figure 1 shows the distribution of estimates in our dataset. The raw mean and median estimates of α are, respectively, 0.39 and 0.19 with around a quarter of estimates (21 out of 81) which are equal to or less than 0 (in contrast with the assumption in FS). The raw mean and median estimtes of β are, respectively, 0.17 and 0.23, with a bell-shaped distribution and, again, many negative observations ($\beta < 0$ in 16 out of 85 estimates). Focusing on studies which estimate both parameters, disadvantageous inequality matters more than advantageous inequality only half of the time (44 out of 81 estimates) and the correlation between the two parameters is indistinguishable from 0. In the non-parametric analysis,

summary of past research and suggesting future avenues, narrative reviews do not systematically analyze all studies asking the same research question in order to test a statistical hypothesis like meta-analyses do.

all estimates are given equal weight (even if the parameters computed in some studies are more reliable than others) and assumed to be independent from one another (even if the same study provides multiple estimates). To tackle these issues, we compute a "weighted average" for α and β using a multi-level random-effects model and a Bayesian hierarchical model. The two approaches give nearly identical results and suggest that inequality aversion is a strong driver of human behavior: the meta-synthetic average for the disadvantageous inequality coefficient is 0.43 while the meta-synthetic average for the advantageous inequality coefficient is 0.29 (and both are strongly statistically significant).

While we use weighted averages to summarize the information in our dataset, we observe high level of heterogeneity in estimates, both across studies and across individuals in a single study. To explain this heterogeneity, we use the features of the studies and of the estimates we coded in our dataset as mediating variables. These meta-regressions reveal interesting patterns: estimates of α computed using choices from strategic environments are larger than estimates computed using choices from individual decision-making tasks, while the reverse is true for estimates of β; adults are less concerned about disadvantageous inequality than college students; and experimental subjects from Southern Europe (France, Italy, Spain, and Turkey) are more averse to advantageous inequality than subjects from the US and Northern Europe (UK, Germany, Netherlands, Sweden, and Switzerland).

Finally, one aspect to keep in mind when conducting a meta-analysis is the problem of selective reporting and publication bias which arise when the probability of a study being published is affected by its results. In order to detect selective reporting, we use funnel plots and apply the Funnel Asymmetry Testing and Precision Effect Testing (FAT-PET) procedure (Stanley and Doucouliagos, 2012, 2017). On one hand, funnel plots highlight the absence of studies estimating (large in magnitude and imprecisely estimated) negative values of α and positive values of β. On the other hand, the FAT-PET procedure suggests that the asymmetry in the funnel plots could be generated in the absence of publication bias — for example, because of feasibility constraints in the estimation of the parameters due to the experimental tasks employed or because of the implausible preferences implied by the missing values of α and β.

While meta-analysis is not as common in economics as in other disciplines (e.g., medicine and public policy), its popularity has increased in the last decade, especially after concerns have been raised regarding the replicability of results in the social sciences.[3] Examples of meta-analyses in experimental and behavioral economics are Zelmer (2003) on linear public good games, Embrey, Fréchette and Yuksel (2018) on the finitely repeated prisoner's dilemma, Baranski and Morton (2021) on multilateral alternating-offer bargaining, Imai, Rutter and Camerer (2018, 2021) on time preferences, and Brown et al. (2021) on loss aversion. To the best of our knowledge, this is the first work that uses meta-analysis techniques to summarize empirical estimates of outcome-based inequity aversion. Our work builds on the narrative reviews on other-regarding preferences by Fehr and Schmidt (2006) and Cooper and Kagel (2016), the meta-analysis on dictator games by Engel (2011) and the meta-analysis on ultimatum games by Oosterbeek, Sloof and Van De Kuilen (2004) and Cooper and Dutcher (2011). These meta-analyses summarize the behavior observed in laboratory experiments testing ultimatum and dictator games and investigate the explanatory power of mediating variables (e.g., the size of the pie and the location of the experiment) but do not discuss structural estimates of a model.

The rest of this paper is organized as follows. Section 2 describes the model of inequity aversion proposed by FS and its variations structurally estimated in the literature. Section 3 describes how the data was assembled and coded. Section 4 presents the results and Section 5 concludes.

2 The FS Model of Other-Regarding Preferences

In this section, we describe the original model in FS and the variations whose parameters are structurally estimated by the studies in our dataset. Consider a set of N players indexed

[3] See Dreber and Johannesson (2019) and Camerer et al. (2016).

by i and a vector of outcomes (e.g., monetary payoffs), $x = (x_1, x_2, ..., x_N)$. FS assume that player i derives the following utility from x:

$$U_i(x) = x_i - \alpha_i \frac{1}{N-1} \sum_{j \neq i} \max[x_j - x_i, 0] - \beta_i \frac{1}{N-1} \sum_{j \neq i} \max[x_i - x_j, 0], \tag{1}$$

where $\alpha_i \geq \beta_i$ and $1 > \beta_i \geq 0$. With only two players, this simplifies to

$$U_i(x) = x_i - \alpha_i \max[x_j - x_i, 0] - \beta_i \max[x_i - x_j, 0],\ i \neq j. \tag{2}$$

The first term in equations (1) and (2) captures the utility from one's own outcome; the second term measures the disutility from being behind in pairwise comparisons (i.e., aversion to disadvantageous inequality); and the third term measures the disutility from being ahead in pairwise comparisons (i.e., aversion to advantageous inequality).

The assumptions in FS are worth a brief discussion. The non-negativity constraints, $\alpha \geq 0$ and $\beta \geq 0$, imply that this is a model of *inequality aversion*: fixing her own payoff, x_i, player i's utility is maximized when $x_j = x_i$ (see Figure 2). The assumption $\alpha \geq \beta$ implies that disadvantageous inequality hurts more than advantageous inequality and it is inspired by earlier work in behavioral and experimental economics (Kahneman and Tversky, 1979; Loewenstein, Thompson and Bazerman, 1989). Finally, constraining β to be smaller than 1 is meant to avoid an implausible scenario: agents with $\beta > 1$ are willing to burn money in order to reduce the favorable gap between their allocation and the allocation to others.

While this can be interpreted as a model of inequality aversion when $\alpha > 0$ and $\beta > 0$, the framework can be used to model different kinds of other-regarding preferences: if $\alpha < 0$ and $\beta < 0$, this is a model of *inequality seeking*; if $\alpha < 0$ and $\beta = 0$, this is a model of *altruistic preferences*; if $\alpha > 0$ and $\beta < 0$, this is a model of *spiteful preferences*; and if $\alpha < 0$ and $\beta > 0$, this is a model of *efficiency concerns*. Our meta-analysis will reveal which type of other-regarding preferences is more common in the populations that have been sampled

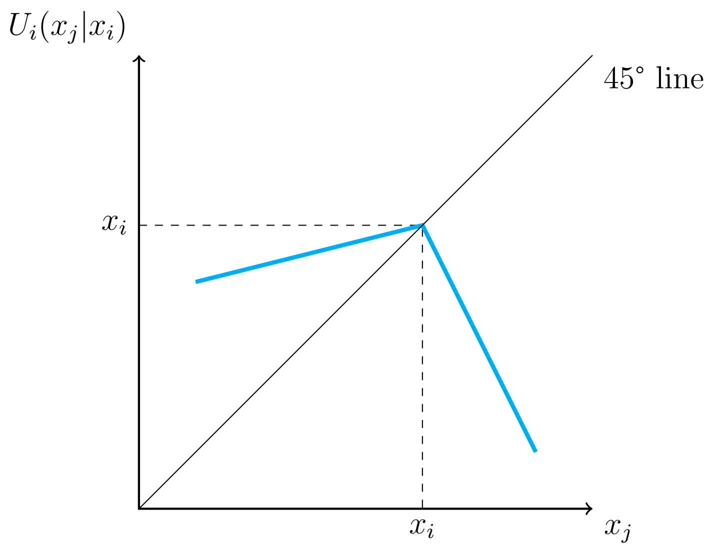


Figure 2: Utility of Inequality Averse Player i in Game with 2 Players ($\alpha = 2$, $\beta = 0.5$).

in 20 years of social sciences experiments.

Most studies in our dataset estimate α and β assuming the utility function specification in FS. However, some studies explore variations of the original framework. First, for the sake of parsimony and mathematical tractability, FS assumed a piece-wise linear utility function. This predicts corner solutions in decision environments where we usually observe interior choices.[4] To improve on this, Bellemare, Kröger and van Soest (2008) assume a non-linear disutility from inequality and estimate the following utility function:

$$U_i(x) = x_i - \alpha_{1i} \max[x_j - x_i, 0] - \alpha_{2i} \max[x_j - x_i, 0]^2 - \beta_{1i} \max[x_i - x_j, 0] - \beta_{2i} \max[x_i - x_j]^2$$

If $\alpha_{2i} = \beta_{2i} = 0$, this model simplifies to FS. Bellemare and coauthors find the disutility from advantageous inequality to be nearly linear, while the disutility from advantageous inequality to be an increasing and concave function of the gap in outcomes.

A second simplification of the original model is the lack of any role for reciprocal motives. Morishima, Schunk, Bruhin, Ruff and Fehr (2012) and Bruhin, Fehr and Schunk (2019) augment FS to incorporate reciprocity, adopting the following utility function inspired by

[4]Consider, for example, a dictator game. If $\beta < 0.5$, the dictator keeps the whole budget; if $\beta > 0.5$, instead, the dictator shares the budget equally.

Fehr and Schmidt (1999) and Charness and Rabin (2002):

$$U_i(x_i, x_j) = (1 - \beta r - \alpha s - \theta q + \delta v)x_i + (\beta r + \alpha s + \theta q - \delta v)x_j,$$

where r, s, q, v are indicators for advantageous inequality, disadvantageous inequality, positive reciprocity and negative reciprocity respectively. Here, α and β have the usual meaning while θ and δ are reciprocity parameters. For example, if $\theta > 0$ and $\delta < 0$, an agent rewards kind actions at a cost (i.e., he displays positive reciprocity) and punishes selfish actions at a cost (i.e., he displays negative reciprocity). Note that, in this model, the sign of the disadvantageous inequality coefficient has the opposite meaning compared to the standard FS model: here, inequity aversion is captured by $\alpha < 0$ and $\beta > 0$.[5] Bellemare, Kröger and van Soest (2011) follow another route to introduce reciprocity in FS and assume the following utility function:

$$U_i(x_i, x_j) = x_i - (\alpha_i + l_i) \max[x_j - x_i, 0] - (\beta_i + k_i) \max[x_i - x_j, 0]$$

Here, depending on the intentions of the other players, l_i and k_i change the marginal disutility of disadvantageous or advantageous allocations.

Finally, the baseline FS model is sufficiently tractable to easily incorporate concerns in addition to or different from inequality aversion or reciprocity. For example, Alger and van Leeuwen (2021) augment the model by adding Kantian morality, whereby an individual evaluates her actions by considering what her payoff would be if others behaved in the same way; and Boun My, Lampach, Lefebvre and Magnani (2018) estimate a model of advantageous inequality aversion which includes loss aversion.

3 Data

[5] We take this into account when using the estimates from these papers in our meta-analysis.

((''inequality aversion'' OR ''inequity aversion'' OR ''envy'' OR ''guilt''
OR ''advantageous inequality'' OR ''disadvantageous inequality''
OR ''advantageous inequity'' OR ''disadvantageous inequity'')
AND
(experiment* OR estimat* OR surve*))

Figure 3: Query Used for Search on Web of Science and Google Scholar

3.1 Identification and Selection of Relevant Studies

In order to perform an unbiased meta-analysis, it is important to define a precise and unambiguous inclusion criteria. Our criterion is to include "all papers that estimated the parameters for disadvantageous inequality, α, and/or advantageous inequality, β, using the model by Fehr and Schmidt (1999)".[6]

The search procedure followed five steps. First, we read the narrative reviews by Fehr and Schmidt (2006) and Cooper and Kagel (2016) and searched on Google Scholar to find a first seed of papers that estimated α and β. Second, we read these papers to identify the best possible combination of keywords for a more detailed search. Third, we searched the scientific citation indexing database Web of Science using the query in Figure 3 . Since we are interested in estimations of the FS parameters, we restricted the search to papers that cite FS. This search was performed on February 8, 2022 and returned 433 articles. Fourth, we read these articles and excluded papers that were clearly irrelevant for our analysis — for example, articles that measured inequality aversion in animals or studies that, while reporting the results of dictator and ultimatum games, did not estimate the parameters of interest. Finally, we performed another search on Google Scholar using the same query from Figure 3 to find unpublished work or papers missing from the Web of Science database. The final dataset consists of 26 articles and the complete list is available in Appendix A.

We included in the dataset only studies reporting a precise measure of the parameters — for example, the value of an aggregate estimate or the mean of individual-level estimates.

[6] This definition includes also the models that use FS as baseline and augment it by adding other parameters as discussed in Section 2.

Four studies computed individual-level estimates for α and β but did not provide the mean or median for the parameters: Teyssier (2012), Corgnet, Espín and Hernán-González (2015) and Yang, Onderstal and Schram (2016) report only a scatter plot or a bar graph of the results; Müller and Rau (2019) discuss an imprecise distributions of the parameters based on the classification used in Blanco, Engelmann and Normann (2011). While it would be possible to recover an imprecise mean or median for the estimates in these studies, given the high level of arbitrariness this exercise would entail (for example, in assuming a uniform distribution of the parameter within each bin of a bar graph, or in evaluating the exact location of dots in a scatter plot), we decided not to include these papers in the dataset.

3.2 Data Construction

After identifying the relevant articles, we assembled the dataset for the meta-analysis by coding the estimates for α and β, the features of the studies and the features of the estimation methodology. The main variables of interest are the structural estimates for the two coefficients of advantageous and disadvantageous inequality aversion. In our 26 articles, these estimates take four forms: (i) *aggregate*, where a single value for α and β is estimated for the pooled data of all subjects in the study; (ii) *finite-mixture*, where a finite number of values for α and β alongside their distributions are estimated from the pooled data of all subjects; (iii) *individual-level mean*, where α and β are estimated separately for each subject and the mean value of the parameters is reported; and (iv) *individual-level median*, same as iii) but where the median (rather than the mean) is reported. The first, third and fourth types of estimates are ready to be used in the meta-analysis. For the finite-mixture estimates, we computed and coded a weighted average for each parameter.[7]

[7]For example, consider one of the finite-mixture estimates of α from Bruhin, Fehr and Schunk (2019) which reports the presence of three types in the population: $\alpha_1 = -0.159$, $\alpha_2 = -0.065$, and $\alpha_3 = 0.437$. The estimated frequencies associated with each of these types are $p_1 = 0.405$, $p_2 = 0.474$, and $p_3 = 0.121$. We construct a single estimate which is given by $\widehat{\alpha} = p_1\alpha_1 + p_2\alpha_2 + p_3\alpha_3 = -0.042$. Moreover, we construct a measure of estimation uncertainty as follows: first, we compute the standard deviation as $\text{SD} = \sum_i p_i(\alpha_i - \widehat{\alpha})^2$; second, we compute the standard error as $\text{SD}/\sqrt{n}$, where n is the sample size. This procedure disregards the estimated uncertainty of each α_i and the associated p_i but it greatly simplifies our

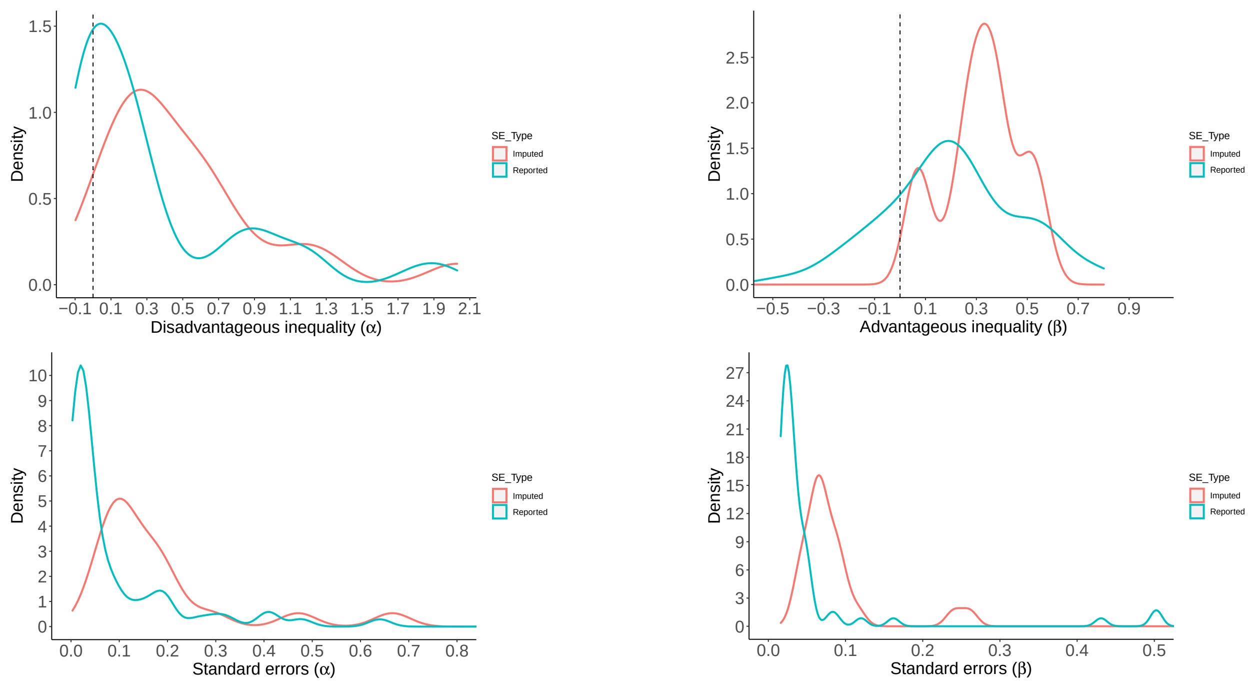


Figure 4: Distribution of Estimates and SEs for α and β As Function of SE Type. Note: The top two graphs show kernel density estimates (Gaussian with Silverman's rule of thumb) for the subsets of parameters with reported vs. imputed SEs; the bottom two graphs show kernel density estimates of SEs in the two subgroups; the x-axis in the density plot for β is truncated at -0.5 for better visual rendering but the kernel density uses all estimates; dotted vertical lines are at 0.

The measure of estimation uncertainty is another important variable to code in the dataset. This information is fundamental when conducting a meta-analysis: instead of simply averaging estimates from various studies, our aggregation procedure gives more weight to estimates that have lower SEs and, thus, are more precisely estimated (for example, because they are computed from experiments with a larger sample size). Out of 85 estimates in our dataset, the source reported the SEs for 23 estimates and, in other 38 cases, we were able to compute the SEs using the reported standard deviation and sample size. For the remaining 24 estimates, we did not have (direct or indirect) information about the SEs.[8]

We had two options: either drop the 24 estimates without SEs or approximate the SEs

analysis and it is similar to the procedure used by studies that report an individual-level mean.

[8]This usually happens for articles that compute individual-level estimates but report only the mean or median without the standard deviation. In one case, the standard deviation was reported but the sample size was unclear.

and keep these estimates in the dataset. We chose the latter option, especially since the observations would not be dropped randomly: as the density plots in the top row of Figure 4 show, there is a significant difference in the distribution of α and β between studies that report SEs and studies that did not and, thus, dropping the latter subset of estimates would introduce a bias in our results. For this reason, while using approximated SEs is a second-best, we deemed this as the more sensible option. Nonetheless, we present the main results of our meta-analysis both for the full sample and for the restricted sample that considers only estimates with reported (i.e., not approximated) SEs. For the approximation procedure, we followed Brown, Imai, Vieider and Camerer (2021): we first estimated the parameters characterizing the distribution in the data as $log(se_o) \sim \mathcal{N}(\mu_{se}, \sigma^2_{se})$; and we then used these distributional parameters to estimate the missing SEs as $log(se_m) \sim \mathcal{N}(\hat{\mu}_{se}, \hat{\sigma}^2_{se})$, where o stands for observed and m stands for missing. In order for this procedure to give a good approximation of the SEs, we need variables that are significantly associated with them. In our dataset, the values of the parameters are the best predictors for the values of their SEs, while other information available to us does not improve the estimates. We, thus, run the two following regressions to find $\hat{\mu}^{\alpha}_{se}$, $\hat{\mu}^{\beta}_{se}$ and their respective variances:[9]

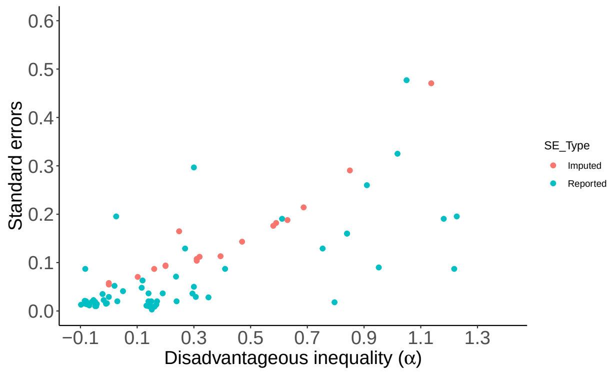
$$log(se^{\alpha}_o) = \delta_0 + \delta_1 \alpha_o + \delta_2 \beta_o$$

$$log(se^{\beta}_o) = \gamma_0 + \gamma_1 \alpha_o + \gamma_2 \beta_o$$

The two parameters explain 58% of the variance in the SEs for α and 31% of the variance in the SEs for β. Our approximation is, thus, better for α than for β.

Finally, we coded variables describing features of the studies and of the estimates. These variables include the paper publication status, the methodology (e.g., laboratory experiment, classroom experiment, online experiment), the subject population (e.g., non-representative sample of college students, non-representative sample of adults, sample representative of

[9]There are 4 estimates for which we only have a value for β. In this case, we only use β as a regressor other than the constant.



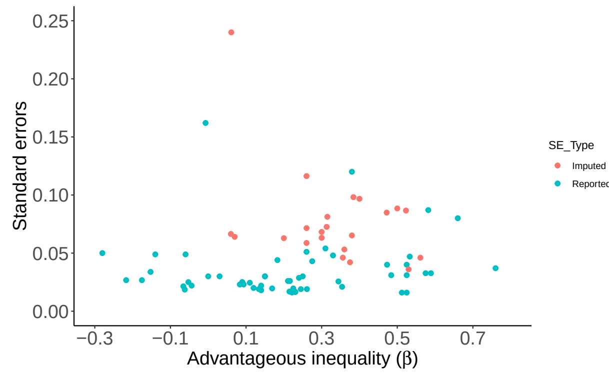


Figure 5: Scatter Plots of α and β SEs as a Function of SE Type. Note: The x-axis in the plot for α is truncated at 1.3 and the x-axis in the plot for β is truncated at -0.3 for better visual rendering.

a target population), subjects' location of residence, the task used to elicit the parameters (e.g., dictator game, ultimatum game, etc.), the reward type, and the utility function posited for the estimation (e.g., FS, FS plus Kantian morality, etc.). The next subsection discusses the distribution of the main features in our dataset. The full list is available in Appendix B.

3.3 Features of Studies and Estimates in the Dataset

As discussed in Section 3.1, we identified 26 articles which estimated the advantageous and disadvantageous inequality parameters in FS. In our dataset, we use as unit of measure a single *study* rather than a single *paper*. These two objects usually coincide but there is one exception: Beranek, Cubitt and Gächter (2015) report results of three distinct laboratory experiments conducted in the UK, the US and Turkey with three different samples. In our terminology, each of these three laboratory experiments comes from the same paper but corresponds to a different study. This means that, overall, we have 28 studies (discussed in 26 papers). These studies report 85 estimates of the advantageous and disadvantageous inequality parameters.

Table 1 reports the coded features of the 28 studies in our dataset. Among the 28 studies, 24 were presented in papers published (as of 08 February 2022) in economics, psychology, neuroscience and computer science journals. The majority of these 28 studies conducted tra-

Table 1: Features of the Studies ($N = 28$) in the Dataset

	Frequency	Proportion
Publication Status		
Published (as of February 8, 2022)	24	0.86
Unpublished	4	0.14
Methodology		
Laboratory Experiment	22	0.78
Classroom Experiment	1	0.04
Online Experiment	4	0.14
Multiple Methodologies	1	0.04
Geographic Location		
United States	6	0.21
Northern Europe (CH, DE, NL, SE, UK)	14	0.50
Southern Europe (FR, IT, ES, TR)	5	0.18
China	1	0.04
Multiple Locations	2	0.07
Subject Population		
College Students	22	0.79
Adults	3	0.11
Representative Sample of Dutch Population	2	0.07
Multiple Populations	1	0.04
Experimental Task Used to Estimate α		
Standard Dictator Game	1	0.03
Mini Dictator Game	2	0.07
Mini Dictator Game with equality-efficiency trade-off	9	0.31
Ultimatum Game	9	0.28
Other Game	10	0.31
Experimental Task Used to Estimate β		
Standard Dictator Game	1	0.03
Mini Dictator Game	2	0.06
Mini Dictator Game with equality-efficiency trade-off	15	0.48
Ultimatum Game	4	0.13
Other Game	10	0.32
Reward Type		
Money	27	0.96

Note: ‘Adults’ refers to mTurk workers, members of the Intergovernmental Panel on Climate Change, and workers who are part-time students; ‘Other Game’ includes bargaining game, gift exchange game, sequential prisoner dilemma, trust game, sequential public good game, and Stackelberg game; we label as ‘Mini Dictator Game’ a task where a single decision-maker chooses from a finite set of (exogenous) self/other allocations; in the papers, this task has different labels (‘ultimatum game abstracted from strategic interactions’, ‘choice menu’, ‘equality equivalence test’, ‘inequality list’, and ‘random ultimatum game’).

Table 2: Features of the Estimates ($N = 85$) in the Dataset.

	Frequency	Proportion
Utility Function in Estimated Model		
Linear FS	52	0.61
Non-Linear FS	2	0.02
Linear FS + Reciprocity	12	0.14
Linear FS + Kantian Morality	15	0.18
Linear FS + Intentions	2	0.02
Linear FS + Loss Aversion	2	0.02
Type of Estimates		
Aggregate	23	0.27
Finite Mixture	15	0.18
Individual Mean	36	0.42
Individual Median	11	0.13
Standard Errors		
Reported	61	0.72
Imputed	24	0.28

Notes: To avoid showing two separate tables, we use the 85 estimates for β; 'Linear FS' refers to the baseline model of inequality aversion in Fehr and Schmidt (1999).

ditional in-person laboratory experiments, while 4 studies conducted experiments online: one recruiting participants from mTurk, two using CentERpanel (an internet survey consisting of a representative sample of the adult Dutch population), and one contacting climate negotiators from the Intergovernmental Panel on Climate Change directly via email. The studies were conducted in 11 different countries (China, France, Germany, Italy, Netherlands, Spain, Sweden, Switzerland, Turkey, UK, and US) and involved mostly college students (22 studies out of 28), with 2 studies using a representative sample of the Dutch population (Bellemare, Kröger and van Soest, 2008, 2011) and 3 studies using a non-representative sample of adults (Dannenberg, Sturm and Vogt, 2010; He and Wu, 2016; Beranek, Cubitt and Gächter, 2015). All studies offered a monetary reward for participating in the experiments.

Table 2 reports the coded features of the 85 estimates in our dataset. Around 50% of the estimates (47 out of 85) come from studies that compute individual-level estimates of α and β and then report the mean and/or the median; 15 come from two studies, Bruhin, Fehr and Schunk (2019) and Alger and van Leeuwen (2021), which use finite-mixture models; while

23 come from studies which estimate parameters for a "representative" agent by pooling together all the available data. Around 60% of the estimates (52 out of 85) are computed assuming the original utility function specification from Fehr and Schmidt (1999); 12 and 15 estimates are computed assuming the model of inequity aversion augmented with, respectively, reciprocity parameters or Kantian morality; and the remaining 6 estimates use the baseline inequity aversion model in FS plus intentions, non-linearity or loss aversion. The parameters were elicited using choice data from a variety of games. However, even if some studies use more complex games (e.g., sequential prisoner's dilemmas or sequential public good games), more than half of the estimates come from experiments where subjects play a combination of ultimatum games and dictator games or variations of these.

4 Results

In this section, we first provide a non-parametric description of the 81 estimates of α and 85 estimates of β in our dataset (Section 4.1). We then fit a random-effects multi-level model to find average values for the advantageous and disadvantageous inequality aversion coefficients which take into account the different degree of precision of the various estimates and the correlation between multiple estimates from the same study. This analysis, which is presented in Section 4.2, provides the main results of the paper. In addition, we try to understand the heterogeneity across studies using the features coded in our dataset (Section 4.3). Finally, in Section 4.4, we investigate the issue of publication bias and selective reporting with the use of funnel plots and the FAT-PET procedure.

4.1 Non-Parametric Analysis

Figure 1 shows the distribution of the 81 estimates of α and of the 85 estimates of β in our dataset. The raw mean and median for α are, respectively, 0.39 and 0.19. Around a quarter of the estimates (21 out of 81) are equal to or less than 0 (in contrast with the assumption

Table 3: Summary Statistics for Disadvantageous Inequality (α)

	N	Min	1st Q	2nd Q	Mean	3rd Q	Max	SD
Estimate Type								
Aggregate	23	-0.10	0.01	0.16	0.32	0.30	1.89	0.48
Finite Mixture	15	-0.08	-0.05	0.14	0.09	0.16	0.35	0.14
Individual Mean	34	-0.08	0.20	0.52	0.63	1.03	2.03	0.59
Individual Median	9	-0.06	0.00	0.03	0.15	0.27	0.61	0.23
Experimental Task								
Game	45	-0.08	0.14	0.27	0.54	0.85	2.03	0.59
Individual Choice	36	-0.10	-0.32	0.11	0.20	0.32	1.05	0.30
Complete Dataset	81	-0.10	0.00	0.19	0.39	0.61	2.03	0.51

in FS). This suggests that some individuals are not hurt by unfavorable comparisons with others' outcomes. Table 3 shows that the estimates of α differs depending on whether the parameter is elicited in strategic environments (i.e., situations where the decision-maker's earnings depend also on the actions of others; e.g., the ultimatum game or the prisoner's dilemma) or in individual decision-making tasks (e.g., the dictator game or choice menus).[10] In the former case, the mean and the median of α are, respectively, 0.54 and 0.27; in the latter case, instead, the mean is 0.20 and the median is 0.11. This result is in line with the discussion in Dannenberg, Sturm and Vogt (2007), Dannenberg, Sturm and Vogt (2010), Kleine, Königstein and Rozsnyói (2014), Yang, Onderstal and Schram (2016), and He and Wu (2016) and it contributes to an ongoing debate on the economic construct captured by estimates of α. The significant difference observed in our dataset supports the hypothesis that, in strategic environments, α captures both equity and reciprocity concerns.

The estimates of β feature a bell-shaped distribution with a fatter left tail: the raw mean and median are, respectively, 0.23 and 0.17. While there are no estimates greater than 1 (as assumed in FS), around a fifth of the estimates (16 out of 85) are less than 0 (in contrast with the assumption in FS). This suggests that some individuals have "competitive" or "spiteful" preferences, so that they strictly prefer reducing other earnings (while keeping their own earnings unchanged). As shown in Table 4, contrary to α, estimates of β computed

[10] The full list of games used in the 28 studies from our dataset and whether they are considered strategic environments or individual decision-making tasks can be found in Table 9 in the Appendix.

Table 4: Summary Statistics for Advantageous Inequality (β)

	N	Min	1st Q	2nd Q	Mean	3rd Q	Max	SD
Estimate Type								
Aggregate	23	−0.46	0.04	0.15	0.20	0.35	0.80	0.32
Finite Mixture	15	−0.22	−0.06	0.08	0.03	0.15	0.23	0.15
Individual Mean	36	−2.12	0.18	0.29	0.15	0.36	0.59	0.55
Individual Median	11	−0.14	0.30	0.50	0.39	0.52	0.57	0.21
Experimental Task								
Game	34	−1.27	−0.62	0.09	0.07	0.23	0.80	0.34
Individual Choice	51	−2.12	0.21	0.30	0.24	0.50	0.76	0.45
Complete Dataset	85	−2.12	0.08	0.23	0.17	0.38	0.80	0.41

using choices from strategic environments are smaller than estimates computed using choices from individual decision-making tasks. This difference, which has not been discussed in the literature, can be rationalized by a higher discomfort from a favorable comparison with others' outcomes when the outcome is entirely attributable to one's own action and others only play a passive role (because of, e.g., image concerns).

Finally, we look at the joint distribution of the two parameters. Figure 6 shows a scatter plot of all 81 estimates for which we have a value for both α and β. We highlight two features of the joint distribution. First, a large number of observations (37 out of 81) lie above the 45-degree line where $\alpha \leq \beta$. This is in contrast with the assumption in FS and reflects the estimates from studies which compute individual-level estimates using choices in individual decision-making tasks (rather than in strategic environments). Second, the correlation between the two parameters is slightly positive but not significantly different from 0 ($\rho = 0.12$; $p = 0.28$). This is in line with the results discussed in Dannenberg, Sturm and Vogt (2007) Dannenberg, Sturm and Vogt (2010), Daruvala (2010), Blanco, Engelmann and Normann (2011), Morishima et al. (2012) and Beranek, Cubitt and Gächter (2015). This evidence suggests that the two parameters capture two separate traits of an individual's social preferences which are uncorrelated with each other or, at least, whose relationship is unclear.

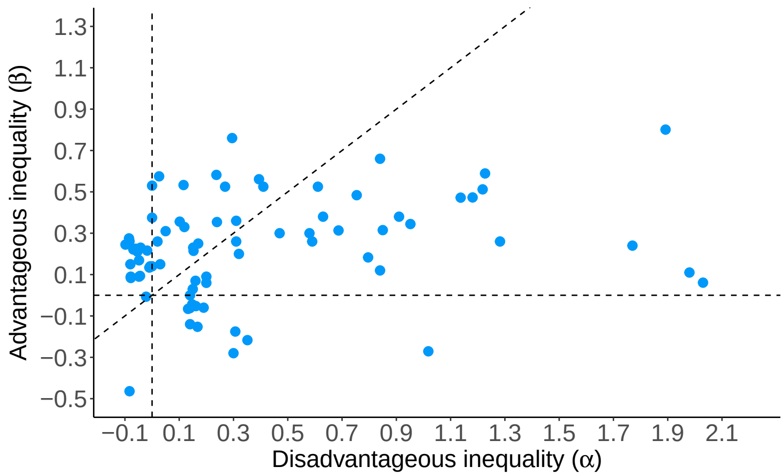


Figure 6: Scatter Plot of Disadvantageous (α) and Advantageous Inequality (β) Coefficients. Notes: We use the 81 estimates for which we have both a value for α and β; the vertical axis is truncated at -0.5 for better visual rendering.

4.2 Meta-Analytic Synthesis

The non-parametric analysis from Section 4.1 suffers from two potential pitfalls. First, all estimates are given equal weight, even if the information available to us suggests that the parameters computed in some studies are more reliable (i.e., more precisely estimated) than others. Second, estimates are assumed to be independent from one another, even if the same source and experimental study provides multiple estimates which are likely correlated with one another (e.g., because they are meant to capture the same subjects' underlying preferences). The econometric techniques adopted in this section tackle both issues.

In particular, we provide a meta-analytic estimation of a "weighted average" for α and β. There are two possible methodological approaches to this task. The first approach is a *frequentist analysis* that uses fixed- or random-effects (two-level or multi-level) models to find an average for a parameter. The second approach is aggregating the data using a *Bayesian hierarchical model*. We use both procedures and show that they return nearly identical results. We present the frequentist analysis in this section while we refer the reader to Appendix D for a detailed presentation of the Bayesian hierarchical model and its results.

Since there are two parameters, we have two options for how to conduct the meta-analytic synthesis with the frequentist approach. We can either estimate two univariate models or a single multivariate model that considers both parameters at the same time. While the latter procedure is the first-best (since it takes into account the possible inter-dependency between α and β), it is unfeasible in our case: we would need not only a measure for the variance of α and β but also a measure for their covariance, an information none of the studies in our dataset provides. For this reason, we conduct two separate univariate meta-analysis, one for α and one for β. While ignoring the dependence between the two variables might introduce a bias in our results, we note that the non-parametric analysis from the previous subsection suggests the correlation between α and β is weak and this reduces the concern. In Appendix E, we present the results of a multivariate model, which we estimate under the assumption that the covariance for each pair of parameters is 0.

We now describe our meta-analytic framework, which follows Imai, Rutter and Camerer (2021). We start from the simplest fixed-effects model (which we do not estimate but provides a building block for the ensuing discussion), continue with the two-level random-effects model and conclude with the more sophisticated model, the multi-level random-effects model. From this point on, our discussion of the methodology will focus on α, considering that the same concepts and equations (up to replacing α with β) also apply to β.

The fixed-effects model assumes the following:

$$\alpha_j = \alpha_0 + \epsilon_j, \tag{3}$$

where α_j is the parameter measured in study j, with $j = \{1, ..., k\}$, k being the total number of studies in the dataset; and α_0 is the "true" disadvantageous inequality parameter. The fixed-effects model assumes that all the parameters in the dataset come from a single homogeneous population and the reason why the value of α_j changes among studies is because of sampling errors, represented here by ϵ_j. It is assumed that $\epsilon_j \sim \mathcal{N}(0, v_j^2)$, where v_j^2 is the known sampling variance (i.e., the variance of the estimates). One way to get an estimate of α_0 is then to compute a weighted average of the α_j, with weights given by their precision:

$$\alpha_0^{FE} = \frac{\sum_{j=1}^{k} p_j^{FE} \alpha_j}{\sum_{j=1}^{k} p_j^{FE}} \tag{4}$$

where $p_j^{FE} = \frac{1}{v_j^2}$. This equations says that parameters with a lower variance are given more weight in the aggregation. Given its assumptions, a fixed-effects model performs well only if there is no heterogeneity across studies, since the only reason for the parameters to differ is due to sampling variance. If the studies are not homogeneous — as it is the case in our dataset because different articles employ different subject populations, experimental tasks, utility specifications, etc. — then a fixed-effects model would perform rather poorly.

Alternatively, we can estimate a two-level random-effects model (DerSimonian and Laird, 1986). This model assumes that:

$$\alpha_j = \mu_j + \epsilon_j \tag{5}$$

$$\mu_j = \alpha_0 + \xi_j. \tag{6}$$

The observed parameter, α_j, is an estimator of the study's true effect size, μ_j, plus a sampling error, ϵ_j. The true effect size, μ_j, comes from a homogeneous population with a "grand mean", α_0, plus a second source of error, ξ_j, which is assumed to be distributed as $\xi_j \sim \mathcal{N}(0, \tau^2)$, where τ^2 captures between-observations heterogeneity. We can combine the two equations above to get:

$$\alpha_j = \alpha_0 + \xi_j + \epsilon_j \tag{7}$$

This equation makes clear that ϵ_j is the sampling error for α_j, which is an estimate for μ_j (the true effect size). This is, in turn decomposed into the grand mean, α_0, and the second error term, ξ_j. If $\tau^2 = 0$, meaning that there is no between-observations heterogeneity, the two-level random-effects model coincide with the fixed-effects model. Endowed with this model, we can get an estimate for α_0 by taking again a weighted average of the form:

$$\alpha_0^{RE} = \frac{\sum_{j=1}^{k} p_j^{RE} \alpha_j}{\sum_{j=1}^{k} p_j^{RE}}, \tag{8}$$

where, in this case, the weights are given by $p_j^{RE} = \frac{1}{v_j^2 + \hat{\tau}^2}$, with $\hat{\tau}^2$ being an estimate of τ^2. The weights take into account both the precision of the observed parameters and the between-observation heterogeneity. The two-level random-effects model assumes that observations are independent. In our dataset, this is most likely not the case, since many articles provide more than one estimate — for example, by computing α and β using different econometric approaches or utility function specifications. In order to account for the for the possible correlation across estimates from the same study, we fit a random-effects model that uses cluster-robust variance estimation at the study level.[11]

A third alternative is a multi-level random-effects model as in Konstantopoulos (2011) and Van den Noortgate et al. (2013). A multi-level model is another way to handle estimates that are statistically dependent. Denote with α_{ij} the jth estimate of parameter α from study i. Then, the first level is defined as:

$$\alpha_{ij} = \mu_{ij} + \epsilon_{ij}, \tag{9}$$

where μ_{ij} is the "true" effect size (in this case, the "true" disadvantageous inequality parameter) and the error term is distributed as $\epsilon_{ij} \sim \mathcal{N}(0, v_{ij}^2)$. The second level is:

$$\mu_{ij} = \theta_i + \xi_{ij}^{(2)}, \tag{10}$$

where θ_i represents the average disadvantageous inequality in study i and $\xi_{ij}^{(2)} \sim \mathcal{N}(0, \tau_{(2)}^2)$.

[11]We adopt the cluster-robust correction in Hedges, Tipton and Johnson (2010). We also implement a small-sample adjustment as suggested in Bell and McCaffrey (2002), Tipton (2015) and Pustejovsky and Tipton (2018).

The last level is:

$$\theta_i = \alpha_0 + \xi_j^{(3)}, \tag{11}$$

where α_0 is the population mean of α (what we are interested in) and $\xi_j^{(3)} \sim \mathcal{N}(0, \tau_{(3)}^2)$. We can combine the three levels into a single equation to have

$$\alpha_{ij} = \alpha_0 + \xi_{ij}^{(2)} + \xi_j^{(3)} + \epsilon_{ij}. \tag{12}$$

Compared to the two-level random-effects model, here there are two heterogeneity terms in addition to the sampling error: $\xi_{ij}^{(2)}$ represents the within-cluster heterogeneity, i.e., the heterogeneity that is present among different estimates in a single study; $\xi_j^{(3)}$, instead, stands for the between-cluster heterogeneity, with a large value for $\tau_{(3)}^2$ indicating that the "true" disadvantageous inequality parameter varies a lot between different studies.

Before fitting the two-level and the multi-level random-effects models described above, we run some diagnostic checks to exclude potentially "overly influential" observations by computing DFBETAS (Belsley, Kuh and Welsch, 1980), which measure the effect of dropping one observation on a regression coefficient. We use the classification in Bollen and Jackman (1985) and identify an observation to be influential if $|\text{DBETAS}| > 1$. Since none of the coefficients exceed the threshold, we do not remove any observation from the analysis.

Tables 5 and 6 report the results of the meta-analytic synthesis. In discussing these results, we focus on the estimates obtained in the full sample, that is, without removing studies whose SEs we had to approximate. Results for the restricted sample of studies with reported SEs are available in the same tables and are qualitatively identical. Starting with the disadvantageous inequality parameter (α), both the two-level and the multi-level random-effect specifications return an estimate that is positive and significantly different from zero. Our meta-analysis, thus, supports the hypothesis that people are concerned about equity when they are in a disadvantageous situation. The coefficient in the two-level model is 0.279 while the coefficient in the multi-level model is 0.425. The difference

Table 5: Meta-Analytic Average of Disadvantageous Inequality (α)

	(1)	(2)	(3)	(4)
Disadvantageous Inequality Coefficient (α_0)	0.279	0.264	0.425	0.448
	(0.098)	(0.122)	(0.091)	(0.129)
p-value	0.009	0.046	< 0.0001	0.001
$\hat{\tau}^2$	0.110	0.134		
I^2	99.69	99.81		
I^2_{within}			10.14	7.20
$I^2_{between}$			89.68	92.70
Observations	81	61	81	61
Model	RE	RE	ML	ML
Sample	Full	Restricted	Full	Restricted

Notes: Columns (1) and (3) estimate a two-level random-effects (RE) and multi-level random-effects (ML) model on the full sample; columns (2) and (4) focus on studies with reported (i.e., non-approximated) SEs; p-values are for a two-sided test with null hypothesis $H_0 : \alpha_0 = 0$; SEs in parenthesis are cluster-robust (Hedges, Tipton and Johnson, 2010) with a small-sample adjustment; in both RE and ML models, we use the restricted maximum likelihood method.

between the two specifications is due to the fact that many estimates come from a single paper — for example, Alger and van Leeuwen (2021) report 21 values for α. Even if the cluster-robust SEs try to address this issue, the results from the two-level model might be driven by these observations. Both estimates are smaller than the average value from the distribution reported in FS (0.850). From the I^2 statistics (Higgins and Thompson, 2002), we learn that nearly all of the variability (99%) in the two-level random-effects model is due to between observations heterogeneity rather than sampling variance.[12] In the multi-level model, instead, around 10% of the variability in the data is due to heterogeneity within studies (I^2_{within}), 89% to heterogeneity across studies ($I^2_{between}$) and the remainder to sampling variance.

The meta-analytic average of β in the two-level random-effects model is 0.226, smaller than in the multi-level specification for the same reason discussed above. Given the larger estimation uncertainty due to the cluster-robust SEs, the parameter is statistically different

[12]The I^2 statistics is computed as $I^2 = 100\left(\frac{\hat{\tau}^2}{\hat{\tau}^2+s^2}\right)$ where $s^2 = \frac{(k-1)\sum p_j}{(\sum p_j)^2+\sum p_j^2}$ with $p_j = \frac{1}{v_j^2}$.

Table 6: Meta-Analytic Average of Advantageous Inequality (β)

	(1)	(2)	(3)	(4)
Advantageous Inequality Coefficient (β_0)	0.226	0.199	0.291	0.278
	(0.091)	(0.110)	(0.037)	(0.048)
p-value	0.020	0.090	< 0.0001	< 0.0001
$\hat{\tau}^2$	0.054	0.057		
I^2	98.35	98.78		
I^2_{within}			34.50	36.53
$I^2_{between}$			63.10	62.80
Observations	85	61	85	61
Model	RE	RE	ML	ML
Sample	Full	Restricted	Full	Restricted

Notes: Columns (1) and (3) estimate a two-level random-effects (RE) and a multi-level random-effects (ML) model on the full sample; columns (2) and (4) focus on studies with reported (i.e., non-approximated) SEs; p-values are for a two-sided test with null hypothesis $H_0 : \beta_0 = 0$; SEs in parenthesis are cluster-robust (Hedges, Tipton and Johnson, 2010) with a small-sample adjustment; in both RE and ML models, we use the restricted maximum likelihood method.

from zero only at the 5% significance level (10% in the restricted sample with reported SEs). In the preferred specification with a multi-level random-effects model, the estimate of β is 0.291 and this is statistically different from zero at any conventional significance level. This value is in line with the weighted average and the median of β from the distribution reported in FS (0.315 and 0.290). We, thus, find evidence of equity concerns in the realm of advantageous situations. While the FS assumptions hold in our meta-analysis, since $\alpha \geq \beta$ and $0 \leq \beta < 1$, we cannot claim that the estimate of α is statistically greater than the estimate of β. The I^2 statistics shows that, in the two-level random-effects specification, 98.35% of the variability in β can be attributed to between observations heterogeneity; in the multi-level model, instead, around 35% of the variability is due to within study heterogeneity and around 63% to between studies heterogeneity.

4.3 Explaining Heterogeneity

The estimates in our dataset come from studies that are very different from each other, for example, because of subject population, the tasks subjects performed during the experiment, the utility function that was assumed in the estimation procedure and so on. It is then far fetched that the estimates for α and β depend mainly on sampling errors, either at the observation or study level, as we did previously. In order to explain the heterogeneity, we estimate a model of this form (for α; the equation for β is analogous):

$$\alpha_{ij} = \delta_0 + \delta_1 SE^{\alpha}_{ij} + \phi X_{ij} + \epsilon_{ij} \tag{13}$$

where X_{ij} is a set of moderator variables coded in our dataset. Given the high amount of coded variables and the few observations for some of these, it is unclear what model should we use to explain the heterogeneity in the parameters. We then run three different regressions varying the number of explanatory variables, from the most parsimonious model to the one including all moderator variables in the dataset.

Since X_{ij} is composed of dummy variables, each coefficient represents the additional effect on the dependent variable with respect to the baseline condition. We chose the baseline conditions as follows: for methodology, the omitted category is laboratory experiment; for subject population, the omitted category is college students; for geographic location, the omitted category is the US, for the utility function specification, the omitted category is linear FS; and for experimental task, the omitted category is non-strategic environment. Thus, a positive coefficient indicates more aversion to disadvantageous or advantageous inequality compared to a laboratory study with college students conducted in the US which estimated α and β in a non-strategic environment using the utility function in FS.

The meta-regressions for α and β are presented in Table 7. Columns 1 and 4 consider as moderator variables only the SEs and the dummy for the strategic environment; columns 2 and 5 add dummies for the different utility specifications; and columns 3 and 6 contain

Table 7: Explaining Heterogeneity

	Disadvantageous Inequality (α)			Advantageous Inequality (β)		
	(1)	(2)	(3)	(4)	(5)	(6)
Estimates' SE	2.771*	3.413***	2.909***	0.156	-1.026	-1.324
	(1.185)	(0.485)	(0.463)	(1.273)	(1.319)	(1.223)
Strategic Environment	0.202***	0.002	-0.011	-0.237**	-0.186**	-0.102*
	(0.030)	(0.024)	(0.007)	(0.076)	(0.060)	(0.051)
FS + Reciprocity		-0.072***	-0.059***		-0.152*	-0.070
		(0.014)	(0.006)		(0.070)	(0.066)
FS + Kantian Morality		0.182***	0.200***		-0.178***	-0.164**
		(0.025)	(0.004)		(0.035)	(0.059)
FS + Non-Linearity		-0.083	-0.518**		0.366	0.356
		(0.194)	(0.171)		(0.748)	(0.765)
FS + Intentions		0.772***	0.150***		-0.030	-0.203
		(0.028)	(0.035)		(0.040)	(0.109)
Online Experiment			0.646***			0.195***
			(0.036)			(0.051)
Classroom Experiment			0.073			0.213**
			(0.045)			(0.073)
Adults			-0.199*			0.047
			(0.077)			(0.055)
Northern Europe			-0.152			0.025
			(0.080)			(0.073)
Southern Europe			-0.104			0.319***
			(0.113)			(0.068)
Multiple Countries			-0.647***			0.077
			(0.049)			(0.043)
Constant	-0.091**	-0.052***	0.099	0.314***	0.410***	0.295***
	(0.029)	(0.015)	(0.081)	(0.090)	(0.097)	(0.077)
Observations	81	81	81	85	85	85
R^2	0.312	0.868	0.905	0.357	0.505	0.655
Adjusted R^2	0.294	0.857	0.888	0.342	0.467	0.597

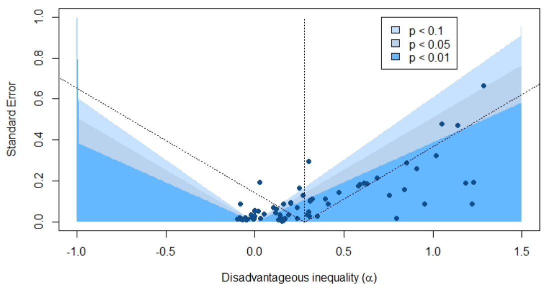
Notes: SEs are clustered at the study level. *p<0.05; **p<0.01; ***p<0.001.

all moderator variables. While we have a small number of observations for some categories and should thus be cautious in inferring too much from these coefficients, we nonetheless highlight some interesting patterns. Adults are less concerned about (disadvantageous) inequality than college students and participants to classroom and online experiments are more concerned about inequality than participants to traditional laboratory experiments. Interestingly, participants from Southern Europe (France, Italy, Spain, and Turkey) are more averse to advantageous inequality than participants from the US. A regression analogous to the one in column 6 where Southern Europe is the baseline geographic location shows that participants from Southern Europe are more averse to advantageous inequality than participants from both the US and Northern Europe (UK, Germany, Netherlands, Sweden, and Switzerland).

4.4 Identifying Selective Reporting and Publication Bias

One aspect to keep in mind when conducting a meta-analysis is the problem of selective reporting or publication bias. The main concern arises when a theory strongly predicts certain results — for example, the magnitude or significance of some statistical relationships — and the literature anchors itself towards the same findings. This causes problems when, for example, new evidence reporting "unusual" or "unconventional" results is not taken in consideration because it goes against this norm. Articles are, then, either rejected and not published in journals or simply not written to begin with (the "file-drawer" problem). Beyond biases in the publication process, there are other sources of selective reporting that go from conscious frauds to more morally gray actions like "p-hacking".

In order to gauge the occurrence of publication bias in studies estimating inequality aversion coefficients, we first look at funnel plots. Funnel plots are scatter plots of the parameter estimates and of their SEs. The idea is that estimates with a higher precision should lie close to the meta-synthetic mean of the parameters, while estimates far from this mean should show a lower precision. Without selective reporting, we expect to see a funnel-



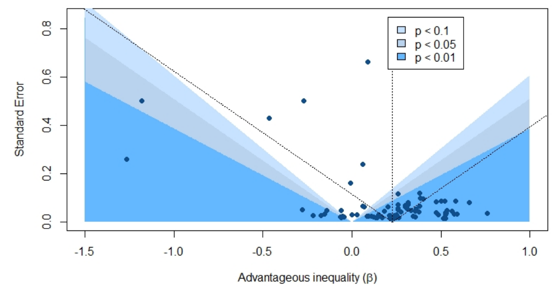


Figure 7: Funnel Plots of Disadvantageous (α) and Advantageous Inequality (β) Coefficients. Notes: Shaded blue areas correspond to p-values where H_0 is the parameter being equal to zero. The horizontal axis is truncated at 1.5 (α) and -1.5 (β) for better visual rendering.

shaped distribution which is symmetric around the "average" parameter value. An absence of symmetry can hint to "missing" studies and so to the presence of publication bias. Figure 7 shows the funnel plots for the advantageous and disadvantageous inequality coefficients. The distribution for α looks highly asymmetric: observations with a negative (and large in magnitude) value of α which is imprecisely estimated are "missing". A similar, albeit more attenuated, effect is present also for β: there are no studies reporting a large and imprecisely estimated positive value of this coefficient.

A second approach to detect selective reporting is the FAT-PET procedure, which consists in regressing the parameters on their SEs. If there is no publication bias, the reported estimates should be uncorrelated with the SEs. We then estimate the two following equations:

$$\alpha_{ij} = \delta_0 + \delta_1 SE^{\alpha}_{ij} + \epsilon_{ij} \tag{14}$$

$$\beta_{ij} = \gamma_0 + \gamma_1 SE^{\beta}_{ij} + \xi_{ij} \tag{15}$$

In this model, δ_1 and γ_1 capture the degree of selective reporting bias while δ_0 and γ_0 represent the selection-bias-corrected value of the parameters. The regression is usually run using a weighted least square procedure, where the weights are given by the inverse of the variance (Stanley and Doucouliagos, 2012). This exercise tests at the same time for asymmetry in the funnel plots (FAT; Egger et al. 1997; Stanley 2005; Stanley and Doucouliagos 2017) and for a "true effect" of the parameters beyond publication selection (PET). The coefficients δ_1 and γ_1 are positive but statistically indistinguishable from zero ($\delta_1 = 0.82$ with p-value$= 0.451$; $\gamma_1 = 0.55$ with p-value$= 0.692$). On the other hand, the constants δ_0 and γ_0 are positive and statistically significant, indicating the presence of inequity aversion even after correcting for possible publication bias ($\delta_0 = 0.1$ with p-value < 0.0001; $\gamma_0 = 0.2$ with p-value < 0.0001).

The results from the FAT-PET procedure suggests that the asymmetry in the funnel plots could be generated in the absence of publication bias — for example, because of constraints in the estimation of α and β when eliciting these parameters with the experimental tasks typically employed by the literature.[13] Moreover, while the funnel plot procedure assumes that the two parameters can take any value, some values are more plausible than others since these coefficients are meant to capture social preferences. In particular, it would be surprising to find values of α smaller than -1 and values of β larger than 1, which imply that an individual is willing to burn money just to increase the gap in outcomes when ahead or just to reduce the gap when behind. Indeed, the 81 estimates of α and the 85 estimates of β in our dataset never take values beyond those thresholds and this can hardly be deemed proof of publication bias.

5 Conclusion

In this paper, we reported the results of a meta-analysis of empirical estimates of the inequality aversion coefficients in models of outcome-based other-regarding preferences à la Fehr and Schmidt (1999). We conduct both a frequentist analysis (using a multi-level random-effects model) and a Bayesian analysis (using a Bayesian hierarchical model) to provide a "weighted average" for α and β. The results from the two approaches are nearly identical and support the hypothesis of inequality concerns. From the frequentist analysis, we learn that the mean

[13] For example, the ultimatum and dictator games used in Blanco, Engelmann and Normann (2011) lead to feasible estimates in the following ranges: $\alpha \in [0, 4.5]$ and $\beta \in [0, 1]$.

envy coefficient is 0.425 with a 95% confidence interval of $[0.244, 0.606]$; the mean guilt coefficient is, instead, 0.291 with a 95% confidence interval $[0.218, 0.363]$.[14] This means that, on average, an individual is willing to spend €0.41 to increase others' earnings by €1 when ahead, and €0.74 to decrease others' earnings by €1 when behind. The theoretical assumptions $\alpha \geq \beta$ and $0 \leq \beta < 1$ are upheld in our empirical analysis, but we cannot conclude that the disadvantageous inequality coefficient is statistically greater than the coefficient for advantageous inequality. We also observe no correlation between the two parameters.

Our analysis suggests two avenues for further research on social preferences. First, while this is not always a clean comparison (since studies conducted in different countries differ also in other dimensions), the analysis of heterogeneity in Section 4.3 shows that participants from Southern Europeans are more sensitive to advantageous inequality than participants from Northern Europe and the US. The variation of inequality aversion across (and within) countries should be explored in experimental studies which allow the estimation of parameters using the same methodology and reaching participants from a wider set of countries and cultures. Second, the sensitivity of the estimates to the experimental task (strategic versus non-strategic) and to the utility function specification (e.g., whether Kantian morality is included or not) points to the inter-dependency between different facets of social preferences and to the crucial role played by the decision environment in making one more salient than others. We believe that studying outcome-based social preferences (e.g., inequality aversion), intention-based social preferences (e.g., reciprocity), and image concerns in the same theoretical framework and designing experiments which allow the joint estimation of parameters from these models is an important step for a better understanding of social preferences.

References

Alger, Ingela and Boris van Leeuwen, "Estimating Social Preferences and Kantian

[14] In the Bayesian analysis, the mean envy coefficient is 0.426 with a 95% probability that the true value lies in the interval $[0.240, 0.620]$; the mean guilt coefficient is, instead, 0.290 with a 95% probability that the true value lies in the interval $[0.212, 0.366]$.

Morality in Strategic Interactions," 2021. Unpublished Manuscript.

Anthoff, David, Cameron Hepburn, and Richard SJ Tol, "Equity Weighting and the Marginal Damage Costs of Climate Change," *Ecological Economics*, 2009, *68* (3), 836–849.

Azar, Christian and Thomas Sterner, "Discounting and Distributional Considerations in the Context of Global Warming," *Ecological Economics*, 1996, *19* (2), 169–184.

Baranski, Andrzej and Rebecca Morton, "The Determinants of Multilateral Bargaining: A Comprehensive Analysis of Baron and Ferejohn Majoritarian Bargaining Experiments," *Experimental Economics*, 2021, pp. 1–30.

Bell, Robert M and Daniel F McCaffrey, "Bias Reduction in Standard Errors for Linear Regression with Multi-Stage Samples," *Survey Methodology*, 2002, *28* (2), 169–182.

Bellemare, Charles, Sabine Kröger, and Arthur van Soest, "Measuring Inequity Aversion in a Heterogeneous Population Using Experimental Decisions and Subjective Probabilities," *Econometrica*, 2008, *76* (4), 815–839.

_ , _ , and _ , "Preferences, Intentions, and Expectation Violations: A Large-Scale Experiment with a Representative Subject Pool," *Journal of Economic Behavior & Organization*, 2011, *78* (3), 349–365.

Belsley, David A, Edwin Kuh, and Roy E Welsch, *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, Hoboken, NJ: Wiley, 1980.

Beranek, Benjamin, Robin Cubitt, and Simon Gächter, "Stated and Revealed Inequality Aversion in Three Subject Pools," *Journal of the Economic Science Association*, 2015, *1* (1), 43–58.

Berg, Joyce, John Dickhaut, and Kevin McCabe, "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 1995, *10* (1), 122–142.

Berkey, CS, DC Hoaglin, A Antczak-Bouckoms, F Mosteller, and GA Colditz, "Meta-Analysis of Multiple Outcomes by Regression with Random Effects," *Statistics in Medicine*, 1998, *17* (22), 2537–2550.

Blanco, Mariana, Dirk Engelmann, and Hans Theo Normann, "A Within-Subject

Analysis of Other-Regarding Preferences," *Games and Economic Behavior*, 2011, *72* (2), 321–338.

Bollen, Kenneth A and Robert W Jackman, "Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases," *Sociological Methods & Research*, 1985, *13* (4), 510–542.

Brown, Alexander L, Taisuke Imai, Ferdinand Vieider, and Colin F Camerer, "Meta-Analysis of Empirical Estimates of Loss-Aversion," 2021. Unpublished Manuscript.

Bruhin, Adrian, Ernst Fehr, and Daniel Schunk, "The Many Faces of Human Sociality: Uncovering the Distribution and Stability of Social Preferences," *Journal of the European Economic Association*, 2019, *17* (4), 1025–1069.

Burks, Stephen V, Jeffrey P Carpenter, and Eric Verhoogen, "Playing Both Roles in the Trust Game," *Journal of Economic Behavior & Organization*, 2003, *51* (2), 195–216.

Camerer, Colin F, Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan, Emma Heikensten, Taisuke Imai, Siri Isaksson, Thomas Pfeiffer, Michael Razen, and Hang Wu, "Evaluating Replicability of Laboratory Experiments in Economics," *Science*, 2016, *351* (6280), 1433–1436.

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell, "Stan: A Probabilistic Programming Language," *Journal of Statistical Software*, 2017, *76* (1).

Charness, Gary and Matthew Rabin, "Understanding Social Preferences with Simple Tests," *The Quarterly Journal of Economics*, 2002, *117* (3), 817–869.

Cooper, David J and E Glenn Dutcher, "The Dynamics of Responder Behavior in Ultimatum Games: a Meta-Study," *Experimental Economics*, 2011, *14* (4), 519–546.

_ **and John H Kagel**, *Other Regarding Preferences: A Selective Survey of Experimental Results. The Handbook of Experimental Economics, Volume 2*, Princeton University Press, 2016.

Corgnet, Brice, Antonio M Espín, and Roberto Hernán-González, "The Cognitive Basis of Social Behavior: Cognitive Reflection Overrides Antisocial but not always Prosocial Motives," *Frontiers in Behavioral Neuroscience*, 2015, *9*, 287.

Dannenberg, Astrid, Bodo Sturm, and Carsten Vogt, "Do equity preferences matter in climate negotiations? An experimental investigation," 2007. Unpublished Manuscript.

_ , _ , and _ , "Do Equity Preferences Matter for Climate Negotiators? An Experimental Investigation," *Environmental and Resource Economics*, 2010, *47* (1), 91–109.

Daruvala, Dinky, "Would the Right Social Preference Model Please Stand Up!," *Journal of Economic Behavior & Organization*, 2010, *73* (2), 199–208.

den Noortgate, Wim Van, José Antonio López-López, Fulgencio Marín-Martínez, and Julio Sánchez-Meca, "Three-Level Meta-Analysis of Dependent Effect Sizes," *Behavior Research Methods*, 2013, *45* (2), 576–594.

DerSimonian, Rebecca and Nan Laird, "Meta-Analysis in Clinical Trials," *Controlled Clinical Trials*, 1986, *7* (3), 177–188.

Dreber, Anna and Magnus Johannesson, "Statistical Significance and the Replication Crisis in the Social Sciences," in "Oxford Research Encyclopedia of Economics and Finance" 2019.

Eckel, Catherine C and Philip J Grossman, "Chivalry and Solidarity in Ultimatum Games," *Economic Inquiry*, 2001, *39* (2), 171–188.

Egger, Matthias, George Davey Smith, Martin Schneider, and Christoph Minder, "Bias in Meta-Analysis Detected by a Simple, Graphical Test," *Bmj*, 1997, *315* (7109), 629–634.

Embrey, Matthew, Guillaume R Fréchette, and Sevgi Yuksel, "Cooperation in the Finitely Repeated Prisoner's Dilemma," *The Quarterly Journal of Economics*, 2018, *133* (1), 509–551.

Engel, Christoph, "Dictator Games: A Meta Study," *Experimental Economics*, 2011, *14*

(4), 583–610.

Epper, Thomas, Ernst Fehr, and Julien Senn, "Other-Regarding Preferences and Redistributive Politics," Technical Report, Working Paper 2020.

Fehr, Ernst, Alexander Klein, and Klaus M Schmidt, "Fairness and Contract Design," *Econometrica*, 2007, *75* (1), 121–154.

_ **and Klaus M Schmidt**, "A Theory of Fairness, Competition, and Cooperation," *The Quarterly Journal of Economics*, 1999, *114* (3), 817–868.

_ **and** _ , "Fairness and Incentives in a Multi-Task Principal-Agent Model," *The Scandinavian Journal of Economics*, 2004, *106* (3), 453–474.

_ **and** _ , "The Economics of Fairness, Reciprocity and Altruism–Experimental Evidence and New Theories," *Handbook of the Economics of Giving, Altruism and Reciprocity*, 2006, *1*, 615–691.

_ **, Susanne Kremhelmer, and Klaus M Schmidt**, "Fairness and the Optimal Allocation of Ownership Rights," *The Economic Journal*, 2008, *118* (531), 1262–1284.

Forsythe, Robert, Joel L Horowitz, Nathan E Savin, and Martin Sefton, "Fairness in simple Bargaining Experiments," *Games and Economic Behavior*, 1994, *6* (3), 347–369.

Gelman, Andrew and Iain Pardoe, "Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models," *Technometrics*, 2006, *48* (2), 241–251.

Glass, Gene V, "Primary, Secondary, and Meta-Analysis of Research," *Educational Researcher*, 1976, *5* (10), 3–8.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze, "An Experimental Analysis of Ultimatum Bargaining," *Journal of Economic Behavior & Organization*, 1982, *3* (4), 367–388.

He, Haoran and Keyu Wu, "Choice Set, Relative Income, and Inequity Aversion: an Experimental Investigation," *Journal of Economic Psychology*, 2016, *54*, 177–193.

Hedges, Larry V, Elizabeth Tipton, and Matthew C Johnson, "Robust Variance Estimation in Meta-Regression with Dependent Effect Size Estimates," *Research Synthesis*

Methods, 2010, *1* (1), 39–65.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger, Natalie Smith Henrich, Kim Hill, Francisco Gil-White, Michael Gurven, Frank W. Marlowe, John Q Patton, and David Tracer, "“Economic Man” in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies," *Behavioral and Brain Sciences*, 2005, *28* (6), 795–815.

Higgins, Julian PT and Simon G Thompson, "Quantifying Heterogeneity in a Meta-Analysis," *Statistics in Medicine*, 2002, *21* (11), 1539–1558.

Hoffman, Elizabeth, Kevin McCabe, Keith Shachat, and Vernon Smith, "Preferences, Property Rights, and Anonymity in Bargaining Games," *Games and Economic Behavior*, 1994, *7* (3), 346–380.

Huck, Steffen, Wieland Müller, and Hans-Theo Normann, "Stackelberg Beats Cournot: On Collusion and Efficiency in Experimental Markets," *The Economic Journal*, 2001, *111* (474), 749–765.

Imai, Taisuke, Tom A Rutter, and Colin F Camerer, "Meta-Analysis of Estimation of Time Discounting of Rewards," *Unpublished manuscript*, 2018.

_ , _ , and _ , "Meta-Analysis of Present-Bias Estimation Using Convex Time Budgets," *The Economic Journal*, 2021, *131* (636), 1788–1814.

Ishak, K Jack, Robert W Platt, Lawrence Joseph, and James A Hanley, "Impact of Approximating or Ignoring Within-Study Covariances in Multivariate Meta-Analyses," *Statistics in Medicine*, 2008, *27* (5), 670–686.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291.

Kirkham, Jamie J, Richard D Riley, and Paula R Williamson, "A Multivariate Meta-Analysis Approach for Reducing the Impact of Outcome Reporting Bias in Systematic Reviews," *Statistics in Medicine*, 2012, *31* (20), 2179–2195.

Kleine, Fabian, Manfred Königstein, and Balázs Rozsnyói, "Voluntary Leadership in an Experimental Trust Game," *Journal of Economic Behavior & Organization*, 2014, *108*, 442–452.

Konstantopoulos, Spyros, "Fixed Effects and Variance Components Estimation in Three-Level Meta-Analysis," *Research Synthesis Methods*, 2011, *2* (1), 61–76.

Ledyard, John O, "Public Goods: A Survey of Experimental Research," in John Kagel and Roth Alvin, eds., *Handbook of Experimental Economics*, Princeton University Press, 1995.

Loewenstein, George F, Leigh Thompson, and Max H Bazerman, "Social Utility and Decision Making in Interpersonal Contexts.," *Journal of Personality and Social psychology*, 1989, *57* (3), 426.

Lü, Xiaobo, Kenneth Scheve, and Matthew J Slaughter, "Inequity Aversion and the International Distribution of Trade Protection," *American Journal of Political Science*, 2012, *56* (3), 638–654.

Morishima, Yosuke, Daniel Schunk, Adrian Bruhin, Christian C Ruff, and Ernst Fehr, "Linking Brain Structure and Activation in Temporoparietal Junction to Explain the Neurobiology of Human Altruism," *Neuron*, 2012, *75* (1), 73–79.

Müller, Stephan and Holger A Rau, "Decisions Under Uncertainty in Social Contexts," *Games and Economic Behavior*, 2019, *116*, 73–95.

My, Kene Boun, Nicolas Lampach, Mathieu Lefebvre, and Jacopo Magnani, "Effects of Gain-Loss Frames on Advantageous Inequality Aversion," *Journal of the Economic Science Association*, 2018, *4* (2), 99–109.

Normann, Hans-Theo and Holger A Rau, "Simultaneous and Sequential Contributions to Step-Level Public Goods: One Versus Two Provision Levels," *Journal of Conflict Resolution*, 2015, *59* (7), 1273–1300.

Oosterbeek, Hessel, Randolph Sloof, and Gijs Van De Kuilen, "Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis," *Experimental*

Economics, 2004, *7* (2), 171–188.

Pustejovsky, James E and Elizabeth Tipton, "Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models," *Journal of Business & Economic Statistics*, 2018, *36* (4), 672–683.

Riley, Richard D, "Multivariate Meta-Analysis: the Effect of Ignoring Within-Study Correlation," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2009, *172* (4), 789–811.

_ , KR Abrams, PC Lambert, AJ Sutton, and JR Thompson, "An Evaluation of Bivariate Random-Effects Meta-Analysis for the Joint Synthesis of Two Correlated Outcomes," *Statistics in Medicine*, 2007, *26* (1), 78–97.

Roth, Alvin E and Ido Erev, "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 1995, *8* (1), 164–212.

Stanley, Tom D, "Beyond Publication Bias," *Journal of Economic Surveys*, 2005, *19* (3), 309–345.

_ and Hristos Doucouliagos, *Meta-Regression Analysis in Economics and Business*, routledge, 2012.

_ and _ , "Neither Fixed nor Random: Weighted Least Squares Meta-Regression," *Research Synthesis Methods*, 2017, *8* (1), 19–42.

Teyssier, Sabrina, "Inequity and Risk Aversion in Sequential Public Good Games," *Public Choice*, 2012, *151* (1), 91–119.

Tipton, Elizabeth, "Small Sample Adjustments for Robust Variance Estimation with Meta-Regression.," *Psychological Methods*, 2015, *20* (3), 375.

Tol, Richard SJ, "International Inequity Aversion and the Social Cost of Carbon," *Climate Change Economics*, 2010, *1* (01), 21–32.

Trikalinos, Thomas A, David C Hoaglin, and Christopher H Schmid, "An Empirical Comparison of Univariate and Multivariate Meta-Analyses for Categorical Outcomes," *Statistics in Medicine*, 2014, *33* (9), 1441–1459.

Vogt, Carsten, "Climate Coalition Formation when Players are Heterogeneous and Inequality Averse," *Environmental and Resource Economics*, 2016, *65* (1), 33–59.

Yamagishi, Toshio and Toko Kiyonari, "The Group as the Container of Generalized Reciprocity," *Social Psychology Quarterly*, 2000, pp. 116–132.

Yang, Yang, Sander Onderstal, and Arthur Schram, "Inequity Aversion Revisited," *Journal of Economic Psychology*, 2016, *54*, 1–16.

Zelmer, Jennifer, "Linear Public Goods Experiments: A Meta-Analysis," *Experimental Economics*, 2003, *6* (3), 299–310.

A Articles Included in Dataset (Chronological Order)

1. **Fehr, Ernst, and Klaus M. Schmidt**, "A Theory of Fairness, Competition, and Cooperation," *The Quarterly Journal of Economics*, 1999, 114(3): 817–868.

2. **Goeree, Jacob, and Charles Holt**, "Asymmetric Inequality Aversion and Noisy Behavior in Alternating-Offer Bargaining Games," *European Economic Review*, 2000, 44(4-6): 1079–1089.

3. **Huck, Steffen, Wieland Müller, and Hans-Theo Normann**, "Stackelberg Beats Cournot: On Collusion and Efficiency in Experimental Markets," *The Economic Journal*, 2001, 111(474): 749–765.

4. **Charness, Gary, and Hernan Haruvy**, "Altruism, Equity, and Reciprocity in a Gift-Exchange Experiment: An Encompassing Approach," *Games and Economic Behavior*, 2002, 40(2): 203–231.

5. **Ellingsen, Tore, and Magnus Johannesson**, "Promises, Threats and Fairness," *The Economic Journal*, 2004, 114(495): 397–420.

6. **Dannenberg, Astrid, Bodo Sturm, and Carsten Vogt**, "Do Equity Preferences Matter in Climate Negotiations? An Experimental Investigation," Unpublished Manuscript, 2007.

7. **Bellemare, Charles, Sabine Kröger, and Arthur Van Soest**, "Measuring Inequity Aversion in a Heterogeneous Population Using Experimental Decisions and Subjective Probabilities," *Econometrica*, 2008, 76(4): 815–839.

8. **Dannenberg, Astrid, Bodo Sturm, and Carsten Vogt**, "Do Equity Preferences Matter for Climate Negotiators? An Experimental Investigation," *Environmental and Resource Economics*, 2010, 47(1): 91–109.

9. **Daruvala, Dinky**, "Would the Right Social Preference Model Please Stand Up!," *Journal of Economic Behavior & Organization*, 2010, 73(2): 199–208.

10. **Lau, Sau-Him Paul, and Felix Leung**, "Estimating a Parsimonious Model of Inequality Aversion in Stackelberg Duopoly Experiments," *Oxford Bulletin of Economics and Statistics*, 2010, 72(5): 6690–686.

11. **Bellemare, Charles, Sabine Kröger, and Arthur Van Soest**, "Preferences, Intentions, and Expectation Violations: A Large-Scale Experiment with a Representative Subject Pool," *Journal of Economic Behavior & Organization*, 2011, 78:(3): 349–365.

12. **Blanco, Mariana, Dirk Engelmann, and Hans-Theo Normann**, "A Within-Subject Analysis of Other-Regarding Preferences," *Games and Economic Behavior*, 2011, 72(2): 321–338.

13. **Morishima, Yosuke, Daniel Schunk, Adrian Bruhin, Christian C.Ruff, Ernst Fehr**, "Linking Brain Structure and Activation in Temporoparietal Junction to Explain the Neurobiology of Human Altruism," *Neuron*, 2012, 75(1): 73–79.

14. **Aksoy, Ozan, and Jeroen Weesie**, "Hierarchical Bayesian Analysis of Biased Beliefs and Distributional Other-Regarding Preferences," *Games*, 2013, 4(1): 66–88.

15. **Kleine, Fabian, Manfred Königstein, and Balázs Rozsnyói**, "Voluntary Leadership in an Experimental Trust Game," *Journal of Economic Behavior & Organization*, 2014, 108: 442–452.

16. **Beranek, Benjamin, Robin Cubitt, and Simon Gächter**, "Stated and Revealed Inequality Aversion in Three Subject Pools" *Journal of the Economic Science Association*, 2015, 1: 43–58.

17. **Ponti, Giovanni, Ismael Rodriguez-Lara**, "Social Preferences and Cognitive Reflection: Evidence from a Dictator Game Experiment," *Frontiers in Behavioral Neuroscience*, 2015.

18. **He, Haoran, Keyu Wu**, "Choice Set, Relative Income, and Inequity Aversion: An Experimental Investigation," *Journal of Economic Psychology*, 2016, 54: 177–193.

19. **Cueva, Carlos, Iñigo Iturbe-Ormaetxe, Esther Mata-Pérez, Giovanni Ponti, Marcello Sartarelli, Haihan Yu, and Vita Zhukova**, "Cognitive (Ir)Reflection: New Experimental Vvidence," *Journal of Behavioral and Experimental Economics*, 2016, 64: 81–93.

20. **De Melo, Celso, Stacy Marsella, and Jonathan Gratch**, "People Do Not Feel Guilty About Exploiting Machines," *ACM Transactions on Computer-Human Interaction*, 2016, 23(2): 1–17.

21. **Tasch, Weiwei, and Daniel Houser**, "Social Preferences and Social Curiosity," Unpublished Manuscript, 2018.

22. **Boun My, Kene, Nicolas Lampach, Mathieu Lefebvre, and Jacopo Magnani**, "Effects of Gain-Loss Frames on Advantageous Inequality Aversion," *Journal of the Economic Science Association*, 2018, 4(2): 99–109.

23. **Bruhin, Adrian, Ernst Fehr, Daniel Schunk**, "The Many Faces of Human Sociality: Uncovering the Distribution and Stability of Social Preferences," *Journal of the European Economic Association*, 2019, 17(4): 1025–1069.

24. **Diaz, Lina, Daniel Houser, John Ifcher, and Homa Zarghamee**, "Estimating Social Preferences Using Stated Satisfaction: Novel Support for Inequity Aversion," Unpublished Manuscript, 2021.

25. **Alger, Ingela, and Boris van Leeuwen**, "Estimating Social Preferences and Kantian Morality in Strategic Interactions," Unpublished Manuscript, 2021.

26. **Sabater-Grande, Gerardo, Aurora García-Gallego, Nikolaos Georgantzís, and Noemi Herranz-Zarzoso**, "The Effects of Personality, Risk and Other-Regarding Attitudes on Trust and Reciprocity," *Journal of Behavioral and Experimental Economics*, 2022, 96.

B Variables Coded in Dataset

Table 8: List of Coded Variables in the Dataset

Variable	Description
paper_id	ID for the 26 paper in the analysis (from 1 to 26)
paper_title	Title of the paper
authors	Authors' first and last names
paper_code	First author's last name + et al. + year
is_published	= 1 if the paper is published
year_published	Year published or last revisited if working paper
journal	Journal
paper_length	Length of the paper (appendix excluded)
affiliations	Affiliations of the authors
is_lab	= 1 if laboratory experiment
is_online	= 1 if online experiment
is_classroom	= 1 if classroom experiment
loc_exp_country	Country location of the experiment
loc_exp_contintent	Continent location of the experiment
is_uni	= 1 if university students population
is_adults	= 1 if adults population (not general or in university)
is_general	= 1 if general population
reward_money	= 1 if monetary reward
strategic_alpha	= 1 if α elicited in a strategic game
strategic_beta	= 1 if β elicited in a strategic game
games_alpha	Games used to elicit α
games_beta	Games used to elicit β
game1-game4	All games played in the experiment
utility_function	Utility function specification used
econometric_strategy	Econometric strategy
estimation_method	Estimation method used
alpha	Disadvantageous inequality coefficient (α)
alpha_se	SE of α
alpha_sd	SD of α
beta	Advantageous inequality coefficient (β)
beta_se	SE of β
beta_sd	SD ofβ
type_se	Type of SE (reported, from SD, from reg)
type_sd	Type of SD (reported, computed)

n	Sample size
is_aggregate	= 1 if aggregate estimates
is_individual	= 1 if individual-level estimates
is_mean	= 1 if individual-level mean
is_median	= 1 if individual-level median
is_finite_mix	= 1 if finite-mixture estimates
p1-p4	mixture probabilities if finite-mixture
p1_se-p4_se	SEs of $p_1 - p_4$ if finite-mixture
alpha1-alpha4	Alpha coefficients if finite-mixture
alpha1_se-alpha4_se	SEs of $\alpha_1 - \alpha_4$ if finite-mixture
beta1-beta4	Beta coefficients if finite-mixture
beta1_se-beta4_se	SEs of $\beta_1 - \beta_4$ if finite-mixture
is_other_param	= 1 if other parameters are estimated
other_param	Names of other parameters
other_info	Other information on the paper

C Experimental Tasks Used To Elicit Parameters

Table 9: Experimental Tasks and Classification as Strategic

Experimental Tasks Used To Elicit Parameters	Strategic Environment
Disadvantageous Inequality Coefficient (α)	
Bargaining game	Yes
Choice menus	No
Dictator game	No
Equality equivalence test	No
Gift exchange game	Yes
Inequality list	No
Modified dictator game	No
Non strategic ultimatum game	No
Random ultimatum game	Yes
Sequential prisoner dilemma	Yes
Sequential public good game	Yes
Stackelberg game	Yes
Trust game	Yes
Ultimatum game	Yes
Advantageous Inequality Coefficient (β)	
Bargaining game	Yes
Choice menus	No
Dictator game	No
Equality equivalence test	No
Gift exchange game	Yes
Inequality list	No
Modified dictator game	No
Random ultimatum game	Yes
Sequential prisoner dilemma	Yes
Sequential public good game	Yes
Stackelberg game	Yes
Trust game	Yes
Ultimatum game	Yes

D Meta-Analysis with Bayesian Hierarchical Model

Here, we explain now the modelling framework of the Bayesian hierarchical model. We will use in the examples the variable α, but the same applies also to β. Consider the dataset $(\alpha_j, se_i^2)_{j=1}^k$, where k is the total number of estimates and α_j the jth observation of the disadvantageous inequality parameter, with its associated standard error se_j. We then assume that the reported estimate α_j is distributed normally around the parameter $\bar{\alpha}_j$:

$$\alpha_j|\bar{\alpha}_j, se_j \sim \mathcal{N}(\bar{\alpha}_j, se_j^2)$$

The variability around $\bar{\alpha}_j$ is due to the sampling variation captured by the standard errors se_j. As in a frequentist random-effects model, we can assume that the sampling variation is not the only source of variability for the estimates, since there could be heterogeneity across measurements due to different settings like subject population, games played etc. This can be modeled by assuming that each $\bar{\alpha}_j$ is normally distributed, adding a second layer to the hierarchy:

$$\bar{\alpha}_j|\alpha_0, \tau \sim \mathcal{N}(\alpha_0, \tau^2)$$

where α_0 is the overall mean of the disadvantageous inequality parameters $\bar{\alpha}_j$, and τ^2 represents the genuine variability across studies. Combining the two expression we get:

$$\alpha_j|\alpha_0, \tau, se_j \sim \mathcal{N}(\alpha_0, \tau^2 + se_j^2)$$

with this formulation being identical to the formulation in the random-effects meta-analysis we explained in the Results section:

$$\alpha_j = \bar{\alpha}_j + \epsilon_j = \alpha_0 + \xi_j + \epsilon_j$$

In Bayesian hierarchical models, each observation ,$\bar{\alpha}_j$, is pooled towards the overall mean with strength depending on the precision of the estimate and on how far the estimate is from the α_0. The pooling equation can be written as follows:

$$\bar{\alpha}_j = (1 - \omega_j)\alpha_j + \omega_j\alpha_0$$

where ω_j is the "pooling factor" (Gelman and Pardoe, 2006), defined as:

$$\omega_j = \frac{se_j^2}{\tau^2 + se_j^2}$$

All others things considered, the more an estimate is imprecise, captured by se_j, the more it will be pooled towards the overall mean. The same effect also happens when τ^2 is low, meaning that if there is low heterogeneity across studies, more weight will be given to α_0.

We now summarize and estimate the model expressed above. We estimate the model in Stan (Carpenter et al., 2017) using the Hamiltonian Monte Carlo simulations and launch it from R (https:// www.r-project.org/) using RStan (Stan Development Team, 2021).

The models we fitted for α and β are the following:

$$\begin{aligned}
\alpha_j|\bar{\alpha}_j, se_j &\sim \mathcal{N}(\bar{\alpha}_j, se_j^2) \\
\bar{\alpha}_j|\alpha_0, \tau &\sim \mathcal{N}(\alpha_0, \tau^2) \\
\alpha_0 &\sim \mathcal{N}(0.5, 1) \\
\tau &\sim half\mathcal{N}(0, 4)
\end{aligned}$$

$$\begin{aligned}
\beta_j|\bar{\beta}_j, se_j &\sim \mathcal{N}(\bar{\beta}_j, se_j^2) \\
\bar{\beta}_j|\beta_0, \tau &\sim \mathcal{N}(\beta_0, \tau^2) \\
\beta_0 &\sim \mathcal{N}(0.25, 1) \\
\tau &\sim half\mathcal{N}(0, 1)
\end{aligned}$$

The priors for the population parameters are mildly regularizing, meaning that they are informative but are chosen in such a way to have a weak effect in the procedure. Looking, for example, at the prior for α_0 and by using the three sigma-rule of thumb, what the prior is saying is that our initial opinion for the true value of α_0 is that the parameter lies between -2.75 and 3.25 with 95% probability. The procedure is not sensitive to the priors we use as long as they are weakly informative.

Looking at the results for the disadvantageous inequality parameter, we observe a mean value for α_0 of 0.281, with a 95% credible interval between [0.2, 0.365]. The frequentist random-effects model returns a value for α_0 of 0.279 with a 95% confidence interval between [0.192, 0.366]. As we can see the two values are nearly identical, and the same happens for the estimate of $\hat{\tau}$ with a mean value in the Bayesian procedure of 0.339 and of 0.332 in the frequentist approach.

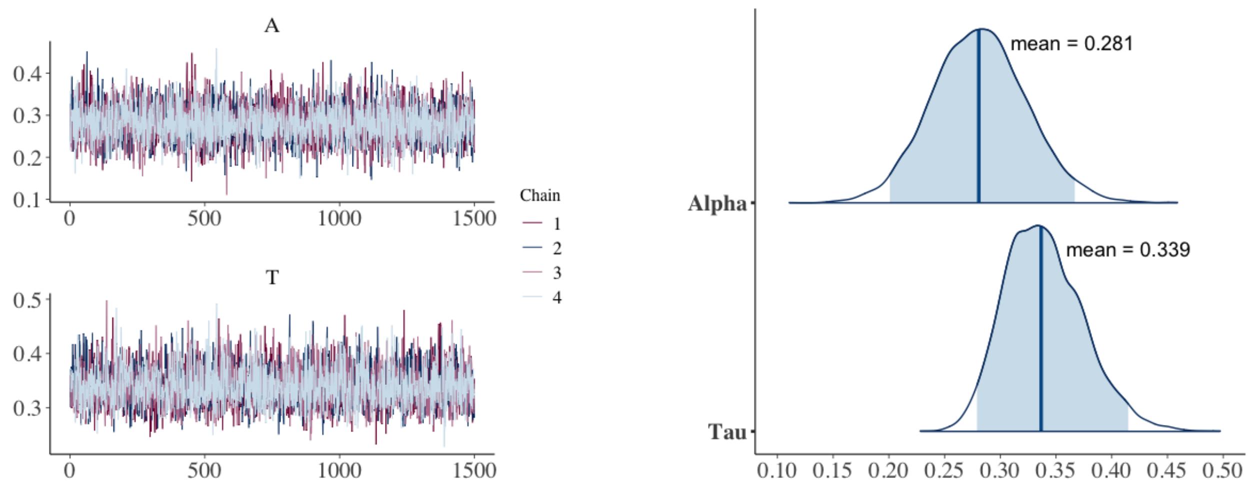


Figure 8: The first figure shows the 1,500 draws for α_0 and τ in the four Markov Chain after the warmup, showing good convergence of the procedure. The second figure shows the posterior distributions of the two parameters. Shaded blue areas correspond to 95% credible intervals.

Table 10: Summary of the Bayesian Hierarchical Model Estimate for α

Parameter	Rhat	ESS	Mean	SD	2.5%	25%	50%	75%	97.5%
α_0	1.000	5088	0.281	0.043	0.2	0.251	0.28	0.309	0.365
$\hat{\tau}$	1.000	3697	0.339	0.035	0.278	0.315	0.337	0.362	0.414

Notes: Rhat is a measure of good convergence of the Markov Chains. As a rule of thumb it should be between 0.9 and 1.05. ESS stands for effective sample size and represents the theoretical number of independent draws. We run four different chains with 3,000 draws each and a warmup of 1,500 draws.

Now looking at the results for the advantageous inequality parameter, we once again observe very similar results between the Bayesian and frequentist methods. The mean value for β_0 is 0.225, while in the random-effects model is 0.226. The same happens for the estimate of $\hat{\tau}$ with a mean value in the Bayesian procedure of 0.238 and of 0.233 in the frequentist approach.

Table 11: Summary of the Bayesian Hierarchical Model Estimate for β

Parameter	Rhat	ESS	Mean	SD	2.5%	25%	50%	75%	97.5%
β_0	1.000	6112	0.225	0.027	0.17	0.208	0.226	0.244	0.277
$\hat{\tau}$	1.000	4612	0.238	0.022	0.199	0.222	0.236	0.251	0.288

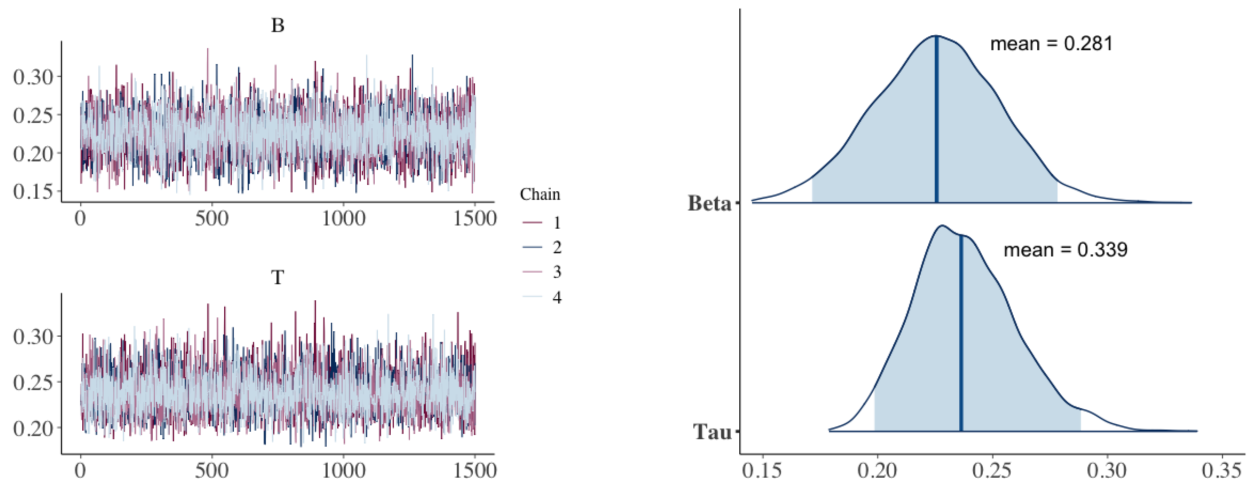


Figure 9: The first figure shows the 1,500 draws in the four Markov Chain after the warmup, showing good convergence of the procedure. The second figure shows the posterior distributions of β_0 and τ. Shaded blue areas correspond to 95% credible intervals.

The model we just estimated does not take into account the possible correlation among estimates that come from the same study. One way to solve this problem is to introduce a paper level in the hierarchical model as follows:

$$\begin{aligned}
\alpha_{pj}|\bar{\alpha}_{pj}, se_{pj} &\sim \mathcal{N}(\bar{\alpha}_{pj}, se_{pj}^2) \\
\bar{\alpha}_{pj}|\bar{\alpha}_{p}, \sigma_p &\sim \mathcal{N}(, \bar{\alpha}_{p}, \sigma_p^2) \\
\bar{\alpha}_{p}|\alpha_0, \tau_s &\sim \mathcal{N}(\alpha_0, \tau_s^2) \\
\alpha_0 &\sim \mathcal{N}(0.25, 1) \\
\tau &\sim half\mathcal{N}(0, 1) \\
\sigma_p &\sim half\mathcal{N}(0, 1)
\end{aligned}$$

$$\begin{aligned}
\beta_{pj}|\bar{\beta}_{pj}, se_{pj} &\sim \mathcal{N}(\bar{\beta}_{pj}, se_{pj}^2) \\
\bar{\beta}_{pj}|\bar{\beta}_{p}, \sigma_p &\sim \mathcal{N}(, \bar{\beta}_{p}, \sigma_p^2) \\
\bar{\beta}_{p}|\beta_0, \tau &\sim \mathcal{N}(\beta_0, \tau^2) \\
\beta_0 &\sim \mathcal{N}(0.25, 1) \\
\tau &\sim half\mathcal{N}(0, 1) \\
\sigma_p &\sim half\mathcal{N}(0, 1)
\end{aligned}$$

where now we introduced paper level means of the parameters in a single study, $\bar{\alpha}_p$. These models for α and β resemble the multi-level frequentist approach discussed in details in the main body of the paper.

The Bayesian procedure returns a mean disadvantageous inequality coefficient of 0.426, with a 95% probability that the true value falls in the interval [0.24, 0.62]. This is in line with what we found in the frequentist analysis, with an estimate for α of 0.425 and a confidence interval of [0.244, 0.606].

Table 12: Summary of the Bayesian Hierarchical Model Estimate for α with paper level

Parameter	Rhat	ESS	Mean	SD	2.5%	25%	50%	75%	97.5%
α_0	1.000	6931	0.426	0.097	0.24	0.362	0.424	0.488	0.62
$\hat{\tau}$	1.000	4560	0.439	0.086	0.298	0.378	0.43	0.488	0.636

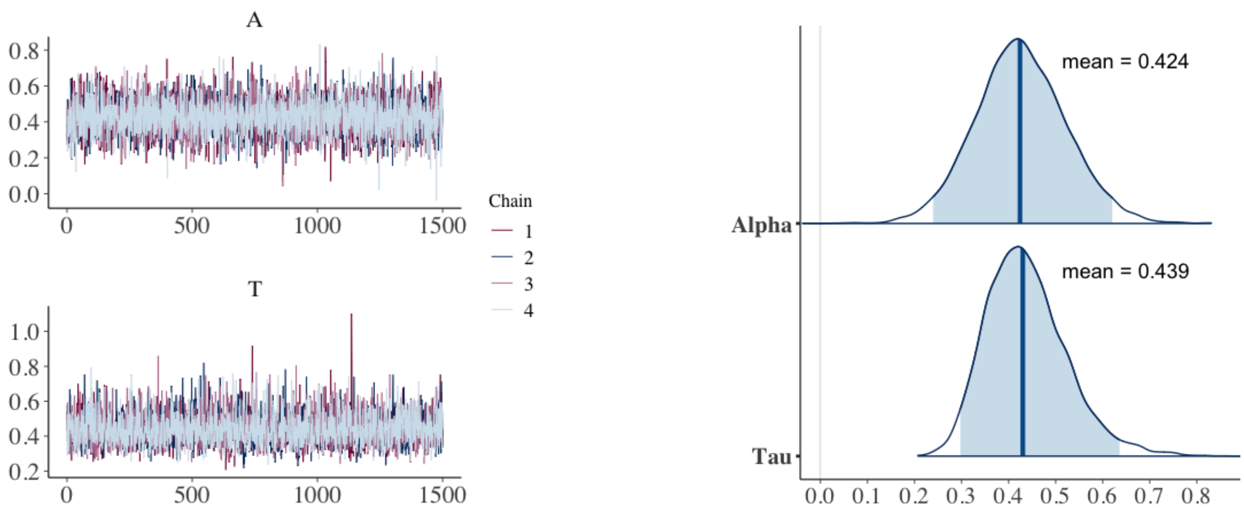


Figure 10: The first figure shows the 1,500 draws in the four Markov Chain after the warmup, showing good convergence of the procedure. The second figure shows the posterior distributions of α_0 and τ. Shaded blue areas correspond to 95% credible intervals.

Now discussing β, the Bayesian procedure returns a mean advantageous inequality coefficient of 0.29, with a 95% probability that the true value falls in the interval [0.212, 0.366]. Once again, this is in line with what we found in the frequentist analysis, with an estimate for β of 0.29 and a confidence interval of [0.218, 0.363].

Table 13: Summary of the Bayesian Hierarchical Model Estimate for β with paper level

Parameter	Rhat	ESS	Mean	SD	2.5%	25%	50%	75%	97.5%
β_0	1.000	7322	0.29	0.039	0.212	0.263	0.289	0.315	0.366
$\hat{\tau}$	1.000	5504	0.164	0.032	0.113	0.143	0.16	0.183	0.237

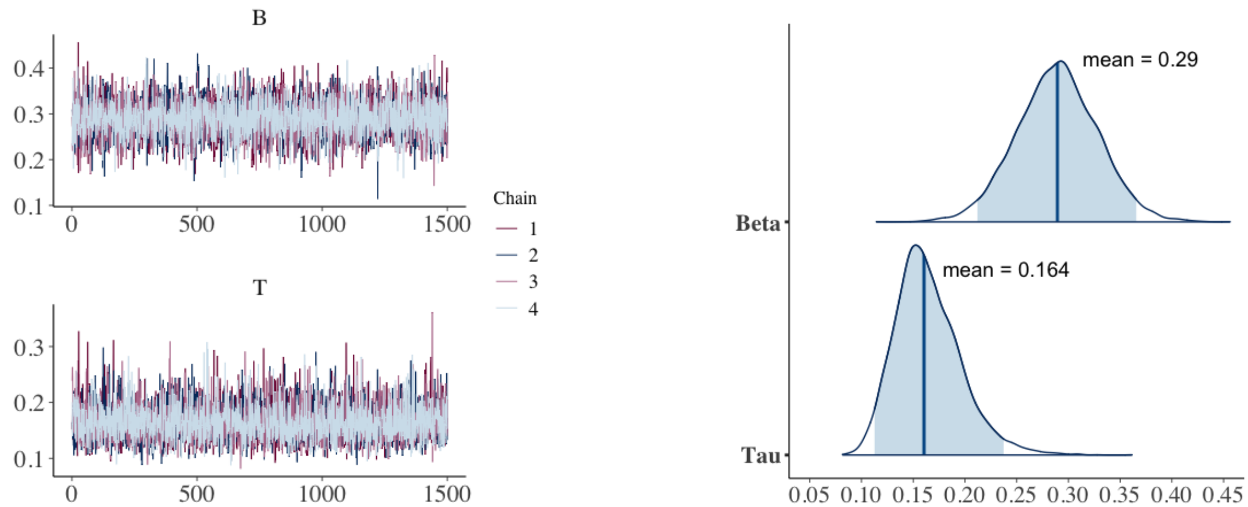


Figure 11: The first figure shows the 1,500 draws in the four Markov Chain after the warmup, showing good convergence of the procedure. The second figure shows the posterior distributions of α_0 and τ. Shaded blue areas correspond to 95% credible intervals.

E Multivariate Meta-Analysis

The standard approach when doing meta-analysis of studies that report multiple effects sizes is to consider each effect size independent of the others and conduct univariate analysis, one for each effect size. Univariate meta-analysis are simple to implement and interpret, but this approach completely disregards possible within-study and between-study outcome correlations that can have a potentially relevant effect on the estimates and their SEs.

The alternative approach is to implement a multivariate meta-analysis by explicitly modelling outcome correlations. While multivariate models are theoretically the first-best, since they can always nest univariate models, they are more difficult and time-consuming to estimate. Moreover, some studies (Trikalinos et al., 2014; Berkey et al., 1998; Ishak et al., 2008) find little to no effect on the parameter estimates between univariate and multivariate meta-analysis, thus supporting the idea of simply using the easier univariate model. Other studies (Riley et al., 2007; Kirkham et al., 2012) find instead a difference between univariate and multivariate estimates, and they argue that a multivariate approach is the correct procedure when dealing with multiple effect sizes in the same study.

Another problem in conducting a multivariate meta-analysis is the need to not only have a measure of the effect sizes and their SEs, but also of their correlation (or covariance), and this information is often not reported. Ishak et al. (2008) suggest that the correlation can be ignored without too much risk of introducing a bias in the analysis, but Riley (2009) finds that this was not true in the studies he analyzed. Nonetheless this is the approach we take in this paper since we do not have in our dataset a measure of the correlation for α and β.

The specification for the multivariate random-effects model applied in our dataset of inequality aversion estimates is the following:

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \sim \mathcal{N}\left\{ \begin{pmatrix} \mu_j^{\alpha} \\ \mu_j^{\beta} \end{pmatrix}, R_j \right\}, \qquad R_j = \begin{bmatrix} SE_{aj}^2 & 0 \\ 0 & SE_{bj}^2 \end{bmatrix}$$

$$\begin{pmatrix} \mu_j^{\alpha} \\ \mu_j^{\beta} \end{pmatrix} \sim \mathcal{N}\left\{ \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}, D \right\}, \qquad D = \begin{bmatrix} D_a^2 & D_a D_b \rho_D \\ D_a D_b \rho_D & D_b^2 \end{bmatrix}$$

Where similarly to the univariate model, we assume that the observed parameters (α_j, β_j) are distributed around the true effect sizes $(\mu_j^{\alpha}, \mu_j^{\beta})$, with known variance-covariance matrix R_j. The diagonal elements are the variance for α and β which are known, while we assumed zero covariance to be able to estimate the model. The true effect sizes are then distributed as a bivariate normal with means (α_0, β_0) and variance-covariance matrix D.

To handle statistically dependent estimates we can add another level to the hierarchy to capture both within-study and between-study heterogeneity, thus getting a multivariate and multi-level specification:

$$\begin{pmatrix} \alpha_{ij} \\ \beta_{ij} \end{pmatrix} \sim \mathcal{N}\left\{ \begin{pmatrix} \mu_{ij}^{\alpha} \\ \mu_{ij}^{\beta} \end{pmatrix}, R_j \right\}, \qquad R_j = \begin{bmatrix} SE_{aij}^2 & 0 \\ 0 & SE_{bij}^2 \end{bmatrix}$$

$$\begin{pmatrix} \mu_{ij}^{\alpha} \\ \mu_{ij}^{\beta} \end{pmatrix} \sim \mathcal{N}\left\{ \begin{pmatrix} \theta_{i}^{\alpha} \\ \theta_{i}^{\beta} \end{pmatrix}, C_i \right\}, \qquad C_i = \begin{bmatrix} C_{aij}^2 & C_{aij}C_{bij}\rho_C \\ C_{aij}C_{bij}\rho_C & C_{bij}^2 \end{bmatrix}$$

$$\begin{pmatrix} \theta_{i}^{\alpha} \\ \theta_{i}^{\beta} \end{pmatrix} \sim \mathcal{N}\left\{ \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}, D \right\}, \qquad D = \begin{bmatrix} D_a^2 & D_a D_b \rho_D \\ D_a D_b \rho_D & D_b^2 \end{bmatrix}$$

Where the observed parameters $(\alpha_{ij}, \beta_{ij})$ are distributed around the true effect sizes $(\mu_{ij}^{\alpha}, \mu_{ij}^{\beta})$, the true effect sizes around paper-level means $(\theta_i^{\alpha}, \theta_i^{\beta})$ and the latter around the population means (α_0, β_0). In this multivariate multi-level model we are estimating in addition to the variance of the within and between study errors for α and β, also their correlation/covariance.

We report the results of the multivariate random-effects and multivariate multi-level random-effects models in both the full and restricted sample in Table 14. Looking at the latter model we observe an estimate of the average disadvantageous inequality parameter equals to 0.425, which is the same as the one obtained in the univariate specification. The average advantageous inequality parameter is instead estimated to be equal to 0.286, slightly lower than the 0.291 found in the univariate case. Also SEs are practically identical.

Estimating both parameters at the same time allows us to correctly test the null hypothesis of $\alpha_0 - \beta_0 = 0$. The t-test statistic and its p-value in the multivariate multi-level specification confirm that the two parameters are not statistically different from zero (p-value=0.134).

Table 14: Meta-Analytic Average of Disadvantageous Inequality (α) and Advantageous Inequality (β)

	RE Full	RE Restricted	ML Full	ML Restricted
Disadvantageous Inequality (α_0)	0.278	0.262	0.425	0.445
	(0.039)	(0.048)	(0.09)	(0.127)
Advantageous Inequality (β_0)	0.218	0.202	0.286	0.282
	(0.027)	(0.032)	(0.037)	(0.048)
p-value α_0	< 0.0001	$< .0001$	< 0.0001	0.0004
p-value β_0	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Observations	81	61	81	61

Notes: The first and the third columns estimate a multivariate random-effects and multivariate multi-level random-effects model on the full sample. The columns in even positions consider only the observations with reported SEs. p-values are from the two sided test of the null hypothesis $H_0 : \alpha_0 = 0$ or $\beta_0 = 0$. In both random effects and multi-level models the restricted maximum likelihood method is used.