# Meta-Analysis of Distributional Preferences Estimates*

Salvatore Nunnari† Massimiliano Pozzi‡

December 2, 2022

**Abstract**

We conduct a large-scale interdisciplinary meta-analysis to aggregate the knowledge from empirical estimates of distributional preferences reported from 1999 to 2022. In particular, we examine 289 estimates of sensitivity to inequality from 40 articles in economics, psychology, neuroscience and computer science which structurally estimate the Fehr and Schmidt (1999) model. Our analysis indicates that individuals are inequality averse: mean sensitivity to disadvantageous inequality is 0.467 with a 95% probability that the true value lies in the interval $[0.302, 0.642]$; mean sensitivity to advantageous inequality is 0.331 with a 95% probability that the true value lies in the interval $[0.266, 0.396]$. We observe high levels of heterogeneity, both across studies and across individuals, with estimated parameters sensitive to some features of the study design (namely, the experimental task and the subject population). Finally, we do not find compelling evidence of selective reporting or publication bias.

**Keywords:** Other-Regarding Preferences, Altruism, Envy, Guilt, Inequality Aversion, Meta-Analysis, Multi-Level Random-Effects Model, Bayesian Hierarchical Model

**JEL Codes:** C90, C11, D63, D91

*A previous version of this paper was circulated under the title "Meta-Analysis of Inequality Aversion Estimates." We are grateful to Jeffrey Carpenter, Lina Diaz, Daniel Houser, John Ifcher, Stephan Mueller, Sander Onderstal, Holger Rau, Andrea Robbett, Arthur Schram, Yang Yang, and Homa Zarghamee for providing additional details on the estimates in their papers. We also thank Colin Camerer, David Cooper, Glenn Dutcher, Catherine Eckel, Ernst Fehr, Guillaume Frechette, Taisuke Imai, Daniel Mueller, Ferdinand Vieider and the audience of the 2022 CESifo Area Conference on Behavioral Economics for helpful comments. Nunnari acknowledges financial support from the European Research Council (Grant No. 852526).

†Bocconi University, Department of Economics, salvatore.nunnari@unibocconi.it

‡Bocconi University, BELSS, pozzi.massimiliano@unibocconi.t

# 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 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 comparison

[1]As of 27 September 2022, FS has 14,196 citations on Google Scholar and 5,064 citations on Web of Science.

with the other agent's payoff, $x_j$. Assuming that $\alpha > \beta > 0$ (as in FS), this is a model of *inequity aversion* (where $\alpha$ can be interpreted as *envy* and $\beta$ as *guilt*), since differences in payoffs cause disutility for agent $i$. At the same time, this simple framework can capture other 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*. 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 $\alpha$ and $\beta$ 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 $\alpha$ 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, $\beta$ was assumed to take three different values — 0, 0.25 and 0.6 — with calibrated shares of, respectively, 30%, 30% and 40%. More recently, Eckel and Gintis (2010) reviewed the mean parameters estimated in four studies other than FS and reported values ranging between 0.31 and 1.89 for $\alpha$, and between $-0.27$ and 0.80 for $\beta$. Blanco, Engelmann and Normann (2011), instead, estimated the coefficients at the individual level using ultimatum and dictator games and reported average estimates of 1.18 for $\alpha$ and 0.47 for $\beta$. 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 other-regarding preferences accumulated in over 20 years of research with the method of meta-analysis, that

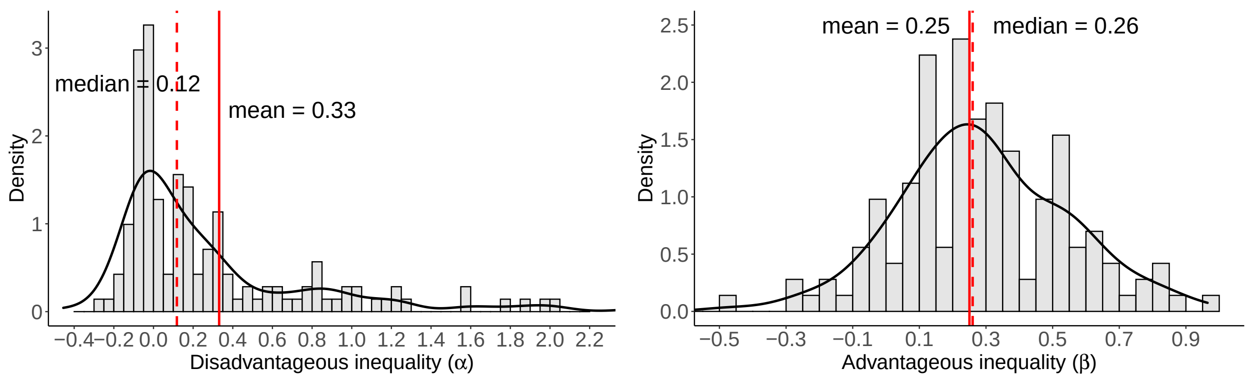


Figure 1: Distribution of Disadvantageous ($\alpha$) and Advantageous Inequality ($\beta$) Coefficients. Notes: Bins for histograms are 0.05 wide; the Gaussian kernel density (solid black line) uses the Silverman's rule of thumb for bandwidth selection; in the panel for $\beta$, the horizontal axis is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

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]

In particular, we collect 145 estimates of sensitivity to disadvantageous inequality and 144 estimates of sensitivity to advantageous inequality from 40 articles in economics, psychology, neuroscience and computer science which 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 $\alpha$ and $\beta$? Second, how do $\alpha$ and $\beta$ vary depending on the characteristics of a study (e.g., the experimental task and the subject population)? 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 $\alpha$ are, respectively, 0.33 and 0.12 with around 40% of the estimates (57 out of 145) equal to or less than 0 (in contrast with the assumption in FS). The raw mean and

[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 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.

median estimates of $\beta$ are, respectively, 0.25 and 0.26, and, again, a sizeable number of observations which do not match the assumption in FS ($\beta \leq 0$ in 21 out of 144 estimates). Focusing on studies which estimate both parameters, disadvantageous inequality matters more than advantageous inequality only around 40% of the time (in 57 out of 140 pairs of estimates) and the correlation between the two parameters is indistinguishable from 0. In the non-parametric analysis, all estimates are given equal weight (even if the parameters computed in some studies are more reliable than others) and are 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 $\alpha$ and $\beta$ 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: according to the multi-level (Bayesian hierarchical) model, the meta-synthetic average for the disadvantageous inequality coefficient is 0.469 (0.467) while the meta-synthetic average for the advantageous inequality coefficient is 0.331 (0.331) 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 $\alpha$ 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 $\beta$; 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 (Denmark, Germany, Netherlands, Sweden, Switzerland and UK).

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 and investigate the incidence of p-hacking, we use funnel plots, the Funnel Asymmetry Testing and Precision

Effect Testing (FAT-PET) procedure (Stanley and Doucouliagos, 2012, 2017), histograms of z-statistics and the p-curve (Simonsohn, Nelson and Simmons, 2014). On the one hand, funnel plots highlight the absence of studies estimating (large in magnitude and imprecisely estimated) negative values of $\alpha$ and positive values of $\beta$ and this is confirmed by the FAT-PET procedure. Moreover, we observe a jump around the threshold for statistical significance in the histograms of z-statistics for both parameters, which is a hint of p-hacking. On the other hand, 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 $\alpha$ and $\beta$ — and the publication-bias corrected meta-synthetic averages of the two parameters are still positive and strongly statistically significant (0.340 for $\alpha$ and 0.400 for $\beta$). In addition, the p-curves for both $\alpha$ and $\beta$ are highly right-skewed which strongly supports the hypothesis that both parameters are different from zero and that researchers did not engage in p-hacking. We, thus, conclude that there is no compelling evidence of selective reporting or publication bias.

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 (2021) on time preferences, Brown et al. (2021) on loss aversion, and Meager (2019, 2022) on the effect of microcredit. To the best of our knowledge, this is the first work that uses meta-analysis techniques to summarize empirical estimates of other-regarding preferences. 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

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

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 other-regarding preferences 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 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., sensitivity to disadvantageous inequality); and the third term measures the disutility from being ahead in pairwise comparisons (i.e., sensitivity to advantageous inequality).

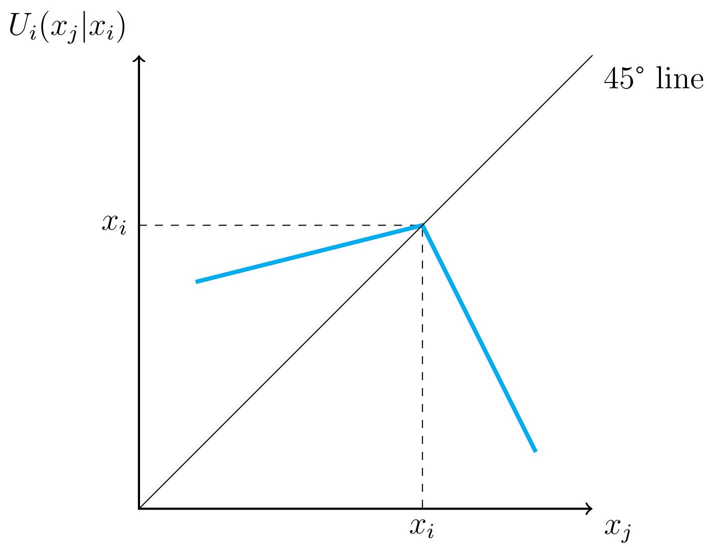


Figure 2: Utility of Inequality Averse Player $i$ in Game with 2 Players ($\alpha = 2$, $\beta = 0.5$).

We briefly discuss the assumptions made in the original contribution by FS. First, FS assume $\alpha > 0$ and $\beta > 0$, making this 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). FS further assume that $\alpha \geq \beta$. This assumption implies that disadvantageous inequality hurts more than advantageous inequality and is inspired by earlier work in behavioral and experimental economics (Kahneman and Tversky, 1979; Loewenstein, Thompson and Bazerman, 1989). Finally, FS constrain $\beta$ to be smaller than 1 in order 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. As discussed in the Introduction, while this is interpreted as a model of inequality aversion when $\alpha > 0$ and $\beta > 0$, this parsimonious framework can be used to model different kinds of other-regarding preferences. Our meta-analysis will reveal which type of other-regarding preferences is more common in the populations that have been sampled in 20 years of social sciences experiments.

Most studies in our dataset estimate $\alpha$ and $\beta$ 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 sensitivity to advantageous inequality to be nearly linear, while the sensitivity to 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 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, $\alpha$ and $\beta$ are inequality sensitivity parameters while $\theta$ and $\delta$ 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:

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

[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.

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

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 sensitivity 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

## 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 sensitivity to disadvantageous inequality, $\alpha$, and/or advantageous inequality, $\beta$, using the model by Fehr and Schmidt (1999)".[6]

The search procedure followed four 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 $\alpha$ and $\beta$. 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 databases Web of Science (February 8, 2022), Google Scholar (February 8, 2022) and Scopus (7 September 2022) 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 returned 1,916 articles. We then read these articles and excluded papers that were clearly irrelevant for our analysis — for example, articles that measured other-

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

(("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, Google Scholar and Scopus

regarding preferences in animals or studies that, while reporting the results of dictator and ultimatum games, did not estimate the parameters of interest. Finally, we read through the remaining articles and applied our inclusion criteria. The final dataset consists of 40 articles and the complete list is available in Appendix A.[7]

## 3.2 Data Construction

After identifying the relevant articles, we assembled the dataset for the meta-analysis by coding the estimates for $\alpha$ and $\beta$, 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 sensitivity to advantageous and disadvantageous inequality. In our 40 articles, these estimates take four forms: (i) *aggregate*, where a single value for $\alpha$ and $\beta$ is estimated for the pooled data of all subjects in the study; (ii) *finite-mixture*, where a finite number of values for $\alpha$ and $\beta$ alongside their distributions are estimated from the pooled data of all subjects; (iii) *individual-level mean*, where $\alpha$ and $\beta$ 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.[8] For the finite-mixture

[7]When a precise measure of the estimated parameters was not available (e.g., because the article reported only a scatter plot or a bar chart of individual-level estimates), we contacted the authors to get additional details. This procedure led us to exclude a single study which computes individual-level estimates for $\alpha$ and $\beta$ but reports only a bubble plot of these estimates (Teyssier, 2012). While it would be possible to recover an imprecise mean or median for the estimates in this study, given the high level of arbitrariness this exercise would entail (for example, in evaluating the exact location of bubbles in the graph and their relative size), we decided not to include this paper in the dataset.

[8]The 2 estimates from Corgnet, Espín and Hernán-González (2015) and 2 out of 4 estimates from Hedegaard, Kerschbamer, Müller and Tyran (2021) are an exception: they report set-valued individual-level

estimates, we computed and coded a weighted average for each parameter.[9]

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 289 estimates in our dataset, the source reported the SEs for 81 estimates and, in other 130 cases, we were able to compute the SEs using the reported standard deviation and sample size. For the remaining 78 estimates, we did not have (direct or indirect) information about the SEs.[10]

We had two options: either drop the 78 estimates without SEs or approximate the SEs 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 difference in the distribution of $\alpha$ and $\beta$ 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.[11] 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,

estimates and the frequency of individuals in each set. In this case, we identity the interval where the median individual is located and we approximate the median value of the parameter with the median point of this interval. For example, consider an hypothetical study which estimates 6 participants have $\alpha \in [0, 0.2)$, 4 participants have $\alpha \in [0.2, 0.4)$, and 4 participants have $\alpha \in [0.4, 0.8]$. In this case, the median individual has $\alpha \in [0.2, 0.4)$ and we approximate the median individual-level estimate with 0.3.

[9]For example, consider one of the finite-mixture estimates of $\alpha$ 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 $\mathrm{SD} = \sqrt{\sum_i p_i(\alpha_i - \widehat{\alpha})^2}$; second, we compute the standard error as $\mathrm{SD}/\sqrt{n}$, where $n$ is the sample size. This procedure disregards the estimated uncertainty of each $\alpha_i$ and the associated $p_i$ but it greatly simplifies our analysis and it is similar to the procedure used by studies that report an individual-level mean.

[10]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.

[11]The two distributions of $\beta$ are statistically different according to a Wilcoxon rank sum test.

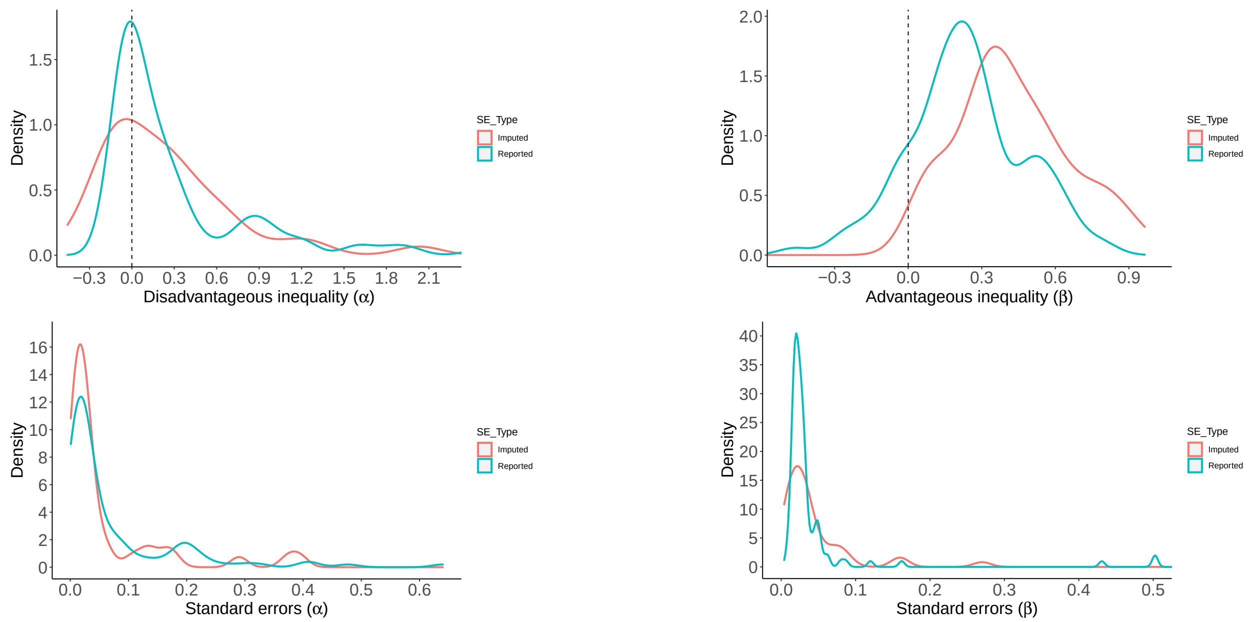


Figure 4: Distribution of Estimates and SEs for $\alpha$ and $\beta$ as Function of SE Type. Note: The top two graphs show kernel density estimates (Gaussian with Silverman's rule of thumb) for the subsets of parameters with reported vs. imputed SEs; the bottom two graphs show kernel density estimates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

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:[12]

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

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

[12]For studies estimating a single parameter, we use only this estimate (and a constant) as regressor.

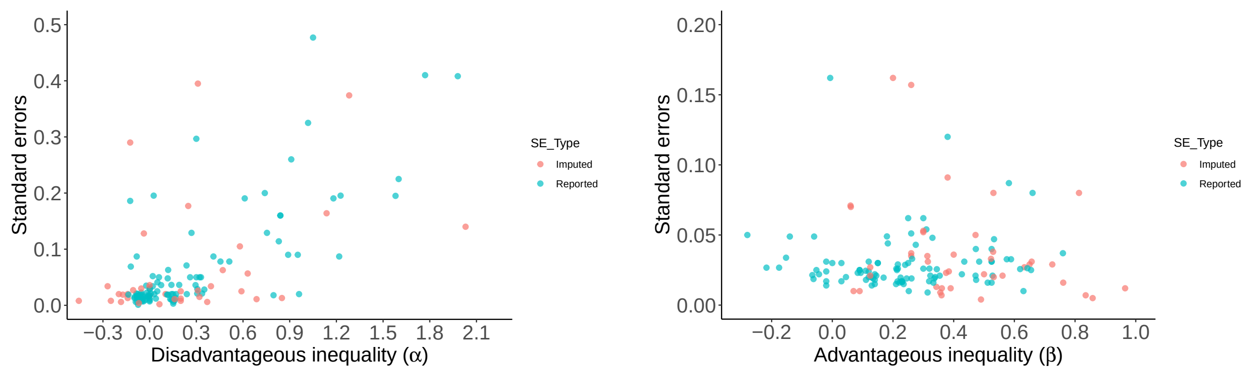


Figure 5: Scatter Plots of $\alpha$ and $\beta$ SEs as a Function of SE Type. Note: The x-axis in the plot for $\beta$ is truncated at $-0.3$ for better visual rendering.

The two parameters explain 43% of the variance in the SEs for $\alpha$ and 25% of the variance in the SEs for $\beta$. Our approximation is, thus, better for $\alpha$ than for $\beta$.

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 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 40 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 42 studies (discussed in 40 papers). These studies report estimates for 149 models of social preferences à la FS: 140 models estimate both the advantageous and disadvantageous inequality parameters, 5 models estimate only $\alpha$, and 4 models estimate only $\beta$.

Table 1 reports the coded features of the 42 studies in our dataset. Among the 42 studies, 37 were presented in papers published (as of 07 September 2022) in economics, psychology, neuroscience and computer science journals. The majority of these 42 studies conducted traditional in-person laboratory experiments, while 7 studies conducted experiments online.[13] The studies were conducted in 11 different countries (China, Denmark, France, Germany, Italy, Netherlands, Spain, Sweden, Switzerland, Turkey, UK, and US) and involved mostly college students (33 studies out of 42), with 4 studies recruiting a sample representative of the Danish, Dutch or German general population (Bellemare, Kröger and van Soest, 2008, 2011; Kerschbamer and Müller, 2020; Hedegaard, Kerschbamer, Müller and Tyran, 2021), and 6 studies recruiting a non-representative sample of adults (Dannenberg, Sturm and Vogt, 2010; Beranek, Cubitt and Gächter, 2015; Sáez, Zhu, Set, Kayser and Hsu, 2015; He and Wu, 2016; Hu, He, Zhang, Wölk, Dreher and Weber, 2018; Carpenter and Robbett, 2022). All studies offered monetary rewards for participating in the experiments.

Table 2 reports the coded features of the 289 estimates in our dataset. Around 55% of the estimates come from studies that compute individual-level estimates of $\alpha$ and $\beta$ and then report the mean and/or the median; around 20% come from four studies which use finite-mixture models (Bruhin, Fehr and Schunk, 2019; Alger and van Leeuwen, 2021; Hedegaard, Kerschbamer, Müller and Tyran, 2021; Carpenter and Robbett, 2022); and around 25% come from studies which estimate parameters for a "representative" agent by pooling together all the available data. Around 75% of the estimates are computed assuming the original utility function specification from FS; around 13% and 10% estimates are computed assuming the

[13]One study recruited participants from mTurk, one from Prolific, two using CentERpanel (an online survey consisting of a representative sample of the adult Dutch population), one using the German Internet Panel (an online survey consisting of a representative sample of the adult German population), one using the internet Laboratory for Experimental Economics (iLEE) at the University of Copenhagen (with subjects selected to be a random sample from the general Danish population), and one contacting climate negotiators from the Intergovernmental Panel on Climate Change directly via email.

Table 1: Features of the Studies ($N = 42$) in the Dataset

| | Frequency | Percentage |
|---|---|---|
| **Publication Status** | | |
| Published (as of September 7, 2022) | 37 | 0.88 |
| Unpublished | 5 | 0.12 |
| **Methodology** | | |
| Laboratory Experiment | 33 | 0.79 |
| Classroom Experiment | 1 | 0.02 |
| Online Experiment | 7 | 0.17 |
| Multiple Methodologies | 1 | 0.02 |
| **Geographic Location** | | |
| United States | 10 | 0.24 |
| Northern Europe (CH, DE, DK, NL, SE, UK) | 20 | 0.48 |
| Southern Europe (FR, IT, ES, TR) | 6 | 0.14 |
| China | 3 | 0.07 |
| Multiple or Unspecified Locations | 3 | 0.07 |
| **Subject Population** | | |
| College Students | 31 | 0.74 |
| Non-Representative Sample of Adults | 6 | 0.14 |
| Representative Sample (of DE, DK, or NL) | 4 | 0.10 |
| Multiple Populations | 1 | 0.02 |
| **Experimental Task Used To Estimate $\alpha$** | | |
| Standard Dictator Game | 3 | 0.07 |
| Mini Dictator Game | 2 | 0.04 |
| Mini Dictator Game with Equality-Efficiency Trade-Off | 18 | 0.39 |
| Ultimatum Game | 12 | 0.26 |
| Other Game | 11 | 0.24 |
| **Experimental Task Used To Estimate $\beta$** | | |
| Standard Dictator Game | 3 | 0.06 |
| Mini Dictator Game | 2 | 0.04 |
| Mini Dictator Game with Equality-Efficiency Trade-Off | 26 | 0.56 |
| Ultimatum Game | 5 | 0.11 |
| Other Game | 11 | 0.23 |
| **Reward Type** | | |
| Money | 42 | 1.00 |

Note: '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').

model in FS is augmented with, respectively, reciprocity parameters or Kantian morality; and the remaining 3% of estimates use the baseline model in FS plus intentions, non-linearity or

Table 2: Features of the Estimates ($N = 289$) in the Dataset.

| | $\alpha$ ($N = 145$) | | $\beta$ ($N = 144$) | |
|---|---|---|---|---|
| | Frequency | Proportion | Frequency | Proportion |
| **Utility Function in Estimated Model** | | | | |
| Linear FS | 107 | 0.74 | 104 | 0.73 |
| Non-Linear FS | 2 | 0.01 | 2 | 0.01 |
| Linear FS + Reciprocity | 19 | 0.14 | 19 | 0.13 |
| Linear FS + Kantian Morality | 15 | 0.10 | 15 | 0.11 |
| Linear FS + Intentions | 2 | 0.01 | 2 | 0.01 |
| Linear FS + Loss Aversion | 0 | 0.00 | 2 | 0.01 |
| **Type of Estimates** | | | | |
| Aggregate | 38 | 0.26 | 37 | 0.25 |
| Finite Mixture | 24 | 0.17 | 24 | 0.17 |
| Individual Mean | 60 | 0.41 | 60 | 0.42 |
| Individual Median | 23 | 0.16 | 23 | 0.16 |
| **Standard Errors** | | | | |
| Reported | 108 | 0.74 | 103 | 0.72 |
| Imputed | 37 | 0.26 | 41 | 0.28 |

loss aversion. The parameters are elicited using choice data from a variety of games. However, even if some studies do 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 145 estimates of $\alpha$ and 144 estimates of $\beta$ 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 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,

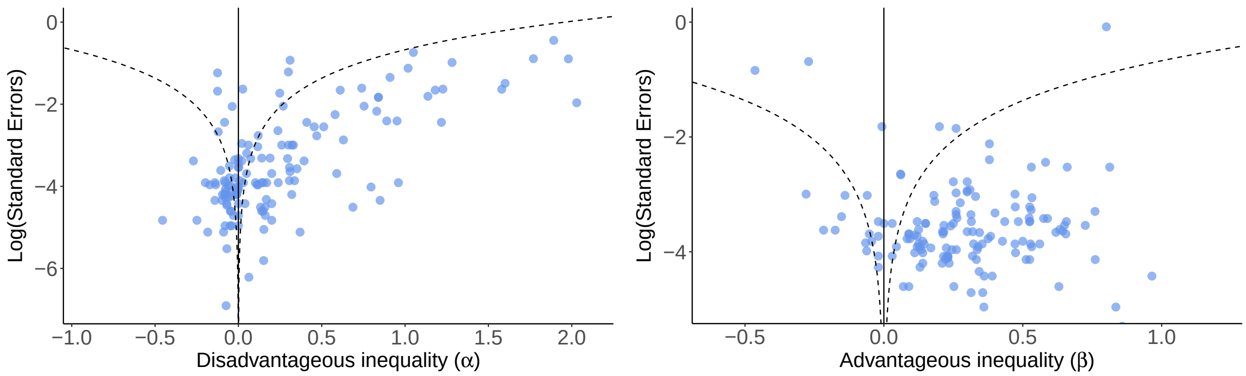


Figure 6: Funnel Plots of Disadvantageous ($\alpha$) and Advantageous Inequality ($\beta$) Coefficients. Notes: The vertical continuous line is at 0 and the diagonal dotted lines represent a p-value of 0.05 for a z-test whose null hypothesis is that the estimate is equal to 0 (i.e., estimates below each dotted line are statistically different from 0).

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 145 estimates of $\alpha$ and of the 144 estimates of $\beta$ in our dataset.[14] The raw mean and median for $\alpha$ are, respectively, 0.33 and 0.12. In contrast with the assumption in FS ($\alpha > 0$), around a third of the estimate (57 out of 145) are equal to or less than 0.[15] This suggests that some individuals are not hurt by unfavorable comparisons with others' outcomes. Table 3 shows that the estimates of $\alpha$ 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).[16] In the former case, the mean and the median of $\alpha$ are, respectively, 0.63 and 0.30; in the

[14]Boxplots of the estimates reported in each paper can be found in Appendix H.

[15]In particular, 88 estimates are greater than 0, 8 estimates are equal to 0 and 49 estimates are smaller than 0. As shown in the left-hand panel of Figure 6, a z-test reveals that 76 estimates are positive and significantly (i.e., p-value $< 0.05$) different from 0, 33 estimates are indistinguishable from 0 and 36 estimates are negative and significantly different from 0.

[16]The full list of games used in the 42 studies from our dataset and whether they are considered strategic environments or individual decision-making tasks can be found in Table 10 in the Appendix.

Table 3: Summary Statistics for Disadvantageous Inequality ($\alpha$)

| | N | Min | 1st Q | 2nd Q | Mean | 3rd Q | Max | SD |
|---|---|---|---|---|---|---|---|---|
| **Estimate Type** | | | | | | | | |
| Aggregate | 38 | −0.14 | −0.07 | 0.12 | 0.24 | 0.28 | 1.89 | 0.44 |
| Finite Mixture | 24 | −0.09 | −0.05 | 0.00 | 0.05 | 0.15 | 0.35 | 0.13 |
| Individual Mean | 60 | −0.46 | −0.03 | 0.30 | 0.50 | 0.84 | 2.80 | 0.72 |
| Individual Median | 23 | −0.13 | 0.00 | 0.03 | 0.33 | 0.33 | 4.50 | 0.94 |
| **Experimental Task** | | | | | | | | |
| Game | 60 | −0.14 | 0.14 | 0.30 | 0.63 | 0.87 | 4.50 | 0.84 |
| Individual Choice | 85 | −0.46 | −0.07 | −0.07 | 0.12 | 0.20 | 1.60 | 0.36 |
| **Complete Dataset** | 145 | −0.46 | −0.05 | 0.12 | 0.33 | 0.39 | 4.50 | 0.65 |

Table 4: Summary Statistics for Advantageous Inequality ($\beta$)

| | N | Min | 1st Q | 2nd Q | Mean | 3rd Q | Max | SD |
|---|---|---|---|---|---|---|---|---|
| **Estimate Type** | | | | | | | | |
| Aggregate | 37 | −0.46 | 0.11 | 0.26 | 0.28 | 0.49 | 0.80 | 0.29 |
| Finite Mixture | 24 | −0.22 | −0.05 | 0.18 | 0.11 | 0.23 | 0.23 | 0.17 |
| Individual Mean | 60 | −2.12 | 0.18 | 0.31 | 0.27 | 0.5 | 0.97 | 0.48 |
| Individual Median | 23 | −0.14 | 0.05 | 0.32 | 0.27 | 0.53 | 0.58 | 0.23 |
| **Experimental Task** | | | | | | | | |
| Game | 40 | −1.27 | −0.06 | 0.12 | 0.11 | 0.31 | 0.80 | 0.33 |
| Individual Choice | 104 | −2.12 | 0.19 | 0.31 | 0.31 | 0.53 | 0.97 | 0.36 |
| **Complete Dataset** | 144 | −2.12 | 0.12 | 0.26 | 0.25 | 0.47 | 0.97 | 0.36 |

latter case, instead, the mean is 0.12 and the median is −0.07. This result is in line with the discussion in Dannenberg, Riechmann, 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 $\alpha$. The significant difference observed in our dataset supports the hypothesis that, in strategic environments, $\alpha$ captures both equity and reciprocity concerns.

The estimates of $\beta$ feature a bell-shaped distribution with a fatter left tail: the raw mean and median are, respectively, 0.25 and 0.26. While there are no estimates greater than 1 (as assumed in FS), around a seventh of the estimates (21 out of 144) are less than 0 (in contrast with the assumption in FS).[17] This suggests that some individuals have "competitive" or

[17] In particular, 123 estimates are greater than 0, 1 estimate is equal to 0 and 20 estimates are smaller

"spiteful" preferences, so that they strictly prefer reducing other earnings (while keeping their own earnings unchanged). As shown in Table 4, contrary to $\alpha$, estimates of $\beta$ computed 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 7 shows a scatter plot of all 140 estimates for which we have a value for both $\alpha$ and $\beta$. We highlight two features of the joint distribution. First, a large number of observations (83 out of 140) 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 ($\rho = 0.04$; $p = 0.63$). This is in line with the results discussed in Dannenberg, Riechmann, 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.

## 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

than 0. As shown in the right-hand panel of Figure 6, a z-test reveals that 116 estimates are positive and significantly (i.e., p-value $< 0.05$) different from 0, 16 estimates are indistinguishable from 0 and 12 estimates are negative and significantly different from 0.

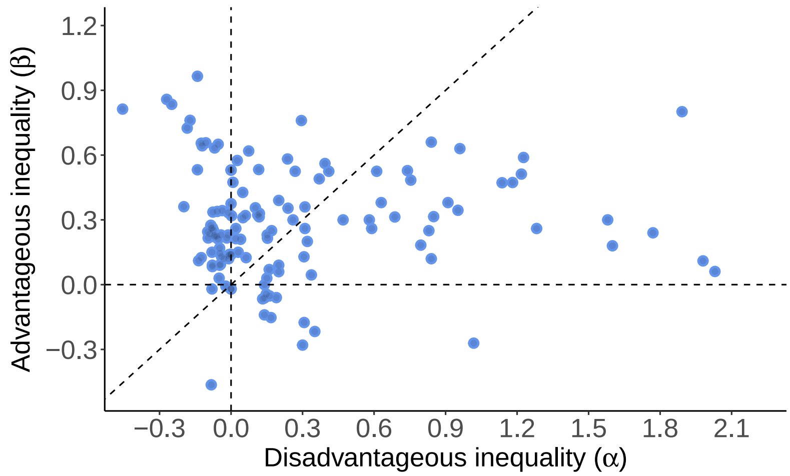


Figure 7: Scatter Plot of Disadvantageous ($\alpha$) and Advantageous Inequality ($\beta$) Coefficients. Notes: We use the 140 estimates for which we have both a value for $\alpha$ and $\beta$; the vertical axis is truncated at $-0.4$ for better visual rendering.

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 $\alpha$ and $\beta$. 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 $\alpha$ and $\beta$), it is unfeasible in our case: we would need not only a measure for the variance of

$\alpha$ and $\beta$ 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 $\alpha$ and one for $\beta$. 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 $\alpha$ and $\beta$ 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 $\alpha$, considering that the same concepts and equations (up to replacing $\alpha$ with $\beta$) also apply to $\beta$.

The fixed-effects model assumes the following:

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

where $\alpha_j$ is the $j$th parameter in the dataset, with $j = \{1, ..., k\}$, $k$ being the total number of estimates; and $\alpha_0$ being 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 $\alpha_j$ varies is because of sampling errors, represented here by $\epsilon_j$. 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 $\alpha_0$ is then to compute a weighted average of the $\alpha_j$, 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, $\alpha_j$, is an estimator of the study's true effect size, $\mu_j$, plus a sampling error, $\epsilon_j$. The true effect size, $\mu_j$, comes from a homogeneous population with a "grand mean", $\alpha_0$, plus a second source of error, $\xi_j$, which is assumed to be distributed as $\xi_j \sim \mathcal{N}(0, \tau^2)$, where $\tau^2$ 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 $\epsilon_j$ is the sampling error for $\alpha_j$, which is an estimate for $\mu_j$ (the true effect size). This is, in turn decomposed into the grand mean, $\alpha_0$, and the second error term, $\xi_j$. 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 $\alpha_0$ 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 $\tau^2$. 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 $\alpha$ and $\beta$ 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.[18]

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 $\alpha_{ij}$ the $j$th estimate of parameter $\alpha$ from study $i$. Then, the first level is defined as:

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

where $\mu_{ij}$ 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 $\theta_i$ represents the average disadvantageous inequality in study $i$ and $\xi_{ij}^{(2)} \sim \mathcal{N}(0, \tau_{(2)}^2)$. The last level is:

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

where $\alpha_0$ is the population mean of $\alpha$ (what we are interested in) and $\xi_i^{(3)} \sim \mathcal{N}(0, \tau_{(3)}^2)$. We can combine the three levels into a single equation to have

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

[18]We adopt the cluster-robust correction in Hedges, Tipton and Johnson (2010).

Table 5: Meta-Analytic Average of Disadvantageous Inequality ($\alpha$)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Disadvantageous Inequality Coefficient ($\alpha_0$) | 0.300 | 0.328 | 0.469 | 0.510 |
| | (0.089) | (0.110) | (0.084) | (0.108) |
| p-value | 0.002 | 0.006 | $< 0.0001$ | $< 0.0001$ |
| $\hat{\tau}^2$ | 0.338 | 0.391 | | |
| $I^2$ | 99.97 | 99.97 | | |
| $I^2_{within}$ | | | 42.25 | 37.07 |
| $I^2_{between}$ | | | 57.73 | 62.90 |
| Observations | 145 | 108 | 145 | 108 |
| 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); in both RE and ML models, we use the restricted maximum likelihood method.

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; $\xi_i^{(3)}$, instead, stands for the between-cluster heterogeneity, with a large value for $\tau^2_{(3)}$ indicating that the "true" disadvantageous inequality parameter varies a lot between different studies.[19]

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.

[19] The three-level model assumes that, conditional on being in the same study, the parameters are independent. In equation (12), this implies that $Cov(\epsilon_{ij}, \epsilon_{ih}) = 0$ for every estimate $j \neq h$ in study $i$. Thus, in this model, the only source of dependence between parameters from the same study is the true value parameter and not the estimation error. The Correlated Hierarchical Effects (CHE) model proposed in Pustejovsky and Tipton (2022) extends the three-level model by allowing estimates from the same study to have correlated estimation errors, i.e. $Cov(\epsilon_{ij}, \epsilon_{ih}) = \rho v_i^2$, where $\rho$ is assumed to be a constant and common correlation coefficient between estimates from the same study $i$ and $v_i^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} v_{ij}^2$.

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 ($\alpha$), 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 0.3 while the coefficient in the multi-level model is 0.469. The difference 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 $\alpha$. 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 (99.97%) in the two-level random-effects model is due to between observations heterogeneity rather than sampling variance.[20] In the multi-level model, instead, around 42.25% of the variability in the data is due to heterogeneity within studies ($I^2_{within}$), 57.73% to heterogeneity across studies ($I^2_{between}$) and the remainder to sampling variance.

The meta-analytic average of $\beta$ in the two-level random-effects model is 0.282, smaller than in the multi-level specification for the same reason discussed above. In our preferred specification with a multi-level random-effects model, the estimate of $\beta$ is 0.331. In both cases, the estimates are statistically different from zero at any conventional significance level. This value is in line with the weighted average and the median of $\beta$ from the distribution reported in FS (0.315 and 0.290). We, thus, find evidence of equity concerns in the realm of advantageous situations. The $I^2$ statistics shows that, in the two-level random-effects

[20]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 ($\beta$)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Advantageous Inequality Coefficient ($\beta_0$) | 0.282 | 0.232 | 0.331 | 0.322 |
| | (0.061) | (0.057) | (0.031) | (0.036) |
| p-value | $< 0.0001$ | 0.002 | $< 0.0001$ | $< 0.0001$ |
| $\hat{\tau}^2$ | 0.060 | 0.047 | | |
| $I^2$ | 99.45 | 98.97 | | |
| $I^2_{within}$ | | | 33.16 | 32.79 |
| $I^2_{between}$ | | | 66.07 | 66.03 |
| Observations | 144 | 103 | 144 | 103 |
| 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); in both RE and ML models, we use the restricted maximum likelihood method.

specification, 99% of the variability in $\beta$ can be attributed to between observations heterogeneity; in the multi-level model, instead, around 33.16% of the variability is due to within study heterogeneity and around 66.07% to between studies heterogeneity. Finally, while the theoretical assumptions in FS hold in our meta-analysis, since $\alpha \geq \beta$ and $0 \leq \beta < 1$, we cannot claim that the estimate of $\alpha$ is statistically greater than the estimate of $\beta$.

## 4.3 Explaining Heterogeneity

The estimates in our dataset come from studies that are very different from each other, for example, because of the 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 $\alpha$ and $\beta$ depend mainly on sampling errors, either at the observation or study level, as we did previously. In order to explain the heterogeneity, we add to the multi-level specification described in equation (12) a set of regressors:

$$\alpha_{ij} = \alpha_0 + \delta X_{ij} + \xi_{ij}^{(2)} + \xi_i^{(3)} + \epsilon_{ij}. \tag{13}$$

Table 7: Explaining Heterogeneity

| | Disadvantageous Inequality ($\alpha$) | | | | Advantageous Inequality ($\beta$) | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Strategic | 0.593*** | | | | −0.138** | | | |
| | (0.198) | | | | (0.058) | | | |
| Non Students | | 0.104 | | | | 0.009 | | |
| | | (0.176) | | | | (0.067) | | |
| North Europe | | | 0.333* | | | | −0.201** | |
| | | | (0.190) | | | | (0.079) | |
| USA | | | 0.178 | | | | −0.198** | |
| | | | (0.218) | | | | (0.078) | |
| China | | | -0.124 | | | | -0.011 | |
| | | | (0.199) | | | | (0.194) | |
| Online | | | | 0.016 | | | | 0.040 |
| | | | | (0.198) | | | | (0.090) |
| Constant | 0.239*** | 0.442*** | 0.265* | 0.467*** | 0.369*** | 0.328*** | 0.482*** | 0.324*** |
| | (0.084) | (0.102) | (0.134) | (0.096) | (0.035) | (0.037) | (0.067) | (0.033) |
| Observations | 145 | 145 | 142 | 145 | 144 | 144 | 141 | 144 |

Notes: SEs in parenthesis are cluster-robust (Hedges, Tipton and Johnson, 2010). *p<0.1; **p<0.05; ***p<0.01.

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 four different regressions with parsimonious models that only include one moderator variable at a time.

Since $X_{ij}$ is composed of dummy variables, each coefficient represents the shift of the population mean $\alpha_0$ with respect to the baseline condition. The meta-regressions for $\alpha$ and $\beta$ are presented in Table 7. A positive coefficient indicates more sensitivity (i.e., stronger aversion) to disadvantageous or advantageous inequality compared to the baseline condition: and we chose the baseline conditions as follows: for the experimental task (columns 1 and 5), the omitted category is individual decision-making tasks; for the subject population (columns 2 and 6), the omitted category is college students (and we pool non-representative samples of adults and representative samples in the non-students category); for geographic location (columns 3 and 7), the omitted category is Southern Europe; and for experimental implementation (columns 4 and 8), the omitted variable is in-person (where we pool laboratory and classroom experiments).[21]

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. First, while $\alpha$ and $\beta$ are both positive and significantly different from 0 regardless of whether they are estimated with games or individual choices, the meta-synthetic averages are strongly affected by the experimental task (see also Figure 8): sensitivity to disadvantageous inequality is stronger in strategic environments than in individual decision-making tasks while the reverse is true for sensitivity to advantageous inequality. Therefore — since both parameters are strictly positive and, thus, capturing inequality aversion in both environments — strategic environments dampen the guilt from being ahead in social comparisons and, at the same time, they enhance the envy from being behind. As conjectured in Fehr, Naef and Schmidt (2006), this suggests that efficiency motives (inducing individuals to values others' payoff positively) may be weakened by the competitive nature of strategic environments, where participants tend to view themselves as opponents rather as partners. Second, populations of college students are not different from other subject populations and participation in person to laboratory and classroom experiments does not affect estimates with respect to experiments conducted remotely. Third, participants from Southern Europe (France, Italy, Spain, and Turkey) are less averse to disadvantageous inequality than participants from Northern Europe (Denmark, Germany, Netherlands, Sweden, Switzerland and UK) and more averse to advantageous inequality than participants from both Northern

[21]When the same study offers to our dataset both estimates computed in strategic environments and estimates computed in individual-decision making environments (Yang, Onderstal and Schram 2016 and Diaz, Houser, Ifcher and Zarghamee 2021), we consider the two estimates as coming from two different studies. This allows for a crispier analysis of estimates' heterogeneity in this dimension, with results similar to those we obtain when estimating a multi-level random-effects model on two sub-samples, one for estimates computed in strategic environments and one for estimates computed in individual-decision making environments (see Appendix F).

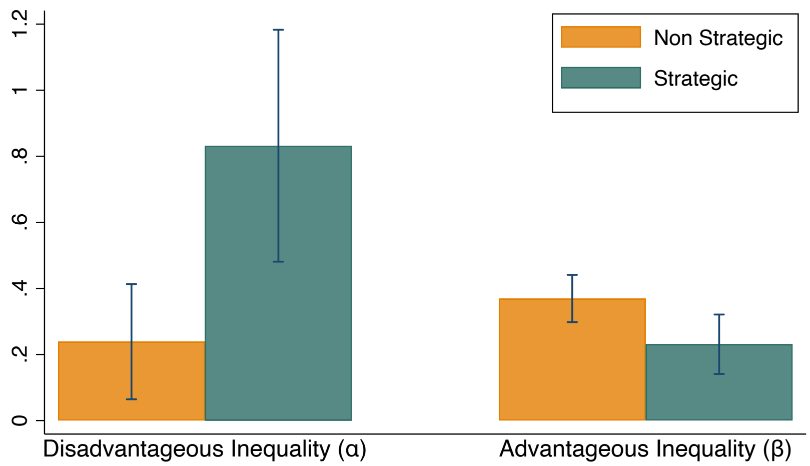


Figure 8: ML Estimates of Disadvantageous Inequality ($\alpha$) and Advantageous Inequality ($\beta$) in the Sub-Samples with Strategic vs Non Strategic Environments. Notes: Estimates are from columns 1 and 5 in Table 7; vertical bars are 95% confidence intervals.

Europe and the US.

## 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 other-regarding preferences parameters, 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

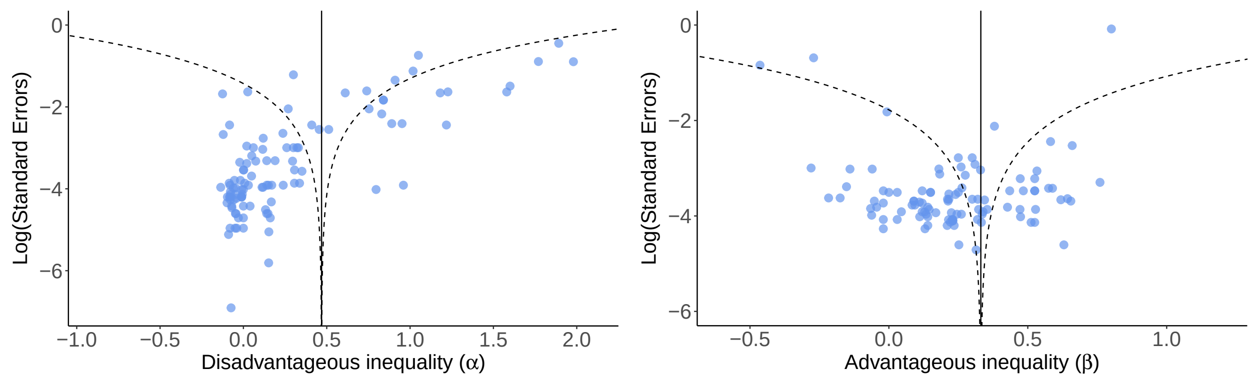


Figure 9: Funnel Plots of Disadvantageous ($\alpha$) and Advantageous Inequality ($\beta$) Coefficients. Notes: The vertical continuous line is at the meta-analytic average from column (3) in Tables 5 and 6 ($\alpha = 0.469$ and $\beta = 0.331$) and the diagonal dotted curves represent a p-value of 0.05 for a z-test whose null hypothesis is that the estimate is equal to the meta-analytic average (i.e., estimates below each dotted line are statistically different from this average). The horizontal axis is truncated at 2.2 ($\alpha$) and $-0.6$ ($\beta$) for better visual rendering. Only those estimates with reported SEs are included.

mean should show a lower precision. Without selective reporting, we expect to see a funnel-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 9 shows the funnel plots for the advantageous and disadvantageous inequality coefficients. The distribution for $\alpha$ looks highly asymmetric: observations with a negative (and large in magnitude) value of $\alpha$ which is imprecisely estimated are "missing". A similar, albeit more attenuated, effect is present also for $\beta$: 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:

$$\alpha_{ij} = \alpha_0 + \delta SE_{ij} + \xi_{ij}^{(2)} + \xi_i^{(3)} + \epsilon_{ij}. \tag{14}$$

$$\beta_{ij} = \beta_0 + \gamma SE_{ij} + \nu_{ij}^{(2)} + \nu_i^{(3)} + \eta_{ij}. \tag{15}$$

In this model, $\delta$ and $\gamma$ capture the degree of selective reporting bias while $\alpha_0$ and $\beta_0$ represent

Table 8: FAT-PET Analysis

| | $\alpha$ | | $\beta$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Constant | 0.340 | 0.192 | 0.404 | 0.400 |
| | (0.075) | (0.067) | (0.037) | (0.043) |
| Standard Errors | 1.580 | 3.039 | -2.027 | -2.260 |
| | (0.590) | (0.646) | (0.762) | (0.772) |
| p-value | $< 0.0001$ | 0.008 | $< 0.0001$ | $< 0.0001$ |
| $I^2_{within}$ | 6.77 | 8.07 | 32.77 | 31.41 |
| $I^2_{between}$ | 93.18 | 91.82 | 66.45 | 67.46 |
| Observations | 143 | 106 | 144 | 102 |
| Model | ML | ML | ML | ML |
| Sample | Full | Restricted | Full | Restricted |

Notes: All columns estimate a multi-level random-effects model with the restricted maximum likelihood method; 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 $H_0$ : Constant = 0; SEs in parenthesis are cluster-robust (Hedges, Tipton and Johnson, 2010); for columns (1) and (2), two estimates from Diaz et al. (2021) were omitted because $|DFBETAS| > 1$; for column (4), one estimate from Bellemare et al. (2008) was removed because $|DFBETAS| > 1$.

the selection-bias-corrected value of the parameters. 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). As reported in Table 8, the coefficient for $\delta$ is positive and statistically significant, while the coefficient for $\gamma$ is negative and statistically significant ($\delta = 1.580$ with p-value$= 0.011$ in the full sample; $\gamma = -2.027$ with p-value$= 0.011$ in the full sample). At the same time, the constants, $\alpha_0$ and $\beta_0$, are positive and highly significant (both in the full and in the restricted sample), indicating the presence of both disadvantageous and advantageous inequity aversion even after correcting for possible publication bias: the publication-bias-corrected 95% confidence intervals for $\alpha$ and $\beta$ are, respectively, $[0.188, 0.492]$ and $[0.33, 0.478]$.

We note that the asymmetry in the funnel plots could be generated also in the absence of publication bias — for example, because of constraints in the estimation of $\alpha$ and $\beta$ when

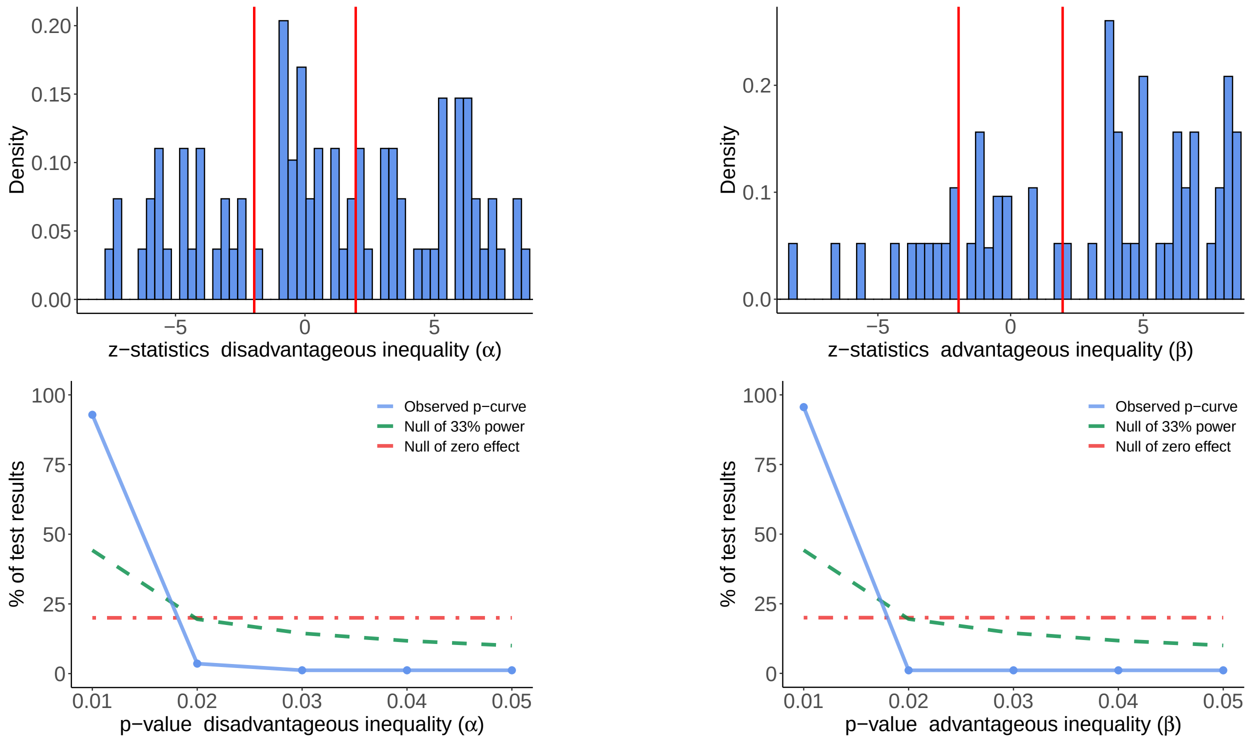


Figure 10: Distribution of z-statistics, top-panels, and p-curves, bottom-panel, of Disadvantageous inequality ($\alpha$) and Advantageous Inequality ($\beta$). Notes: in both figures the test is the null hypothesis of $\alpha = 0$ or $\beta = 0$; red vertical lines are at -1.96 and 1.96. Only those estimates with reported SEs are included.

eliciting these parameters with the experimental tasks typically employed by the literature.[22] 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 $\alpha$ smaller than $-1$ and values of $\beta$ larger than 1, which imply that an individual is willing to burn money just to increase the gap in outcomes when behind or just to reduce the gap when ahead. Indeed, the 145 estimates of $\alpha$ and the 144 estimates of $\beta$ in our dataset never take values beyond those thresholds and this can hardly be deemed proof of publication bias.

Another form of publication bias consists in the practice of p-hacking. Journals might be biased in publishing statistically significant results and, in turn, researchers might be tempted

[22]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 $\beta \in [0, 1]$.

to push analyses just below a threshold (e.g., a p-value of 5%) by, for example, changing econometric specification or the number of covariates in a regression. Two tools employed in the literature to detect publication bias in the form of p-hacking are the histograms of z-statistics and the p-curve (Simonsohn, Nelson and Simmons, 2014). Under the presence of p-hacking, we would see a bunching of z-statistics right above the threshold of statistical significance at the 5% level, i.e., $|1.96|$. This is because researchers who obtain z-statistics just below this value have an incentive to push it right above, thus creating a discontinuity around $|1.96|$ in the histograms. From the top panels of Figure 10, we see a jump just above 1.96 in the histogram for the disadvantageous inequality parameter and a jump right below $-1.96$ in the histogram for the advantageous inequality parameter. While this could suggest the presence of p-hacking (with researchers pushing statistical significance above 5% to show that $\alpha$ is greater than zero and that $\beta$ is less than zero), we must note that we have very few observations around the $|1.96|$ cutoff making this far from a conclusive proof.

The p-curve looks, instead, at the distribution of statistically significant p-values. Under the null hypothesis — which, in our case, is that the parameter is equal to zero — the expected distribution of statistically significant p-values is a uniform (by the definition of p-values) and we expect to see a flat p-curve. If the null hypothesis is false (that is, the parameter is different from zero) and researchers do not engage in p-hacking, we expect to see a right-skewed distribution, since researchers are more likely to find and to report small p-values rather than large ones. If the null hypothesis is true but researchers do engage in p-hacking, researchers try to turn non-significant results into significant ones and, most likely, they stop as soon as they reach this goal. In this case, we expect to see a left-skewed distribution, since researchers add to the true flat distribution of statistically significant p-values, observations that are pushed just above the 5% significance threshold. The bottom panels of Figure 10 show that the p-curves for both $\alpha$ and $\beta$ are highly right-skewed, thus strongly supporting the hypothesis that the parameters are different from zero and the absence of p-hacking.[23]

# 5 Conclusion

In this paper, we reported the results of a meta-analysis of empirical estimates 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 sensitivity to disadvantageous inequality ($\alpha$) and sensitivity to advantageous inequality ($\beta$). 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 sensitivity to disadvantageous inequality is 0.469 with a 95% confidence interval of $[0.298, 0.639]$; the mean sensitivity to advantageous inequality coefficient is, instead, 0.331 with a 95% confidence interval $[0.269, 0.393]$.[24] This means that, on average, individuals feel *guilt* and are willing to pay \$0.49 to increase others' earnings by \$1 when ahead; and that they feel *envy* and are willing to pay \$0.87 to decrease others' earnings by \$1 when behind.[25] The theoretical assumptions originally made in FS — that is, $\alpha \geq \beta$ and $0 < \beta < 1$ — 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 in our dataset.

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

[23]While a useful instrument to detect p-hacking, the p-curve is not a definitive test. For example, if studies are well powered, the p-curve is right-skewed even in the case of a true null and mild p-hacking. Moreover, we note that some of the assumptions in (Simonsohn et al., 2014) are not satisfied in our data: many studies do not test directly the null hypothesis that the parameter is equal to zero and not all p-values come from independent studies.

[24]In the Bayesian analysis, the mean envy coefficient is 0.467 with a 95% probability that the true value lies in the interval $[0.302, 0.642]$; the mean guilt coefficient is, instead, 0.331 with a 95% probability that the true value lies in the interval $[0.266, 0.396]$.

[25]These WTPs are computed as $\beta/(1-\beta)$ when ahead and $\alpha/(1-\alpha)$ when behind.

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) 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 Morality in Strategic Interactions," 2021. Unpublished Manuscript, available at https://www.tse-fr.eu/publications/estimating-social-preferences-and-kantian-morality-strategic-interactions-0.

**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.

**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, available at https://osf.io/preprints/metaarxiv/hnefr/.

**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).

**Carpenter, Jeffrey and Andrea Robbett**, "Measuring Socially Appropriate Social Preferences," 2022. Unpublished Manuscript, available at https://docs.iza.org/dp15590.pdf.

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

_ **, Thomas Riechmann, Bodo Sturm, and Carsten Vogt**, "Inequity Aversion and Individual Behavior in Public Good Games: An Experimental Investigation," 2007. Unpublished Manuscript, available at https://www.econstor.eu/bitstream/10419/24598/1/dp07034.pdf.

**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.

**Diaz, Lina, Daniel Houser, John Ifcher, and Homa Zarghamee**, "Estimating Social Preferences Using Stated Satisfaction: Novel Support for Inequity Aversion," 2021. Unpublished Manuscript, available at https://docs.iza.org/dp14347.pdf.

**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 and Herbert Gintis**, "Blaming the Messenger: Notes on the Current State of Experimental Economics," *Journal of Economic Behavior & Organization*, 2010, *73* (1), 109–119.

**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.

_ **, Michael Naef, and Klaus M Schmidt**, "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment," *American Economic Review*, 2006, *96* (5), 1912–1917.

_ **, 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.

**Hedegaard, Morten, Rudolf Kerschbamer, Daniel Müller, and Jean-Robert Tyran**, "Distributional Preferences Explain Individual Behavior Across Games and Time," *Games and Economic Behavior*, 2021, *128*, 231–255.

**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.

**Hu, Yang, Lisheng He, Lei Zhang, Thorben Wölk, Jean-Claude Dreher, and Bernd Weber**, "Spreading Inequality: Neural Computations Underlying Paying-it-forward Reciprocity," *Social Cognitive and Affective Neuroscience*, 2018, *13* (6), 578–589.

**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 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.

**Kerschbamer, Rudolf and Daniel Müller**, "Social Preferences and Political Attitudes: An online Experiment on a Large Heterogeneous Sample," *Journal of Public Economics*, 2020, *182*, 104076.

**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.

**Meager, Rachael**, "Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments," *American Economic*

*Journal: Applied Economics*, 2019, *11* (1), 57–91.

— , "Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature," *American Economic Review*, 2022, *112* (6), 1818–47.

**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.

**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**, "Meta-Analysis with Robust Variance Estimation: Expanding the Range of Working Models," *Prevention Science*, 2022, *23* (3), 425–438.

**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.

**Sáez, Ignacio, Lusha Zhu, Eric Set, Andrew Kayser, and Ming Hsu**, "Dopamine Modulates Egalitarian Behavior in Humans," *Current Biology*, 2015, *25* (7), 912–919.

**Simonsohn, Uri, Leif D Nelson, and Joseph P Simmons**, "P-Curve: a Key to the File-Drawer," *Journal of Experimental Psychology: General*, 2014, *143* (2), 534.

**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.

**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, Thomas Riechmann, Bodo Sturm, and Carsten Vogt**, "Inequity Aversion and Individual Behavior in Public Good Games: 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. **Chen, Yan, and Sherry Xin Li**, "Group Identity and Social Preferences," *American Economic Review*, 2009, 99: 431–457.
9. **Cabrales, Antonio, Raffaele Miniaci, Marco Piovesan, and Giovanni Ponti**, "Social Preferences and Strategic Uncertainty: An Experiment on Markets and Contracts," *American Economic Review*, 2010, 100: 2261–2278.
10. **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.
11. **Daruvala, Dinky**, "Would the Right Social Preference Model Please Stand Up!," *Journal of Economic Behavior & Organization*, 2010, 73(2): 199–208.
12. **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.

13. **Messer Kent D., Gregory L. Poe, Daniel Rondeau, William D. Schulze, and Christian A. Vossler**, "Social Preferences and Voting: An Exploration using a Novel Preference Revealing Mechanism," *Journal of Public Economics*, 2010, 94 (3-4): 308–317.

14. **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.

15. **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.

16. **Wright, Nicholas D., Mkael Symmonds, Stephen M. Fleming, and Raymond J. Dolan**, "Neural Segregation of Objective and Contextual Aspects of Fairness," *Journal of Neuroscience*, 2011, 31 (14): 5244-5252.

17. **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.

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

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

20. **Wu, Yan, Hongbo Yu, Bo Shen, Rongjun Yu, Zhiheng Zhou, Guoping Zhang, Yushi Jiang, and Xiaolin Zhou**, "Neural Basis of Increased Costly Norm Enforcement under Adversity," *Social Cognitive and Affective Neuroscience*, 2014, 9 (12): 1862-1871.

21. **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.

22. **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.

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

24. **Saez, Ignacio, Lusha Zhu, Eric Set, Andrew Kayser, and Ming Hsu**, "Dopamine Modulates Egalitarian Behavior In Humans," *Current Biology*, 2015, 30;25 (7): 912-919.

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

26. **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.

27. **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.

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

29. **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.

30. **Hu, Yang, Lisheng He, Lei Zhang, Thorben Wölk, Jean-Claude Dreher, and Bernd Weber**, "Spreading Inequality: Neural Computations Underlying Paying-It-Forward Reciprocity," *Social Cognitive and Affective Neuroscience*, 2018, 13 (6): 578-589.

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

32. **Gao, Xiaoxue, Hongbo Yu, Ignacio Sáez, Philip R Blue, Lusha Zhu, Ming Hsu, and Xiaolin Zhou**, "Distinguishing Neural Correlates of Context-dependent Advantageous- and Disadvantageous-inequity Aversion," *Proceedings of the National Academy of Sciences*, 2018, 115 (33): E7680-E7689.

33. **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.

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

35. **Kerschbamer, Rudolf, and Daniel Müller**, "Social Preferences, Political Opinions and Charitable Giving: An Online Experiment on a Large Heterogeneous Sample," *Journal of Public Economics*, 2020, 182.

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

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

38. **Hedegaard, Morten, Rudolf Kerschbamer, Daniel Müller, and Jean-Robert Tyran**, "Distributional Preferences Explain Individual Behavior across Games and Time," *Games and Economic Behavior*, 2021, 128: 231–255.

39. **Carpenter, Jeffrey, and Andrea Robbett**, "Measuring Socially Appropriate Social Preferences," Unpublished Manuscript, 2022.

40. **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 9: List of Coded Variables in the Dataset

| Variable | Description |
|---|---|
| study_id | ID for the 42 studies in the analysis (from 1 to 42) |
| 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 $\alpha$ elicited in a strategic game |
| strategic_beta | = 1 if $\beta$ elicited in a strategic game |
| games_alpha | Games used to elicit $\alpha$ |
| games_beta | Games used to elicit $\beta$ |
| 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 ($\alpha$) |
| alpha_se | SE of $\alpha$ |
| alpha_sd | SD of $\alpha$ |
| beta | Advantageous inequality coefficient ($\beta$) |
| beta_se | SE of $\beta$ |
| beta_sd | SD of$\beta$ |
| 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 |
| t-stat | t-statistics of the estimate |
| 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 10: Experimental Tasks and Classification as Strategic

| Experimental Tasks Used To Elicit Parameters | Strategic Environment |
|---|---|
| Disadvantageous Inequality Coefficient ($\alpha$) | |
| 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 ($\beta$) | |
| 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 $\alpha$, but the same applies also to $\beta$. Consider the dataset $(\alpha_j, se_i^2)_{j=1}^k$, where $k$ is the total number of estimates and $\alpha_j$ the $j$th observation of the disadvantageous inequality parameter, with its associated standard error $se_j$. We then assume that the reported estimate $\alpha_j$ 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 $\alpha_0$ is the overall mean of the disadvantageous inequality parameters $\bar{\alpha}_j$, and $\tau^2$ 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 $\alpha_0$. The pooling equation can be written as follows:

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

where $\omega_j$ 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 $\tau^2$ is low, meaning that if there is low heterogeneity across studies, more weight will be given to $\alpha_0$.

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 $\alpha$ and $\beta$ are the following:

$$\begin{aligned} \alpha_j|\bar{\alpha}_j, se_j &\sim \mathcal{N}(\bar{\alpha}_j, se_j^2) \\ \bar{\alpha}_j|\alpha_0, \tau &\sim \mathcal{N}(\alpha_0, \tau^2) \\ \alpha_0 &\sim \mathcal{N}(0.25, 1) \\ \tau &\sim half\mathcal{N}(0, 1) \end{aligned}$$

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

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 $\alpha_0$ and by using the three sigma-rule of thumb, what the prior is saying is that our initial opinion for the true value of $\alpha_0$ is that the parameter lies between $-1.75$ and 2.25 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 $\alpha_0$ of 0.301, with a 95% credible interval between [0.201, 0.401]. The frequentist random-effects model returns a value for $\alpha_0$ of 0.300 with a 95% confidence interval between [0.198, 0.401]. As we can see the two values are nearly identical, and the same happens for the estimate of $\hat{\tau}^2$ with a mean value in the Bayesian procedure of 0.345 and of 0.338 in the frequentist approach.

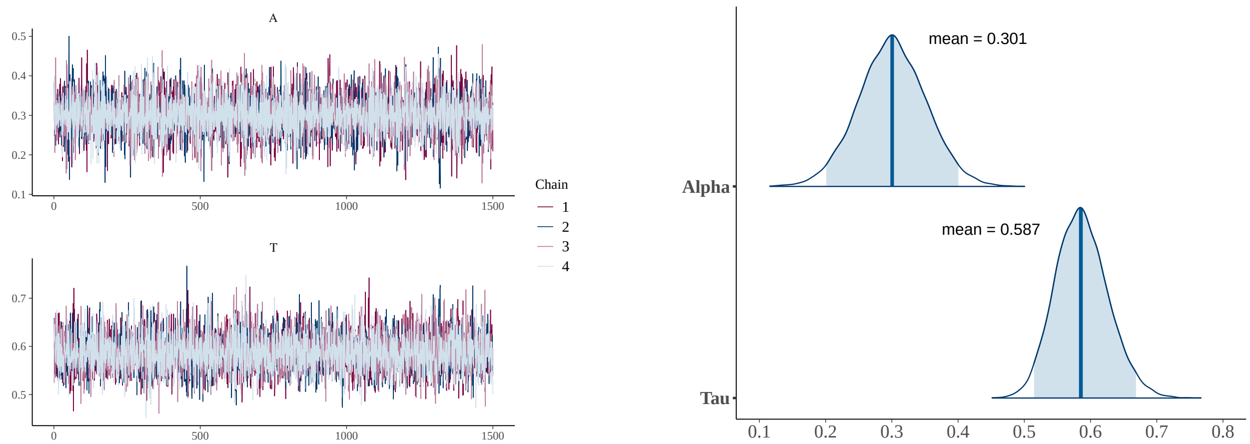


Figure 11: The first figure shows the 1,500 draws for $\alpha_0$ and $\tau$ in the four Markov Chain after the warmup, showing good convergence of the procedure. The second figure shows the posterior distributions of the two parameters. Shaded blue areas correspond to 95% credible intervals.

Table 11: Summary of the Bayesian Hierarchical Model Estimate for $\alpha$

| Parameter | Rhat | ESS | Mean | SD | 2.5% | 25% | 50% | 75% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha_0$ | 1.000 | 14723 | 0.301 | 0.051 | 0.201 | 0.267 | 0.300 | 0.335 | 0.401 |
| $\hat{\tau}$ | 1.000 | 10594 | 0.587 | 0.039 | 0.515 | 0.559 | 0.585 | 0.612 | 0.669 |

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 $\beta_0$ is 0.282, the same as in the random-effects model. The estimates of $\hat{\tau}^2$ are equal, with a mean value in the Bayesian procedure of 0.06 and of 0.06 in the frequentist approach.

Table 12: Summary of the Bayesian Hierarchical Model Estimate for $\beta$

| Parameter | Rhat | ESS | Mean | SD | 2.5% | 25% | 50% | 75% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|
| $\beta_0$ | 1.000 | 7023 | 0.282 | 0.021 | 0.237 | 0.268 | 0.282 | 0.296 | 0.323 |
| $\hat{\tau}$ | 1.000 | 5649 | 0.248 | 0.017 | 0.218 | 0.237 | 0.248 | 0.259 | 0.282 |

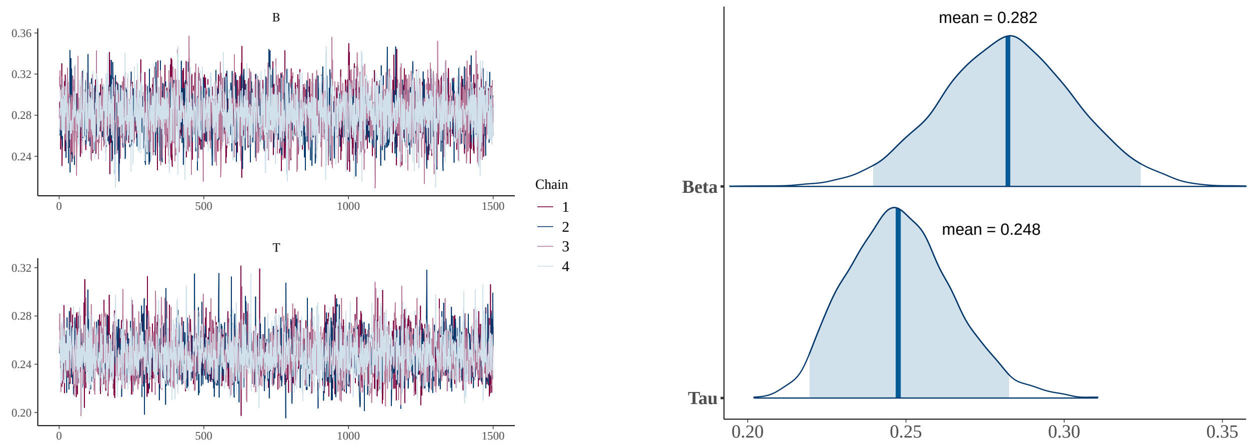


Figure 12: The first figure shows the 1,500 draws in the four Markov Chain after the warmup, showing good convergence of the procedure. The second figure shows the posterior distributions of $\beta_0$ and $\tau$. Shaded blue areas correspond to 95% credible intervals.

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 $\alpha$ and $\beta$ 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 0.467, with a 95% probability that the true value falls in the interval [0.302, 0.642]. This is in line with what we found in the frequentist analysis, with an estimate for $\alpha$ of 0.469 and a confidence interval of [0.298, 0.539].

Table 13: Summary of the Bayesian Hierarchical Model Estimate for $\alpha$ with paper level

| Parameter | Rhat | ESS | Mean | SD | 2.5% | 25% | 50% | 75% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha_0$ | 1.000 | 7834 | 0.467 | 0.087 | 0.302 | 0.409 | 0.467 | 0.524 | 0.642 |
| $\hat{\tau}$ | 1.000 | 4556 | 0.471 | 0.079 | 0.332 | 0.416 | 0.466 | 0.519 | 0.642 |

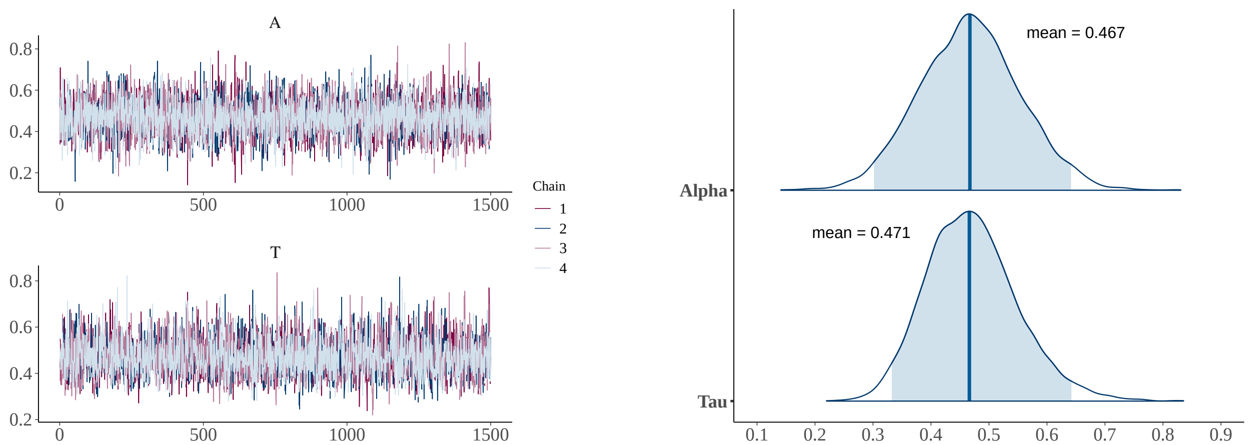


Figure 13: The first figure shows the 1,500 draws in the four Markov Chain after the warmup, showing good convergence of the procedure. The second figure shows the posterior distributions of $\alpha_0$ and $\tau$. Shaded blue areas correspond to 95% credible intervals.

Now discussing $\beta$, the Bayesian procedure returns a mean advantageous inequality coefficient of 0.331, with a 95% probability that the true value falls in the interval [0.266, 0.396]. Once again, this is in line with what we found in the frequentist analysis, with an estimate for $\beta$ of 0.331 and a confidence interval of [0.269, 0.393].

Table 14: Summary of the Bayesian Hierarchical Model Estimate for $\beta$ with paper level

| Parameter | Rhat | ESS | Mean | SD | 2.5% | 25% | 50% | 75% | 97.5% |
|---|---|---|---|---|---|---|---|---|---|
| $\beta_0$ | 1.000 | 8857 | 0.331 | 0.033 | 0.266 | 0.309 | 0.331 | 0.352 | 0.396 |
| $\hat{\tau}$ | 1.000 | 6521 | 0.176 | 0.025 | 0.133 | 0.159 | 0.174 | 0.192 | 0.231 |

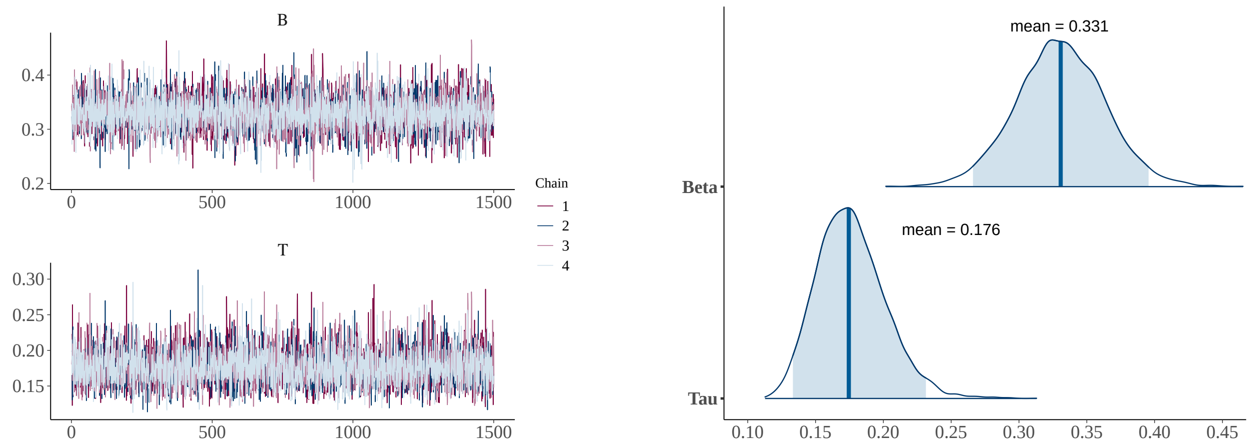


Figure 14: The first figure shows the 1,500 draws in the four Markov Chain after the warmup, showing good convergence of the procedure. The second figure shows the posterior distributions of $\alpha_0$ and $\tau$. Shaded blue areas correspond to 95% credible intervals.

We provide here a table with a sensitivity analysis on the priors chosen for the bayesian models by looking at the average of the parameters and their 95% credible intervals in the different specifications.

Table 15: Sensitivity Analysis on Priors for the Random Effects Model

| *Prior* | Disadvantageous Inequality ($\alpha_0$) | | Advantageous Inequality ($\beta_0$) | |
|---|---|---|---|---|
| $sd = 2, \phi_p = 0.25$ | 0.300 | [0.205,0.397] | 0.282 | [0.240,0.324] |
| $sd = 0.5, \phi_p = 0.25$ | 0.299 | [0.206,0.394] | 0.282 | [0.239,0.323] |
| $sd = 1, \phi_p = 0$ | 0.299 | [0.204,0.396] | 0.282 | [0.240,0.323] |
| | Average | 95% Credible | Average | 95% Credible |

Table 16: Sensitivity Analysis on Priors for the Multi-Level Model

| *Prior* | Disadvantageous Inequality ($\alpha_0$) | | Advantageous Inequality ($\beta_0$) | |
|---|---|---|---|---|
| $sd = 2, \phi_p = 0.25$ | 0.468 | [0.300,0.643] | 0.330 | [0.267,0.393] |
| $sd = 0.5, \phi_p = 0.25$ | 0.461 | [0.302,0.624] | 0.330 | [0.271,0.391] |
| $sd = 1, \phi_p = 0$ | 0.466 | [0.295,0.647] | 0.330 | [0.267,0.392] |
| | Average | 95% Credible | Average | 95% Credible |

Notes: sd is the standard deviation used for all priors. $\phi_p$ is the mean of the normal prior on the parameter, for both $\alpha$ and $\beta$.

# 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 $\alpha$ and $\beta$.

The specification for the multivariate random-effects model applied in our dataset of inequality sensitivity 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 $(\alpha_j, \beta_j)$ 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 $\alpha$ and $\beta$ 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 $(\alpha_0, \beta_0)$ 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 $(\alpha_0, \beta_0)$. In this multivariate multi-level model we are estimating in addition to the variance of the within and between study errors for $\alpha$ and $\beta$, 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 17. Looking at the latter model we observe an estimate of the average disadvantageous inequality parameter equals to 0.461, which is very close to the one obtained in the univariate specification. The average advantageous inequality parameter is instead estimated to be equal to 0.325, slightly lower than the 0.331 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=0.147).

Table 17: Meta-Analytic Average of Disadvantageous Inequality ($\alpha$) and Advantageous Inequality ($\beta$)

| | RE Full | RE Restricted | ML Full | ML Restricted |
|---|---|---|---|---|
| Disadvantageous Inequality ($\alpha_0$) | 0.293 | 0.320 | 0.461 | 0.503 |
| | (0.051) | (0.064) | (0.088) | (0.116) |
| Advantageous Inequality ($\beta_0$) | 0.277 | 0.234 | 0.325 | 0.323 |
| | (0.022) | (0.022) | (0.031) | (0.036) |
| p-value $\alpha_0$ | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ |
| p-value $\beta_0$ | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ |
| Observations | 140 | 103 | 140 | 103 |

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. In both random effects and multi-level models the restricted maximum likelihood method is used.

# F Non-Strategic vs Strategic Environment

## F.1 Meta-Analytic Averages in Non-Strategic-Environments:

Table 18: Meta-Analytic Average of Disadvantageous Inequality ($\alpha$)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Disadvantageous Inequality Coefficient ($\alpha_0$) | 0.099 | 0.102 | 0.234 | 0.248 |
| | (0.071) | (0.080) | (0.082) | (0.112) |
| p-value | 0.175 | 0.226 | 0.0098 | 0.046 |
| $\hat{\tau}^2$ | 0.092 | 0.098 | | |
| $I^2$ | 99.92 | 99.88 | | |
| $I^2_{within}$ | | | 4.82 | 0.52 |
| $I^2_{between}$ | | | 95.13 | 99.41 |
| Observations | 85 | 54 | 85 | 54 |
| Model | RE | RE | ML | ML |
| Sample | Full | Restricted | Full | Restricted |
| Strategic Environment | No | No | No | No |

Table 19: Meta-Analytic Average of Advantageous Inequality ($\beta$)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Advantageous Inequality Coefficient ($\beta_0$) | 0.346 | 0.293 | 0.370 | 0.362 |
| | (0.058) | (0.048) | (0.035) | (0.041) |
| p-value | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ |
| $\hat{\tau}^2$ | 0.049 | 0.037 | | |
| $I^2$ | 99.40 | 98.81 | | |
| $I^2_{within}$ | | | 18.41 | 9.04 |
| $I^2_{between}$ | | | 80.82 | 89.75 |
| Observations | 104 | 69 | 104 | 69 |
| Model | RE | RE | ML | ML |
| Sample | Full | Restricted | Full | Restricted |
| Strategic Environment | No | No | No | No |

Notes: Columns (1) and (3) in the two tables 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. SEs in parenthesis are cluster-robust.

## F.2 Meta-Analytic Averages in Strategic-Environments:

Table 20: Meta-Analytic Average of Disadvantageous Inequality ($\alpha$)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Disadvantageous Inequality Coefficient ($\alpha_0$) | 0.596 | 0.465 | 0.839 | 0.799 |
| | (0.217) | (0.172) | (0.182) | (0.210) |
| p-value | 0.012 | 0.015 | 0.0002 | 0.0014 |
| $\hat{\tau}^2$ | 0.618 | 0.33 | | |
| $I^2$ | 99.95 | 99.92 | | |
| $I^2_{within}$ | | | 9.05 | 8.23 |
| $I^2_{between}$ | | | 90.91 | 91.74 |
| Observations | 60 | 53 | 60 | 54 |
| Model | RE | RE | ML | ML |
| Sample | Full | Restricted | Full | Restricted |
| Strategic Environment | Yes | Yes | Yes | Yes |

Table 21: Meta-Analytic Average of Advantageous Inequality ($\beta$)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Advantageous Inequality Coefficient ($\beta_0$) | 0.117 | 0.109 | 0.218 | 0.223 |
| | (0.075) | (0.080) | (0.056) | (0.061) |
| p-value | 0.143 | 0.209 | 0.002 | 0.007 |
| $\hat{\tau}^2$ | 0.053 | 0.042 | | |
| $I^2$ | 98.96 | 98.55 | | |
| $I^2_{within}$ | | | 54.71 | 53.02 |
| $I^2_{between}$ | | | 44.11 | 45.45 |
| Observations | 40 | 34 | 40 | 34 |
| Model | RE | RE | ML | ML |
| Sample | Full | Restricted | Full | Restricted |
| Strategic Environment | Yes | Yes | Yes | Yes |

Notes: Columns (1) and (3) in the two tables 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. SEs in parenthesis are cluster-robust.

# G Correlated Hierarchical Effects (CHE) Model

Table 22: Correlated Hierarchical Effects Model

| $\rho$ | Disadvantageous Inequality ($\alpha$) | | Advantageous Inequality ($\beta$) | |
|---|---|---|---|---|
| 0.00 | 0.469 | 0.510 | 0.331 | 0.322 |
| | (0.084) | (0.108) | (0.031) | (0.036) |
| 0.10 | 0.466 | 0.507 | 0.332 | 0.324 |
| | (0.084) | (0.108) | (0.031) | (0.036) |
| 0.20 | 0.463 | 0.503 | 0.334 | 0.325 |
| | (0.083) | (0.107) | (0.030) | (0.036) |
| 0.30 | 0.460 | 0.499 | 0.335 | 0.327 |
| | (0.083) | (0.106) | (0.030) | (0.035) |
| 0.40 | 0.457 | 0.496 | 0.336 | 0.328 |
| | (0.082) | (0.106) | (0.030) | (0.035) |
| 0.50 | 0.454 | 0.492 | 0.338 | 0.329 |
| | (0.082) | (0.105) | (0.030) | (0.035) |
| 0.60 | 0.452 | 0.489 | 0.339 | 0.330 |
| | (0.082) | (0.104) | (0.030) | (0.035) |
| 0.70 | 0.449 | 0.485 | 0.339 | 0.330 |
| | (0.081) | (0.104) | (0.030) | (0.035) |
| 0.80 | 0.446 | 0.481 | 0.340 | 0.331 |
| | (0.081) | (0.103) | (0.030) | (0.035) |
| 0.90 | 0.443 | 0.478 | 0.339 | 0.329 |
| | (0.080) | (0.102) | (0.030) | (0.036) |
| 0.99 | 0.440 | 0.475 | 0.333 | 0.320 |
| | (0.080) | (0.102) | (0.032) | (0.038) |
| Observations | 145 | 108 | 144 | 103 |
| Model | Full | Restricted | Full | Restricted |

Notes: SEs in parenthesis are cluster-robust.

# H Boxplots of Social Preferences Estimates

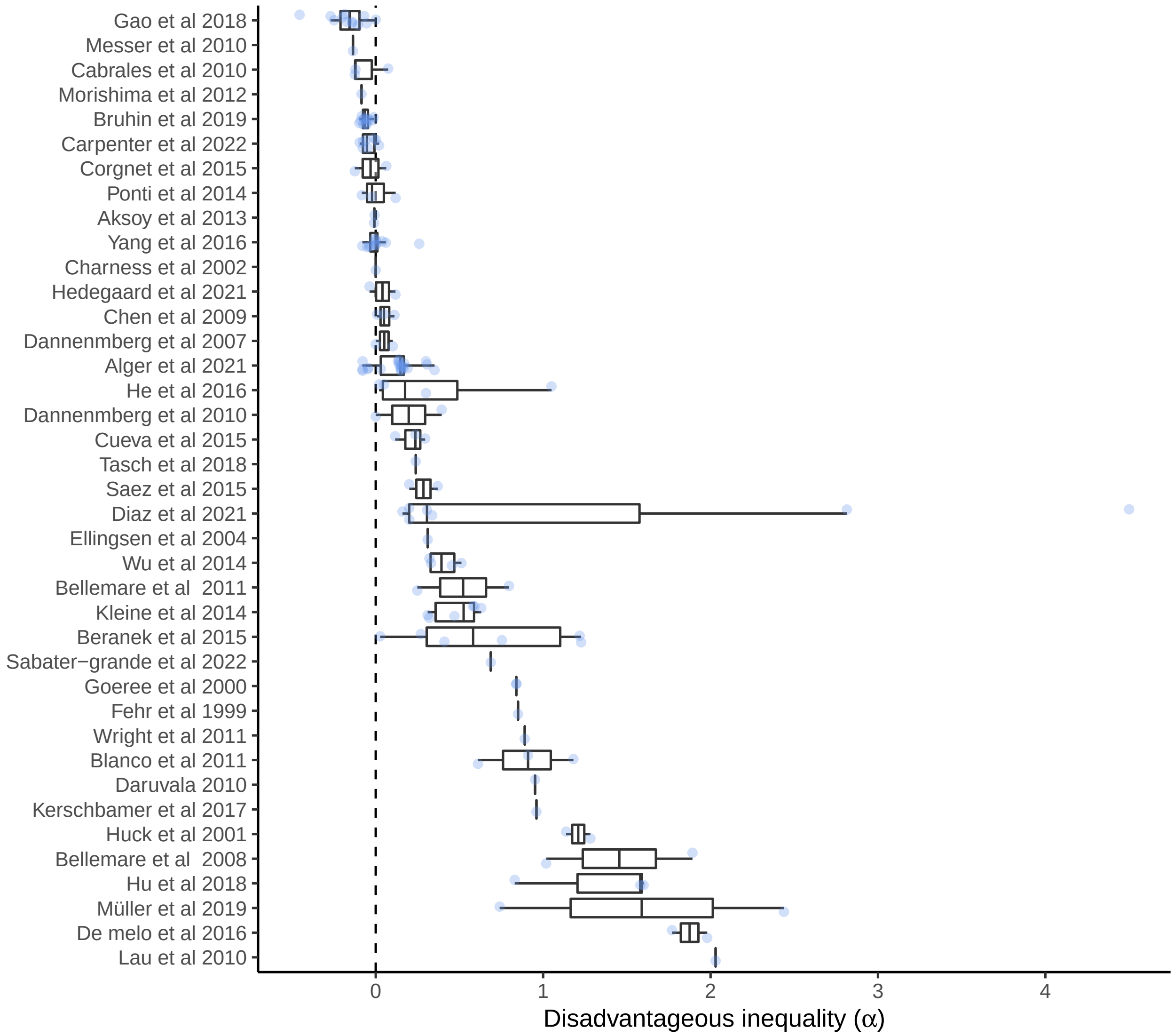


Figure 15: Boxplots of sensitivity to disadvantageous inequality estimates ($\alpha$) by paper.

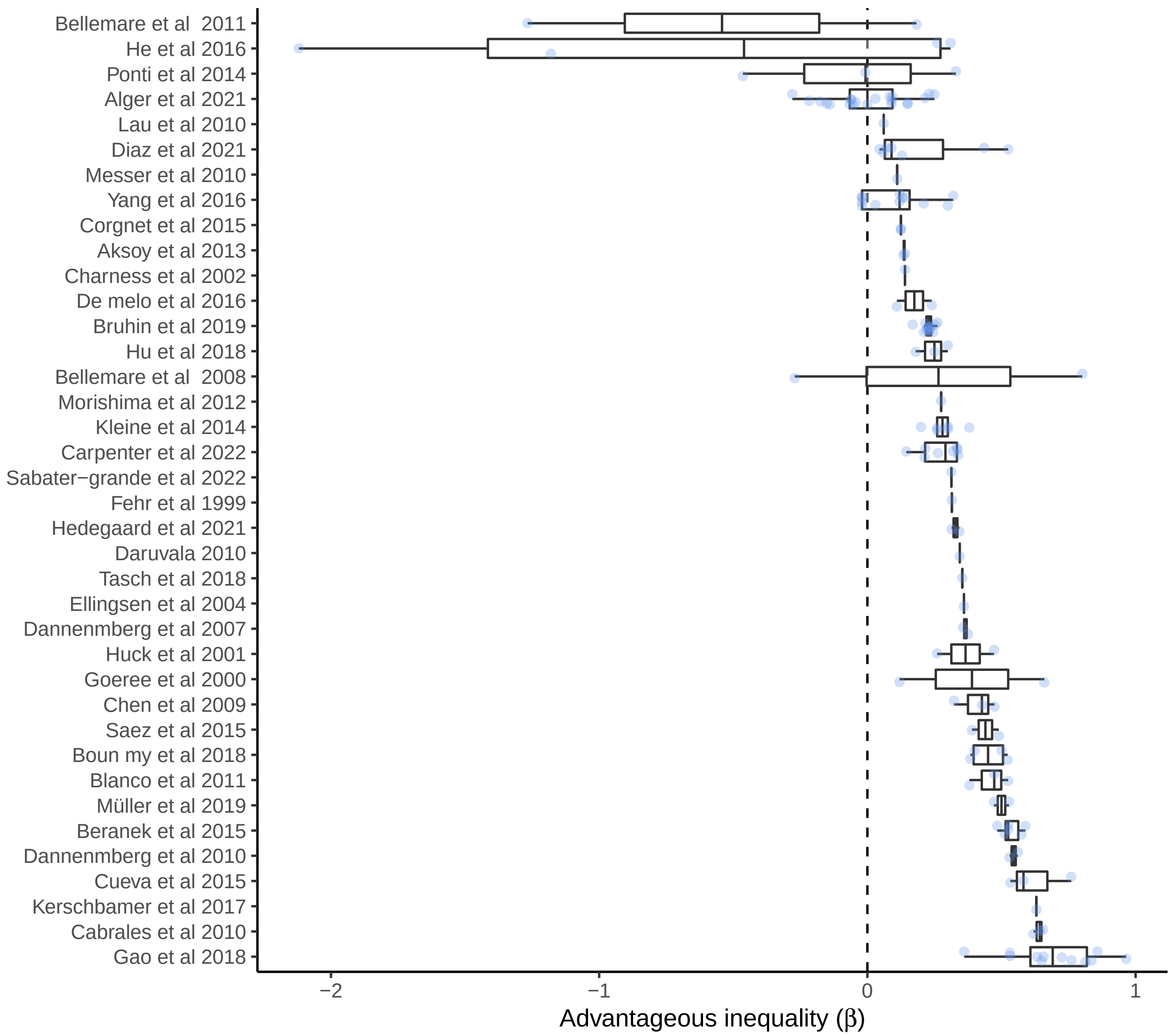


Figure 16: Boxplots of sensitivity to advantageous inequality estimates ($\beta$) by paper.