Inequality, Fairness and Social Capital

Dietmar Fehr[1], Hannes Rau[1], Stefan T. Trautmann[1,2], and Yilong Xu[1]

[1]*University of Heidelberg, Germany*

[2]*Tilburg University, Netherlands*

June 2020

Abstract

Inequality is often associated with negative societal consequences, but identifying a causal relationship is a daunting task. We provide evidence on the impact of unjust economic inequality on social interactions. Using a large-scale controlled experiment, we document that unjust inequality results in a significant decline in trust and trustworthiness. This erosion of social capital is associated with pessimistic beliefs about others' behavior and is muted if there is no direct link between the income-generating process and social interaction. Finally, our data do not support the view that higher status or wealth affects pro-social attitudes: the successful are always more generous, whereas unsuccessful persons display the least efficient and generous behavior regardless of the status of the person who they interact with.

Keywords: inequality, fairness, social capital, experiment

JEL Classifications: C91, D31, D63, D64, H23

---

* We are grateful for comments and suggestions from audiences at Amsterdam, Newcastle, Basel, Heidelberg, Luzern, NYU Abu Dhabi, Tilburg, Renmin University of China, Beijing Information Science and Technology University, WU Vienna, Max-Planck-Institute Bonn, ESA Richmond, FUR Warwick, EEA Cologne, the Social Economy Workshop in Bologna and The Wuhan Experimental Economics Workshop. We thank the Associate Editor and two referees for very constructive comments, and Maria Bigoni, Yan Chen, and Ismael Rodriguez Lara for detailed discussions. Dietmar Fehr thanks the German Science Foundation (DFG) for financial support through individual grant FE 1452/3-1.
Email: dietmar.fehr@awi.uni-heidelberg.de, hannes.rau@awi.uni-heidelberg.de, stefan.trautmann@awi.uni-heidelberg.de, yilong.xu@awi.uni-heidelberg.de

1. Introduction

The recent surge of income and wealth inequality in many developed countries is a widely discussed topic in the media and academic research. Much of these discussions revolve around the gains of the top-income decile and the stagnation of income for the bottom half of the distribution and its implications for society (e.g., Piketty and Saez, 2003; Autor, Katz, and Kearney, 2008; Piketty, 2014; Piketty and Saez, 2014; Piketty, Saez and Zucman, 2018; World Inequality Report, Alvaredo et al., 2017). Indeed, inequality deriving from competitive economic environments is often associated with negative societal consequences (Stiglitz, 2012; Verhaeghe, 2014). In particular, it is sometimes conjectured that inequality may harm the social fabric, destroying social capital (trust, honesty, cooperation) and subsequently affecting economic outcomes (Wilkinson and Pickett, 2010). Two contested hypotheses can be derived from the literature in economics and the social sciences. The first hypothesis states that higher inequality, if perceived as unjust and caused by competition, hampers economic interaction (Alesina and Perotti, 1996; Bénabou, 1996; IPSP, 2017, Chapter 3; Camera, Deck and Porter, 2017). The second hypothesis states that those who are in an advantageous position (of higher status or wealth) in an unequal society, become self-focused and greedy (Piff et al., 2010; 2012).[1] That is, negative social consequences are caused by the egoistic behavior of the successful.

We study the impact of unjust inequality on subsequent social interactions, differentiating between the behavior of the economically successful and the unsuccessful, and thus aim to test both hypotheses within the same setting. Identifying these effects is a challenging task. Empirical assessments of the effects of inequality and the role of the successful typically suffer from an absence of counterfactuals and the endogeneity of status. We overcome these challenges by running a large-scale controlled experiment that allows us to introduce variation in both dimensions: variation in inequality and random assignment of interaction

[1] While this hypothesis received a lot attention in the public and academic sphere, the evidence is mixed (e.g., Erkal, Gangadharan, and Nikiforakis, 2011; Guinote et al., 2015; Martinsson et al., 2015; Nishi et al., 2015; Trautmann et al., 2013; Korndörfer et al., 2015; Smeets et al., 2015; Andreoni et al., 2017). We discuss this in more detail in the conclusions.

partners. Our approach thus complements results from less controlled settings by providing a clear identification of causal effects and underlying processes.

Our experiment consists of two stages. In the first stage, we create income inequality in dyads, using a real-effort task with varying payment schemes. In the second stage, we let these dyads interact in a modified trust game allowing us to measure both players' social trust and trustworthiness. Social trust has been interpreted as an important component of social capital in the literature (Glaeser et al., 2000; Bellemare and Kröger, 2007; Langer et al., 2017; Björnskov, 2019). As higher social capital is typically associated with better-functioning institutions and society in general (Putnam, 2000), social trust is a center piece in the debate on whether inequality erodes the social fabric.[2] In addition, our experimental measure for trustworthiness allows us to quantify subjects' greed or altruism absent strategic motives. It directly tests the hypothesis that higher inequality has a negative impact on social interactions because successful people become less generous, in particular less generous than the unsuccessful.

In many naturally occurring settings inequality arises from unequal opportunities and competition, for example, in school education, universities, workplaces or labor markets more generally.[3] The combination of unequal opportunities and competition can be seen as unfair from a normative perspective (Roemer, 1998) and is frequently blamed for negative consequences of inequality (Stiglitz, 2012). We focus on this kind of "unfair" inequality in investigating the impact of inequality on social interaction. To capture these two features,

---

[2] More precisely, social capital can be defined as values and shared beliefs that help groups to cooperate in situations where contracts are difficult or impossible to enforce (cp., Guiso, Sapienza, and Zingales, 2010). Accordingly, it is possible to measure social capital by eliciting values and beliefs using experimental tools (see e.g., Fehr 2009). In the economic literature social capital has been positively associated with a plethora of economic outcomes, such as economic growth (e.g., Knack and Keefer, 1997), the size of firms (e.g., Bloom, Sadun, and Van Reenen, 2012) or financial development (e.g., Guiso, Sapienza, and Zingales, 2004).

[3] For example, Lemieux, MacLeod, and Parent (2009) document an economy-wide increase of (relative) performance-pay jobs in the U.S. labor market, along with a substantial increase in wage inequality. Moreover, influential work by Chetty and colleagues document the persistence of social status and that better opportunities increase earnings (e.g., Chetty et al. 2011, 2017). The sentiment that inequality is unfair seems also commonplace in public opinion. According to an opinion poll by Pew Research, about 60-70% of Americans across all income levels say that the economic system unfairly favors the powerful (https://pewrsr.ch/2N9m0rB).

we use a relative-payment scheme in the real-effort task combined with an undue time advantage to one participant in the dyad. This particular payment scheme not only results in a more spread pay distribution (and thus more inequality) than the observed performance differences justify, but also undermines equality of opportunity, which together likely triggers strong feelings of injustice.

We then benchmark the trust-game outcomes in this high-inequality setting against the trust-game outcomes in a low-inequality setting, absent competition and unequal opportunities. To generate this low-inequality setting, we use a piece-rate payment scheme in the real-effort task. This benchmark maximizes the power to detect any effects of income inequality on social interaction and reveals the combined effect of high and unjust inequality that is characteristic for many economically relevant settings as noted above.[4] While we are explicitly interested in documenting the impact of unjust inequality on social interactions, we also probe the boundaries of this effect. For this purpose, we run a third condition with the same relative-payment scheme as above, but randomly rematching participants in the trust-game stage (keeping earnings information constant across conditions). This allows us to look at matches with equal and unequal outcomes arising from the *same unfair* institutional features of the real-effort task and eliminates the direct responsibility for each other's outcomes in the dyads.

As intended, we observe a strong "first-stage" effect of our treatment. The variation in payment schemes leads to substantial differences in inequality and to a significant polarization of the perceived fairness of the income-generating process. Inequality is considerably higher in the relative-payment scheme than in the piece-rate payment scheme. Similarly, and in line with normative views and evidence on fairness perceptions of various payment

[4] High inequality is typically not seen as negative per se. Mounting evidence suggests that people tend to accept more inequality if it is the result of work effort (e.g., Cappelen et al., 2007; Cappelen et al., 2013; Bartling et al., 2017; Bartling, Grieder, and Zehnder, 2017; Breza, Kaur, and Shamdasani, 2018) or more generally if they perceive it as the result of effort rather than luck (e.g., Fong, 2001; Alesina and La Ferrara, 2005; Alesina and Giuliano, 2011; Alesina, Stantcheva, and Teso, 2018; Cohn, Jessen, Klasjna, and Smeets, 2019; Fehr, Müller, and Preuss, 2020). However, if higher inequality cannot be clearly attributed to work effort and is possibly the result of immoral behavior, people accept less inequality and engage in more antisocial behaviour (Fehr, 2018; Bortolotti et al., 2017).

schemes (Fehr et al. 2020), subjects consistently perceive the relative-payment scheme as substantially less fair than the piece-rate payment scheme.

This first-stage effects translate into subsequent behavior and we present three main results. First, our findings support the view that unjust inequality can negatively affect social interactions. We document a significant decline in trust and trustworthiness when the income-generating process leads to more inequality that is perceived as unfair. While this finding offers a causal interpretation for correlational evidence from observational studies (e.g., Alesina and La Ferrara, 2002; Haushofer, 2013; Falk et al., 2018), we also move one step further and examine possible mechanisms underlying this result. We find that detrimental effects of unfair inequality are less severe if we break the direct link between the income-generating stage and the subsequent trust game stage by re-matching participants in the latter stage. This muted effect is tied to the behavior of successful participants, who hold more positive beliefs about the behavior of unsuccessful, whereas the beliefs of unsuccessful participants remain largely unchanged.

Second, taking advantage of varying inequality under the same unfair income-generating process in our third condition, we find that inequality on its own does not matter for behavior. Instead, we see that success in the relative-payment scheme results in higher trust and trustworthiness. This is particularly apparent when successful participants interact with other successful participants, whereas dyads of unsuccessful participants display very low levels of trust and trustworthiness. This suggests that despite equal outcomes experiencing unfairness per se seems to affect behavior with the potential to reinforce segregation and inequality.

Third, we find that successful players are more generous than the unsuccessful in absolute terms. This is in strong contrast to an influential literature claiming that the successful behave more selfishly (e.g., Piff et al. 2010, 2012). In general, our results indicate that most successful and unsuccessful players narrowly frame their distribution decisions in the trust game (i.e., they consider this decision in isolation and do not incorporate their first-stage earnings). In our setting, this is particularly evident for the successful, as they are always less generous relative to their total wealth than the unsuccessful. This highlights an important normative question about whether generosity should be judged in absolute or relative terms.

While third-party observers may take a broader perspective incorporating total wealth and thus setting a high normative bar, our stakeholders take a different and narrower perspective focusing primarily on the specific social interaction. In any case, that subjects seemingly frame equity concerns narrowly has significant implications, as such behavior can result in a more unequal society.

In the next section, we introduce the experimental paradigm and design of our study followed by a description of how we induce perceived unjust inequality. Section 3 shows that our experimental paradigm successfully induces inequality differences and a polarization of fairness perceptions. Section 4 presents the effects of unjust inequality in fixed dyads and Section 5 investigates the boundaries of the effects of unjust inequality. We discuss these results in the context of the related literature in section 6.

2. Experimental Paradigm and Design

We employ an experimental paradigm in which dyads of participants interact in two stages. In the first stage, we use a repeated real-effort task involving either an individual piece-rate payment inducing low inequality, or a relative-payment scheme (tournament) with preferential treatment of the initial tournament winner, resulting in high inequality (in a between-subjects design). In the second stage, participants then interact in a trust game. Consequently, we observe trust and trustworthiness depending on stage-1 conditions (low versus high inequality), and depending on stage-1 income. This allows us to gauge the impact of higher inequality on social interactions and how this relates to income. Full instructions are available in the Online Appendix.

2.1. Stage 1: Real-effort Task and Inequality Manipulation

Implementation: We implement a repeated real-effort slider task adapted from Gill and Prowse (2012) and vary the payment scheme to manipulate inequality, i.e., low inequality versus high inequality. In the slider task, participants see a number of sliders on their computer screen and have to adjust each slider to exactly the middle position within a certain time limit (see Section A.8. in the Online Appendix for a screenshot). The goal in this task is to maximize the number of correctly positioned sliders before the allotted time runs out.

Participants are only allowed to use their mouse to drag the sliders into the correct position. To prevent participants from using the arrow keys or the mouse wheel, we used a keyboard locker. The task requires little a-priori knowledge and skills such that outcomes mainly depend on the expended effort of subjects. Unfairness or concerns about unequal opportunities arise only through institutional features, i.e., the details of the two distinct payment schemes we implemented.

In the piece-rate payment scheme (low-inequality condition), participants complete four rounds of this task, each lasting for 120 seconds. In each round, they receive a flat payment of €0.50 plus €0.05 per correctly placed slider. Total earnings are calculated by summing up the earnings in the four rounds. Note that each subject in a dyad individually determines her own earnings, i.e., there is no interaction. However, at the end of each round both subjects in the dyad are informed about the correctly positioned sliders and the resulting earnings of each other. Thus, social comparison is salient in this setting.

In the relative-payment scheme (high-inequality condition), participants in a dyad also complete the slider task four times, with the first round lasting for 120 seconds. In contrast to the piece-rate payment scheme, participants' payoffs in each round are determined through their relative performance. That is, the subject with the higher number of correctly placed sliders in a round receives €3.00, while the subject with the lower number of correctly placed sliders receives €0.30. In the case of equal performance, the two payments are randomly allocated. As in the piece-rate payment scheme, participants receive information on the performance of each subject and the resulting payoffs after each round. To mimic unequal opportunities, we introduce a time bonus for the subject with the higher performance. More specifically, the winner of the first round obtains a time bonus of 8 seconds. The time bonus is subtracted from the time budget of the tournament loser such that the time for completing the second-round task is 128 (112) seconds for the first-round winner (loser). The winners of the second and third round get a time bonus of 6 and 4 seconds, respectively, which is again subtracted from the time budget of the loser in the respective round. The substantial time gap (16 seconds) that arises after the first round makes it nearly impossible for the first-round loser to catch up in the subsequent rounds.

Inequality and Fairness: Note that our relative-payment scheme includes two components – competition and unequal opportunities – that are deliberatively absent in the piece-rate payment scheme. These two components go often hand in hand in real-world settings, where initial (unfair) advantages are often amplified in competitive contexts, leading to enhanced inequality (e.g., Frank and Cook, 1995; Stiglitz, 2012).[5] For example, it is empirically well-established that being born rich is a gateway to better primary and secondary schooling and subsequently to higher-ranked colleges, resulting in better jobs and higher earnings (Chetty et al., 2011) as well as highly persistent socio-economic status (Chetty et al., 2017). As we are specifically interested in how this kind of inequality affects social interaction, we aimed at providing a strong first-stage to demonstrate potential detrimental effects of unjust inequality. Therefore, we stripped off competition and unequal opportunities in our baseline condition (piece-rate payment scheme) with the goal to generate substantial differences in inequality and polarized fairness perceptions.

In a companion paper, we demonstrate that the two payment schemes are indeed at the opposite end of the fairness-perception spectrum (Fehr et al., 2020). Using vignette surveys, we study how different aspects of incentive schemes, such as inequality, competition, unequal opportunities, and pay discontinuities influence fairness perceptions.[6] While all these aspects matter for fairness perceptions, the study reveals that the two schemes used in the current paper are the most extreme in terms of observed fairness judgments. Subjects perceived the low-inequality piece rate as the fairest procedure overall, whereas they considered the relative-payment scheme with time bonus as the least fair procedure.

To examine the robustness of potential effects of unjust inequality, we then take the relative-payment scheme and investigate social interaction after randomly rematching subjects. This allows us to examine two important aspects of our initial investigation of unjust inequality. First, randomly rematching subjects after the real-effort task results in matches

[5] A genuine feature of competition is that there are unequal rewards for winners and losers. Competition without consequences or unequal rewards are rarely seen in real-world settings and arguably of little relevance economically.

[6] We recruited 2,431 participants on the Amazon's Mturk platform and compare 16 vignettes of incentives schemes in a between-subject design (for more details see Fehr et al., 2020).

with equal and unequal incomes. We can thus use within-payment-scheme variation in inequality to examine social interaction holding fairness perceptions fixed (see Section 2.4 below for more details). Second, it breaks the link between the income-generation process and social interaction, i.e., inequality is not the byproduct of prior interaction.

2.2 Stage 1: Measurement of Fairness Perception

We measure subjects' fairness evaluations of the payment schemes to assess whether the piece rate versus tournament manipulation was successful in creating perceptions of unfair inequality. To gauge the impact of the procedures on participants, we measure fairness perceptions both before and after the stage-1 real-effort task. At the beginning of the experiment, participants receive detailed instructions about the stage-1 real-effort task and the payment procedures of their condition. They then answer three control questions about the procedure. Next, they are asked to indicate on a scale from 0 (very unfair) to 10 (very fair) how fair they consider the payment procedures in stage 1. They also indicate their gender, age, field of study, and risk preferences (qualitative measure on a scale from 0-10). Afterwards they start with the real-effort task.

The first assessment provides a fairness judgment based on a verbal description of the mechanism, absent any experience of the task and the outcomes. Our second measurement takes place immediately after the end of stage 1. Subjects have then completed four rounds of the real-effort task and received feedback on the number of correctly placed sliders and the corresponding payoffs of both subjects in the dyad. Thus, we can observe whether and how experiencing the task and the resulting feedback affects subjects' fairness evaluations. Because we collected fairness judgments in all conditions, demand effects cannot drive any observed treatment differences.[7]

2.3. Stage 2: Measurement of Social-Interaction Effects

In the second stage, we use a two-player trust game to measure the effects of the exogenous income variation on social interactions. In this game, we endow the first mover (trustor) with

[7] It is still conceivable though that fairness is made salient by asking. However, saliency of fairness issues is conducive to our goal of studying downstream behavior following allocations perceived as unfair.

€6 and the second mover (trustee) with €0. The first mover decides whether or not to transfer the endowment to the second mover. If she does not transfer, the game ends and the earnings will be €6 for the first mover and €0 for the second mover. In contrast, if she transfers her endowment, the experimenter triples the endowment such that the second mover receives €18 (and first mover has €0 now). The second mover then decides how much of the €18 to send back to the first mover (by the cent). Payoffs follow directly from the second mover's decision.

To obtain information on both decisions and the underlying processes, we use the strategy method. More precisely, we first elicit from each player in the dyad their decision as a first mover, and then their decision as a second mover conditional on having received a transfer (because otherwise there is no decision to be made). The player roles in the game are randomly determined after all decisions have been made and subjects are well aware of this fact. Therefore, this modification allows us to answer our first research question (i.e., the effect of inequality on trust in other individuals in a group; first mover) and the second research question (i.e., the greediness of individuals as a function of stage-1 income; second mover), within the same context.

We also measure participants' beliefs regarding the behavior of the other player in this stage. Specifically, we ask subjects to indicate whether they believe the other player in the dyad transferred her endowment when acting as a first mover (yes/no), and to indicate how much they think the other player sends back when acting as second mover (in six ranges: €0 to €3.00; €3.01 to €6.00; …; €15.01 to €18.00, see Section A.9. in the Online Appendix for the questions and screenshots). We do not incentivize beliefs because the preclusion of hedging opportunities would have required rather complex randomizations. Given the randomization in the implementation of the strategy method, we did not want to complicate matters further.

2.4. Treatments

We implemented three conditions. They all have in common that subjects received exactly the same information at the beginning of the trust-game stage, i.e., we informed them about their own and the other person's stage-1 earnings. Note that it is not possible to attribute

high or low stage-1 earnings to luck or effort as only earnings (but not effort) are communicated. In condition *Piece Rate* (first-stage piece rate – same partner) the piece-rate payment scheme determined subjects' earnings, whereas in *Tournament* (first-stage tournament – same partner) we used the relative-payment scheme. In both conditions, stage-1 dyads remain intact and proceed together to stage 2 to play the trust game as explained above. We emphasized at the very beginning of the experiment that subjects will interact with the same partner throughout the whole experiment and reminded subjects of this fact at the start of stage 2. Condition *Tournament-NEW* (first-stage tournament – new partner) is identical to *Tournament*, except that the dyads are re-matched in stage 2, such that each person will play the trust game with a person with whom she did not interact in stage 1, which was highlighted in the beginning of the experiment. At the beginning of stage 2, subjects were informed about the new match and they received information on their own and the other persons' (the new partner in the dyad) earnings from stage 1. This setting allows us to address two important issues. First, it precludes attributions of responsibility for each other's stage-1 outcomes (Lien et al., 2018). Second, we can observe social interactions in dyads with high and low inequality, which stemmed from the same unfair relative-payment scheme.

2.5. Procedural details and variable definitions

In total, 636 subjects took part in the experiment that was programmed using z-Tree (Fischbacher, 2007): 160 in condition *Piece Rate*, 134 in condition *Tournament*, and 342 in condition *Tournament-NEW*. The first two conditions were run on a subject pool at the Universities in Heidelberg and Mannheim (balanced across conditions). For condition *Tournament-NEW* we recruited 202 new subjects from the same subject pool, and additionally 140 subjects from the laboratory at the Technical University Berlin to increase power for the analysis of various subgroups of matching stage-1 winners and losers in this condition. In the Online Appendix (Section A.4.), we report robustness checks showing that our results and conclusions do not depend on the inclusion of the Berlin sample. Participants were undergraduate students from a wide range of different majors, who were recruited with ORSEE (Greiner, 2015) in Berlin and Mannheim and with Hroot (Bock et al., 2012) in Heidelberg.

In the Online Appendix (Section A.5.), we calculate the minimum detectable effect sizes (MDE) for our main results, that is, the effect that we can detect with 80% power at 5% significance level, given our sample size. For example, the MDE for comparing trust between *Piece Rate* and *Tournament* is 16 percentage points (see Figure A3). Similarly, for comparing trustworthiness between *Piece Rate* and *Tournament* the MDE ranges from €0.55 to €1.16 depending on the assumed standard deviation (see Figure A4).[8] This approach helps us to evaluate whether our sample sizes are sufficient to precisely distinguish the relevant treatment effect from zero.

At the beginning of a session we matched participants in same-gender dyads, with one mixed dyad if there was an uneven number of (fe)males. This was done based on the information about each subjects' gender from the initial questionnaire. Subjects were not aware of this matching procedure; they were only informed that they were matched with another person in the lab. We implemented this matching procedure in the background to control for possible gender differences in the performance in the multiple-round slider task (Gill and Prowse, 2014) and in the behavior in the trust game (Bellemare and Kröger, 2007). Final payoffs were determined by adding payoffs from both the real-effort stage and the trust game. A typical session lasted about 50 minutes, and subjects earned, on average about €13.40 (approximately $14.70 at that time), with final payoffs ranging from €1.20 to €30. There was no show-up fee in addition to the incentivized payoffs; that is, incentives were very salient.

In the presentation of the results we use the following conventions. In the fixed dyads conditions *Piece Rate* and *Tournament* we will call the person with the higher income in a dyad "successful" and the person with the lower income "unsuccessful." In the *Tournament-NEW* condition, participants encounter new partners, leading to various matches based on the stage-1 income. As in the other conditions, we classify participants based on their relative performance in stage-1 and refer to the person with the higher stage-1 income in a dyad as

[8] As our trust outcome is binary, we can simply calculate the MDE by using the proportion of trust in one condition and using the sample sizes in the two groups. For trustworthiness we need to make assumptions on the standard deviation and thus base the calculations on results from a Meta study of trust games (Johnson and Mislin, 2011) and the observed standard deviation. For more details, see Online Appendix A.5.

"successful" and the person with the lower stage-1 income as "unsuccessful." In a few cases there are no stage-1 income differences within a dyad, and thus we cannot apply this classification. Consequently, we ignore these cases (*n=4* in *Piece Rate*, *n=8 in Tournament*, and *n=12 in Tournament-NEW*) when analyzing the question whether negative social consequences are driven by the successful. In the Online Appendix A.6., we present an alternative classification of successful and unsuccessful that keeps income differences (or inequality) constant. This classification only considers subjects with the highest (€12) and lowest (€1.2) stage-1 income and we show that our results are robust to this stricter classification.

3. First-Stage Results: Income Inequality Manipulation

We start by looking at effort levels, i.e., the number of correctly positioned sliders, in the three conditions. The *Piece Rate* and *Tournament* conditions did not result in different levels of effort with an average number of correctly solved sliders of 75 in *Piece Rate* and 76 in *Tournament* in all four rounds (p=0.77, Mann-Whitney U test). Effort in *Tournament-NEW* was somewhat higher at 81 compared to *Tournament* (p=0.03, Mann-Whitney U test). Importantly though, the average difference in effort levels between the two players in a dyad in the first slider task does not differ in all three conditions (4.4 in *Piece Rate*, 4.9 in *Tournament*, and 4.9 in *Tournament-NEW*, Mann-Whitney U tests, all p>0.31) indicating that the higher overall effort level in *Tournament-NEW* is not solely driven by the successful. In the Online Appendix, we show that there is substantial variation in effort levels in each condition (Figure A1) and that the dyads are balanced on a limited set of observables (Table A1), which together suggests that winning the first round is unlikely non-random (conditional on exerting effort).

Table 1: Stage-1 Earnings

	Piece Rate	*Tournament*	*Tournament-NEW*
Earnings: mean	5.77	6.60	6.60
Earnings: median	5.75	6.60	6.60
Earnings: 10% percentile	4.93	1.20	1.20
Earnings: 90% percentile	6.70	12.00	12.00

Notes: Entries are in €.

Table 1 displays stage-1 earnings and shows that the tournament condition has the intended effect on inequality (see Figure A2 in the Online Appendix for the distribution of earnings in all three conditions). While average earnings are comparable across the different treatments, the variation in earnings is much larger in *Tournament* and *Tournament-NEW* than in *Piece Rate*. That is, small initial differences in effort translate into vast income inequality in *Tournament* and *Tournament-NEW*, but not in *Piece Rate*.

Table 2: Fairness Evaluation of Payment Mechanism

Point of evaluation	Evaluators	*Piece Rate*		*Tournament*		*Tournament-NEW*	
Before experience	All	7.17	(n=160)	3.69***	(n=134)	3.91***	(n=342)
After experience	All	6.78$^{\wedge\wedge\wedge}$	(n=160)	2.44$^{\wedge\wedge\wedge,***}$	(n=134)	2.90$^{\wedge\wedge\wedge,***}$	(n=342)
After experience	Successful	7.32	(n=78)	2.98***	(n=63)	3.57***	(n=144)
	Unsuccessful	6.36$^{\#\#}$	(n=78)	1.92$^{\#\#\#,***}$	(n=63)	2.00$^{\#\#\#,***}$	(n=144)

Notes: Entries are fairness ratings ranging from 0 (perceived as very unfair) to 10 (perceived as very fair); *,**,*** indicates significant difference between *Piece Rate* and *Tournament* conditions; #,##,### indicates significant difference between successful and unsuccessful; and ^,^^,^^^ indicates significant difference between evaluation before and after experience; at the 10%, 5%,1% level, Mann-Whitney U test; pairs with equal earnings excluded in analyses of successful and unsuccessful. Note that there are no statistically significant differences between entries for *Tournament* and *Tournament-NEW* (this is robust to restricting to the same locations).

It is conceivable that subjects perceive the high reward for the tournament winner as justified, taking a meritocratic perspective and focus on incentives for performance (see e.g., Cappelen et al., 2007; Bartling et al. 2018). This is not what happens in the current context. Table 2 shows that participants perceive the tournament mechanism as substantially less fair than the piece-rate mechanism. We observe strong treatment differences both before and after the experience of the task and for both the successful and the unsuccessful: the piece-rate scheme always receives much higher fairness evaluations than the two tournament schemes. Experiencing the task leads to lower evaluations compared to the mere verbal description in all three conditions, and in particular, we observe that the unsuccessful perceive the task as less fair than the successful.

We conclude that the first stage succeeded in inducing strong differences in income inequality and fairness perceptions across the three conditions. Moreover, successful and unsuccessful subjects strongly differ in their fairness perceptions, reflecting a self-serving bias that might have lead the successful to perceive the procedures and resulting positional differences as more justifiable than the unsuccessful. This also applies to the successful in the two tournament conditions, despite evaluating the procedure as unfair. Note that self-serving fairness judgments by the successful should facilitate self-centered behavior in the trust game.

4. Results: Social Interaction Effects for Fixed Dyads

4.1. Main Effects

We now turn to the analysis of whether the strong first-stage effects in inequality and fairness perception between *Piece Rate* and *Tournament* affect behavior in the stage-2 trust game. Table 3 shows our main results. We observe strong treatment effects, with the share of trusting participants (i.e., transferring their endowment to the second mover) being almost 20 percentage points lower in *Tournament* than in Piece rate (top panel, Table 3). Trust is significantly lower in *Tournament* for both the successful and the unsuccessful. However, we do not detect significant differences in trust between these subgroups within either condition.

Result 1*: Unjust inequality in stage 1 is detrimental for social trust in stage-2 interaction for fixed dyads.*

The bottom panel of Table 3 shows the amounts returned by the second mover. Remember that there are no strategic considerations at this stage and that these amounts are conditional on the trust decision of the first-mover resulting in a budget of €18 for the second mover and €0 for the first-mover. We observe that amounts returned are almost €1 lower in the *Tournament* than in the *Piece Rate* condition (6.4 vs 5.5). Thus, transferring the budget implies an expected loss for the first-mover in *Tournament*. This effect is mainly driven by the behavior of the unsuccessful stage-1 subjects. While there is no difference in the amounts returned across conditions for the successful, the stage-1 losers strongly reduce these

amounts in *Tournament*. Consequently, amounts returned are significantly lower for the unsuccessful than for the successful in *Tournament*.

Table 3: Social Interaction Effects of Payment Mechanism

	Participants	*Piece Rate*		*Tournament*	
Trusting	All	71%	(n=160)	53%***	(n=134)
	Successful	71%	(n=78)	49%**	(n=63)
	Unsuccessful	71%	(n=78)	56%*	(n=63)
Amount returned	All	€6.41	(n=160)	€5.50**	(n=134)
	Successful	€6.30	(n=78)	€6.10	(n=63)
	Unsuccessful	€6.55	(n=78)	€4.65$^{\#\#,***}$	(n=63)

Notes: *,**,*** indicates significant difference between treatment; #,##,### indicates significant difference between successful and unsuccessful; at the 10%, 5%,1% level (Fisher's exact test for trust and Mann-Whitney U-test for amounts returned). The total number of observations is *n=160* in *Piece Rate* and *n=134* in *Tournament* and we exclude pairs with equal earnings in the analysis of successful and unsuccessful (*n=4* in *Piece Rate* and *n=8* in *Tournament*).

While reduced trustworthiness (generosity) affects the distribution of trust game earnings resulting in a higher variance and skewness, reduced trust affects overall welfare because of the inefficiency of forgoing the tripled payoffs after transfer. Indeed, we observe that the welfare effects are substantial. The average earnings in the trust game are €1.08 lower in the *Tournament* condition (€7.28 vs. €6.04), a 17% loss compared to the *Piece Rate* condition. This corresponds to a 14 percentage point lower efficiency in the *Tournament* than in the *Piece Rate* condition (significant at the 1% level, Mann-Whitney U-test).[9]

Result 2*: Unjust inequality in stage 1 is detrimental for absolute generosity in stage-2 interaction for fixed dyads, which is driven by lower generosity of the unsuccessful.*

In general, the observed returns in the trust game indicate hardly any attempts to equalize overall experimental wealth in our setting in both *Piece Rate* and *Tournament*. It appears that participants narrowly frame their decisions in stage 2 and do not fully take stage-1 income

[9] We calculate efficiency as the ratio between the realized earnings in a dyad and the total pie size if the money is transferred (€18). Accordingly, the efficiency in terms of how much of the total pie is distributed is 81% in *Piece Rate* and 67% in *Tournament*.

into consideration when deciding about how much to return to the trustor. This raises an interesting normative aspect of the observed generosity. Although the successful in *Tournament* are more generous in the trust game than the unsuccessful, they fall short of relevant normative benchmarks. First, they give less than the unsuccessful relative to their total wealth. Second, in spite of having typically earned €12 in stage 1 (vs. €1.20 for their partner), only about 38 percent of the successful share their stage-2 income equally (return €9), while they are far from sharing overall income equally (return €14.40; only 2 percent share more than €12). However, failure to meet such normative criteria is not restricted to participants in *Tournament*. In *Piece Rate*, stage-1 payoff differences are modest in most dyads, and the majority of the successful and the unsuccessful do not share their income equally (about 45 percent return €9 and 2 percent more than €9) and thus fail to equalize their overall earnings. The observer's higher normative expectations towards the stage-1 winners make this behavior look less acceptable for the successful in *Tournament.* Yet, conditional on a narrow framing of decisions in stage 2, the successful return more than the unsuccessful in *Tournament* in absolute and relative terms.

4.2. The Role of Beliefs

In stage 2, we measured subjects' beliefs regarding the other player's behavior as a trustor and as a trustee in a dyad. In the Online Appendix (Table A2), we show that the *Tournament* condition induces more pessimistic beliefs regarding both trust and amounts returned. These effects are mainly driven by the successful subgroup when differentiating by stage-1 outcome. That is, the stage-1 condition affects subjects' beliefs. In Tables 4 and 5 we investigate whether these beliefs can explain the treatment effects on trust and trustworthiness. We present four specifications: specifications 1 and 2 verify the raw comparisons discussed above including various controls, and specifications 3 and 4 include beliefs about trust and trustworthiness. All specifications include controls for gender, location, and session size.

We find a clear correlation between beliefs and behavior. For trust, beliefs about the other person's trust and her trustworthiness relate to higher trust (Table 4). The latter effect makes sense from a strategic point of view (expecting lower returns on trust), while the former effect suggests a conditionally-cooperative or reciprocal view (conditioning on behavior

Table 4: Determinants of Trust

Dependent variable: Transfer (yes/no) to second mover				
	(1)	(2)	(3)	(4)
Tournament	-0.178***	-0.147*	-0.133**	-0.131
	(0.06)	(0.08)	(0.06)	(0.09)
Successful		0.010		0.070
		(0.08)		(0.09)
Tournament × Successful		-0.076		-0.006
		(0.12)		(0.13)
Belief in trust by other			0.428***	0.411***
			(0.06)	(0.06)
Belief in amount returned by other			0.046***	0.047***
			(0.01)	(0.01)
N	294	282	294	282
Joint effect of tournament variable		χ=9.67, p<0.01		χ=4.05, p=0.132

Notes: Marginal effects from Probit regressions with robust standard errors in parentheses. Linear regressions support the sign of the interaction terms in the Probit regressions. Belief in amount returned by other scaled to 100 cents. All regressions control for gender, session size and location.

Table 5: Determinants of Amounts Returned

Dependent variable: Amount returned in cents				
	(1)	(2)	(3)	(4)
Tournament	-101*	-198***	-40	-166***
	(53.20)	(73.26)	(47.50)	(62.07)
Successful		-8		17
		(66.61)		(54.83)
Tournament × Successful		167		253***
		(109.23)		(90.55)
Belief in trust by other			13	22
			(50.00)	(50.58)
Belief in amount returned by other			77***	84***
			(10.45)	(9.75)
N	294	282	294	282
Joint effect of tournament variable		F=3.70, p=.026		F=4.59, p=.011

Notes: Tobit regressions with robust standard errors reported in parentheses. Linear regressions support the sign of the interaction terms in the Tobit regressions. Belief in amount returned by other scaled to 100 cents. Amounts are coded in cents. All regressions control for gender, session size and location.

if the other person were in the trustor's position). Results on trustworthiness support the reciprocal view as well (Table 5). Higher beliefs on amounts returned by the other player relate to higher amounts returned. Because strategic aspects are absent for the second mover, beliefs about the other person's returns can only play a role in terms of reciprocal thinking. Note that while beliefs play a role for both trustor and trustee behavior, the main treatment effects of the *Tournament* condition remain substantial when including the beliefs. That is, beliefs cannot fully explain the effect of unjust inequality on social interactions.

5. Results: Social-Interaction Effects in New Dyads

The previous analysis has revealed strong detrimental effects of unjust inequality on social interactions. In this section, we test the boundaries of this effect by rematching subjects into new dyads in stage 2 (*Tournament-NEW*). First, while the experience and perception of competition and unjust inequality is identical to the *Tournament* condition (see Section 3, Table 2 results), a direct attribution of "responsibility" for the mutual stage-1 outcomes is absent in this condition.[10] Second, the rematching of dyads allows us to examine the impact of inequality in an environment that is perceived as unfair, since a player's own income and the income of the matched partner are not perfectly correlated as it is the case in the *Tournament* condition. That is, the main focus of the analysis of the *Tournament-NEW* condition is the within-condition comparison of different matching groups.

To put the condition in perspective, we first compare behavior in *Tournament-NEW* (trust = 65%; amount returned = €6.61) and *Tournament* (trust = 53%; amount returned = €5.50). Running simple Probit/Tobit regressions with a treatment dummy and controlling for standard covariates as well as locations, we find that trust and trustworthiness are sig-

[10] A negative attribution of high stage-1 earnings for the successful to undeserved luck also becomes more difficult as effort information on the stage-1 dyad is not available (see Section 2.4). König-Kersting et al. (2017) find that outcome information biases the perception of the underlying process ("outcome bias"), which is mainly driven by positive random outcomes being falsely attributed to the decision maker's skill. If this effect transfers to the current setting, we expect that good stage-1 outcomes should more likely be attributed to skill, rather than luck, by stage-2 players.

nificantly higher in *Tournament-NEW* than in *Tournament* (*p=0.016* and *p<0.001*, respectively). The results are robust to restricting the analysis to the same locations (i.e., excluding the Berlin sample).[11]

Table 6: Social Interaction Effects – *Tournament-NEW*

	Participants			vs. successful partner		vs. unsuccessful partner	
		(1)		(2)		(3)	
Trusting	All	65%	(n=342)				
	Successful	66%	(n=165)	68%	(n=82)	64%	(n=81)
	Unsuccessful	64%	(n=165)	67%	(n=81)	62%	(n=76)
Amount returned	All	€6.61	(n=342)				
	Successful	€7.22	(n=165)	€7.43	(n=82)	€7.08	(n=81)
	Unsuccessful	€6.03***	(n=165)	€5.94***	(n=81)	€6.10	(n=76)

Notes: *,**,*** indicates significant difference between successful and unsuccessful, and #,##,### indicates significant difference between successful partner and unsuccessful partner, at the 10%, 5%, 1% level (Fisher's exact test for trust, and Mann-Whitney U test for amounts returned). Unclassified participants (*n=12*, i.e., subjects in dyads with equal stage-1 incomes) are excluded when conditioning on successful and unsuccessful decision maker or successful and unsuccessful partner. This leads to different number of observations across cells, depending on stage-2 matches with unclassified subjects.

Next, Table 6 shows detailed results for Trust and for Amounts Returned in *Tournament-NEW*, separately for successful and unsuccessful decision makers, and successful and unsuccessful partners in the dyad. The upper panel of Table 6 shows trust behavior. There are no significant raw differences in trust between the successful and the unsuccessful (column 1), and neither between situations interacting with a successful partner (column 2), and an unsuccessful partner (column 3). However, there is a tendency to trust the unsuccessful less and also for the unsuccessful to trust less. Accordingly, trust within dyads of unsuccessful participants is lower than trust within dyads of successful participants (62% vs. 68%, Fisher's exact test, p=0.41). This is particularly evident if we only focus on matches with

[11] All results in this section remain qualitatively identical if we exclude the Berlin sample in *Tournament-NEW*, but some of the within-treatment comparisons of successful and unsuccessful become insignificant due to smaller sample sizes in subgroups. Details are available in Section A.4. in the Online Appendix.

equal incomes, in which trust in dyads of successful participants is 71% and in dyads of unsuccessful participants 55% (see Section A.6. in the Online Appendix). Moreover, regressions reveal that the successful are more likely to trust others than the unsuccessful (see Table 7, column 2).

The lower panel of Table 6 shows that stage-1 winners are significantly more generous as second movers than stage-1 losers are. This is mostly driven by the interactions with other successful people, though qualitatively the same when interacting with the unsuccessful. As in the case of trust, these effects lead to an overall large difference of generosity between pairs of unsuccessful people and pairs of successful people (€6.10 vs. €7.43, Mann-Whitney U test, p=0.03).

That dyads of stage-1 losers perform very poorly in terms of trust and trustworthiness suggests that the detrimental effect of inequality on trust and trustworthiness is not driven by inequality within dyads per se. Moreover, because of the low trust and trustworthiness in the interaction between the unsuccessful, stage-2 inequality is larger, and stage-2 welfare is lower in this group compared to interactions between the successfully. The expected earnings in the trust game are €0.36 lower in the unsuccessful dyads compared to the successful dyad (€6.72 vs. €7.08), a 5% loss. This corresponds to a 4 percentage point lower efficiency (7.08*2/18-6.72*2/18) in the unsuccessful group. As in the case of trust, a regression analysis shows that the successful return significantly higher amounts in the trust game (Table 7, columns 3 and 4).

Result 3*: The detrimental effects of unjust inequality on social interactions are dampened in newly assembled dyads. Negative effects derive mainly from interactions among the unsuccessful.*

A closer look at participants' beliefs offers an explanation for the differences in trust game behavior between *Tournament* and *Tournament-NEW*. Table 7 shows that the effect of beliefs on trust and amounts returned emerge in *Tournament-NEW* just as in *Tournament*. However, while in *Tournament* there were substantial negative effects of the stage-1 interaction on beliefs, especially for the successful, there are no such negative effects in *Tournament-NEW* (see Online Appendix, Table A3). Moreover, in *Tournament-NEW* the successful

stage-1 players tend to hold more positive views than the unsuccessful ones, especially when paired with another successful person.

Table 7: Determinants of Trust and Amounts Returned – *Tournament-NEW*

	Trust	Trust	Amounts Returned	Amounts Returned
Successful	0.03	0.11*	146***	122***
	(0.05)	(0.06)	(48.62)	(38.58)
Successful Partner	0.05	-0.02	15.29	-36
	(0.05)	(0.06)	(49.42)	(37.33)
Belief in trust by other		0.598***		95*
		(0.05)		(50.37)
Belief in amount returned by other		0.045***		84***
		(0.01)		(10.80)
N	320	320	320	320

Notes: Marginal effects from Probit regressions for Trust and Tobit regressions for Amounts Returned with robust standard errors reported in parentheses. Amounts Returned are coded in cents. All regressions control for gender, session size, inequality, and location. Belief in amount returned by other scaled to 100 cents (so does the corresponding standard errors).

We also observe that social aspects must be relevant for the observed effects. That is, the negative effects for the unsuccessful stage-1 dyads cannot simply derive from higher risk aversion caused by their lower income. We observe negative effects for the unsuccessful in both trust (potentially affected by risk attitude) and the non-strategic behavior as second mover. Moreover, in the comparison between *Piece Rate* and *Tournament*, in which unsuccessful players were always matched with the successful, we observe no differences in trust. Moreover, recent literature suggests that, if relative position is salient, inequality may lead the poor to take higher levels of risk (Payne et al., 2017; Fehr and Reichlin 2020). We therefore interpret our results in terms of reduced levels of social capital within groups of unsuccessful subjects, rather than in terms of risk attitudes.

6. Discussion

We present an experiment aimed at studying the effects of unjust economic inequality and the potential role of the economically successful in harming the social fabric, in a controlled lab setting. While there is no trivial mapping of laboratory results to behavior in society at

large, the findings we report in this study are consistent with a host of evidence from less controlled settings as we will discuss below. Importantly, our study provides insights into the underlying mechanisms, in particular by analyzing the behavior of the successful and unsuccessful as well as the impact of beliefs about others' behavior.

Our finding that unjust inequality has substantial effects on trust and trustworthiness supports the view that such an environment might be detrimental to social interactions, well-being, and more generally to social capital (Kawachi et al., 1997; Verhaeghe, 2014; Buser and Dreber, 2016). While this is consistent with results from observational studies suggesting that lower trust is associated with poverty within countries and higher inequality across countries (Alesina and La Ferrara, 2002; Haushofer, 2013; Falk et al., 2018), our controlled setting offers insights into why this might be the case. The increased pessimism about others' willingness to cooperate and the lower willingness to take the social risk of trusting a stranger is indicative for a decline in social capital. Indeed, we find that beliefs are correlated with behavior and that they are significantly more pessimistic when experiencing unjust inequality. As a consequence, a vicious cycle of decreasing trust and cooperation may result, leading to a substantial loss of social capital.

A large literature in psychology has argued that rich, high-status individuals are less generous in *absolute* terms than poor, low-status individuals (e.g. Piff et al., 2010, 2012; Guinote et al., 2015, Nishi et al. 2015).[12] In particular, this literature makes the causal claim that increasing wealth or status induces less social behavior. In correlational field data, the existence of a negative correlation between status and prosocial behavior has been questioned (Trautmann et al., 2013, Gsottbauer et al., 2020), and various studies have recently shown that wealthy individuals are often more prosocial and more generous in absolute terms (e.g. Andreoni et al., 2017; Korndörfer et al., 2015; Smeets et al., 2015). Our results confirm this

[12] There are also some economic papers showing that "wealthy" lab subjects give and cooperate less than "poorer" subjects (e.g., Erkal et al 2011; Martinsson et al. 2015). In Table A7 in the Online Appendix, we summarize a larger set of laboratory studies that relate to the question of the impact of inequality and status on cooperation and trust. In contrast to our study, the majority of these studies generates inequality by randomly assigning unequal endowments or by priming. Consequently, we see a mixed picture about the impact of inequality and social status groups on social interactions suggesting that differences in the implementation of inequality are important for the relevant channel driving social capital effects.

field data in a controlled environment: behavior of the successful is indeed not responsible for the negative consequences of unjust inequality on social interactions. Thus, we find no evidence for the hypothesis that the behavior of the successful is mainly responsible for the erosion of the social fabric.

However, we also find that most people frame fairness narrowly (see Exley and Kessler, 2019 for related evidence). As a consequence, despite being more generous in absolute terms, the successful appear more selfish from a broadly framed fairness perspective including all possible benefits. If the broad frame is normatively more compelling, the successful are likely to fall short of the potential normative expectations we may hold with respect to their behavior. This is not the case for the poor, simply because expectations are lower. Such an expectation-behavior gap for the rich may explain the appeal of picturing elites as immoral and selfish in popular discourses, which were eager to pick up the results supporting the view of the selfish elite.

The decline in trustworthiness that we observe is mostly driven by the less well-off. This is consistent with field data on deprived neighborhoods in the UK. Compared to wealthy neighborhoods, social capital is lower in deprived neighborhoods, measured by interactions among people in the same neighborhood and thus social class (Nettle et al., 2011). Jachimowicz et al. (2018) similarly report negative effects on the social fabric among low income households. Our results suggest that these field data may not simply derive from selection of people in or out of certain neighborhoods, but are caused by the local environment itself. That is, betraying the other player's trust is the only way for unsuccessful players to get out of the low-income environment. This is consistent with a prominent finding in social psychology that members of low-status status groups aim to move out of these groups if their group status cannot easily be increased (Tajfel and Turner, 1979).

We also show the boundaries of the detrimental effects of unfair inequality. Negative effects on trust and trustworthiness are muted if the interaction partner has not directly contributed to the existing income inequality within a dyad. This happens despite the fact that subjects perceive the inequality-generating process as equally unfair in the two *Tournament* conditions. This local effect hints at the volatility of the subtle psychological effects caused by inequality or fairness cues. Moreover, our main treatment (*Tournament*) resulted

in a strong divergence of inequality and polarization of fairness perceptions. We have argued that this key feature of our setup is relevant in many contexts outside the lab such as in educational systems, labor markets or one's social environment (e.g., Chetty et al., 2011; Chetty, Hendren, and Katz, 2016; Hanushek and Woessmann, 2006; Lemieux, MacLeod, and Parent, 2009). The more modest inequality emerging in the ideal-prototype condition *Piece Rate* is perceived as fair and allows players to maintain a high level of trust and trustworthiness. The perceived justice of the institution producing unequal outcomes thus seems to constitute an essential aspect, lending support to the conjecture that it is not inequality per se that bothers people in life, but economic unfairness (Starmans et al. 2017). Indeed, dyads of unsuccessful participants in *Tournament-NEW* score low on trust and trustworthiness despite having equal outcomes; their experience of disadvantages caused by unfair economic allocations seems to affect behavior, rather than the experience of inequality per se.

References

Alesina, A., and Giuliano, P. (2011). Preferences for redistribution. In *Handbook of Social Economics*, edited by J. Benhabib, A. Bisin, and M. Jackson, 93–131. North Holland: Elsevier.

Alesina, A. and E. La Ferrara (2002). Who trusts others? *Journal of Public Economics*, 85, 207-234.

Alesina, A., and La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics* 89 (1), 897–931.

Alesina, A. and R. Perotti (1996). Income distribution, political instability, and investment. *European Economic Review,* 40, 1203–1228.

Alesina, A., Stantcheva, S. and E. Teso (2018). Intergenerational mobility and preferences for redis-tribution. *American Economic Review*, 108(2), 521–554.

Alvaredo, F., L. Chanel, T. Piketty, E. Saez, G Zucman, eds. (2017). *World Inequality Report 2018*, Paris.

Andreoni, J., N. Nikiforakis, and J. Stoop (2017). Are the rich more selfish than the poor, or do they just have more money? A natural field experiment. *NBER Working paper No. 23229.*

Autor, D., L. Katz, and M. Kearney (2008). Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics, 90,* 300–323.

Bartling, B., M. Grieder, and C. Zehnder (2017). Competitive Pricing Reduces Wasteful Counterproductive Behavior. *Journal of Public Economics*, 156, 34–47.

Bartling, B., A. Cappelen, M. Ekström, E. Sørensen, and B. Tungodden (2017). Fairness in Winner-Take-All Markets. *Working paper, NHH Bergen*.

Bellemare, C. and S. Kröger (2007). On Representative Social Capital. *European Economic Review,* 51, 183–202.

Bénabou, R. (1996). Inequality and growth. In: B. Bernanke and J. Rotemberg (eds.), *NBER Macroeconomics Annual 1996*. Cambridge, MA: MIT Press.

Björnskov, C. (2019). The Political Economy of Trust. In: R. Congleton, B. Grofman and S. Voigt (eds.), *Oxford Handbook of Public Choice*, 1, 628-648. Oxford: Oxford University Press.

Bloom, N.; R. Sadun, and J. Van Reenen (2012). The Organization of Firms Across Countries, *Quarterly Journal of Economics*, 127(4), 1663–1705.

Bock, O., A. Nicklisch, and I. Baetge (2014). hroot: Hamburg Registration and Organization Online Tool. *European Economic Review,* 71, 117–120.

Bortolotti, S., I. Soraperra, M. Sutter, C. Zoller (2017). Too Lucky to be True Fairness Views under the Shadow of Cheating, *CESifo Working Paper*.

Breza, E., S. Kaur, and Y. Shamdasani (2018). The Morale Effects of Pay Inequality. *Quarterly Journal of Economics*, 133(2), 611-663.

Buser, T. and A. Dreber (2016). The Flipside of Comparative Payment Schemes. *Management Science,* 62, 2626–2638.

Camera, G., C. Deck, and D. Porter (2017). Do Economic Inequalities Affect Long-Run Cooperation? *Experimental Economics, forthcoming*.

Cappelen, A., A. Drange, E. Sorensen, and B. Tungodden (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review,* 97(3), 818–827.

Cappelen, A., J. Konow, E. Sorensen, and B. Tungodden (2013). Just luck: An experimental study of risk-taking and fairness. *American Economic Review,* 103, 1398-1413.

Chetty, R., J. Friedman, N. Hilger, E. Saez, D. Schanzenbach, and D. Yagan (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR, *Quarterly Journal of Economics*, 126(4), 1593-1660.

Chetty, R., N. Hendren, and L. Katz (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment, *American Economic Review*, 106(4), 855-902.

Chetty, R., D. Grusky, M. Hell, N. Hendren, R. Manduca, and J. Narang (2017). The fading American dream: Trends in absolute income mobility since 1940. *Science*, *356*(6336), 398–406.

Cohn, A., Jessen, L. Klasjna, M., and Smeets, P. (2019) Why Do the Rich Oppose Redistribution? An Experiment With America's Top 5%. *Working paper*.

Erkal, N., Gangadharan, L., and Nikiforakis, N., (2011). Relative Earnings and Giving in a Real-Effort Experiment, *American Economic Review*, 101, 3330-3348.

Exley, C. and J. Kessler (2019). Equity Concerns are Narrowly Framed. *Working paper HBS*.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics*, 133 (4), 1645–1692.

Fehr, D. (2018). Is Increasing Inequality Harmful? Experimental Evidence. *Games and Economic Behavior,* Vol 107, 123-134.

Fehr, D., and Reichlin, Y. (2020). Wealth Inequality, Personality and Risk Taking. *Working Paper, University of Heidelberg.*

Fehr, D., D. Müller, and M. Preuss (2020). Social Mobility Perceptions and Inequality Acceptance, *Working Paper, University of Innsbruck*.

Fehr, D., H. Rau, S. T. Trautmann, and Y. Xu (2020). Fairness properties of compensation schemes. *Working paper, University of Heidelberg*.

Fehr, E. (2009). On The Economics and Biology of Trust. *Journal of the European Economic Association,* 7, 235–266.

Fischbacher, U. (2007). Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics,* 10, 171–178.

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics*, 82, 225-246.

Frank, R. H. and P. J. Cook (1995). *The Winner-Take-All society: Why the Few at the Top Get So Much More Than the Rest of Us*. Penguin Books.

Gill, D., and V. Prowse (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review,* 102, 469–503.

Gill, D., and V. Prowse (2014). Gender differences and dynamics in competition: the role of luck. *Quantitative Economics,* 5, 351–376.

Glaeser, E, D. Laibson, J. Scheinkman, and C. Soutter (2000). Measuring Trust. *Quarterly Journal of Economics,* 115, 811–846.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association,* 1(1), 114–125.

Gsottbauer, E., D. Müller, S. Müller, S. Trautmann, G. Zudenkova (2020). Social class and (un)ethical behavior: Causal versus correlational evidence. *Working Paper 2020-10, University of Innsbruck.*

Guinote, A., I. Cotzia, S. Sandhu, and P. Siwa (2015). Social status modulates prosocial behavior and egalitarianism in preschool children and adults. *Proceedings of the National Academy of Science,* 112, 731–736.

Guiso, L., P. Sapienza and L. Zingales (2004). The Role of Social Capital in Financial Development, *American Economic Review*, 94 (3), 526-556.

Guiso, L., P. Sapienza and L. Zingales (2010). Civic Capital as the Missing Link, in: *Handbook of Social Economics*, edited by Benhabib J., A. Bisin and M. Jackson, Oxford: Elsevier Science.

Hanushek, E, and L. Woessmann (2006). Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries, *Economic Journal*, 116 (510), C63-C76.

Haushofer, J. (2013). The Psychology of Poverty: Evidence from 43 Countries. *Mimeo*.

IPSP International Panel on Social Progress Report (2017). Chapter 3 – Inequality and Social Progress.

Jachimowicz, J. M., B. Szaszi, M. Lukas, D. Smerdon, J. Prabhu, and E. Weber (2018). The Rising Tide Sinks the Smallest Boats: Why Higher Economic Inequality Intensifies the Poor's Financial Hardship. *Working paper, Columbia Business School.*

Johnson, N. D., and Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, *32*(5), 865-889.

Kawachi I., B. Kennedy, K. Lochner and D. Prothrow-Stith (1997). Social Capital, Income Inequality and Mortality. *American Journal of Public Health,* 87, 1491–1498.

Knack, S., and P. Keefer (1997), Does Social Capital Have an Economic Payoff? A Cross-Country Investigation‖, *Quarterly Journal of Economics*, 112 (4), 1252–88.

Korndörfer, M., Egloff, B., Schmukle, S. C. (2015). A large scale test of the effect of social class on prosocial behavior. *PLoS One*, 10, e0133193.

König-Kersting, C., M. Pollmann, J. Potters, and S. T. Trautmann (2017). Good decision vs. good results: Outcome bias in the evaluation of financial agents. *Working paper, Tilburg University*.

Langer, A., F. Stewart, K. Smedts, and L. Demarest. (2017). Conceptualising and Measuring Social Cohesion in Africa: Towards a Perceptions-Based Index. *Social Indicators Research*, 131, 321–343

Lemieux, T., B. MacLeod, and D. Parent (2009). Performance Pay and Wage Inequality, *Quarterly Journal of Economics*, 124 (1), 1–49.

Lien, J. W., J. Zheng, and Y. Zhuo (2018). The Cooperative Consequences of Contests. *Working paper, CU Hong Kong*.

Martinsson, P., C. Villegas-Palacio, and C. Wollbrant (2015). Cooperation and social classes: evidence from Colombia. *Social Choice and Welfare* 45, 829–848.

Nettle, D., A. Colleony, and M. Cockerill (2011). Variation in Cooperative Behavior within a Single City. *PLoS ONE* 6, e26922.

Nishi, A., H. Shirado, D. G. Rand, and N. Christakis (2015). Inequality and visibility of wealth in experimental social networks. *Nature, 526* (7573).

Payne, B. K., J. L. Brown-Iannuzzi, J. W. Hanna (2017). Economic inequality increases risk taking. *Proceedings of the National Academy of Sciences* 114, 4643–4648.

Piff, P., M. Kraus, S. Cote, B. Cheng, and D. Keltner (2010). Having Less, Giving More: The Influence of Social Class on Prosocial Behavior. *Journal of Personality and Social Psychology, 99*, 771–784.

Piff, P., D. Stancato, S. Cote, R. Mendoza-Denton, and D. Keltner (2012). Higher Social Class Predicts Increased Unethical Behavior. *Proceedings of the National Academy of Science*, *109*, 4086–4091.

Piketty, T. (2014). *Capital in the twenty-first century*. Cambridge: Harvard University Press.

Piketty, T., and E. Saez (2003). Income Inequality in the United States, 1913-1998. *Quarterly Journal of Economics*, 118(1), 2003, 1-39.

Piketty, T., and E. Saez (2014). Inequality in the long run. *Science*, *344* (6186), 838-843.

Piketty, T., E. Saez, and G. Zucman (2018). Distributional national accounts: Methods and estimates for the United States. *Quarterly Journal of Economics,*133(2), 553-609*.*

Putnam, R. (2000). *Bowling Alone*, New York: Simon and Schuster.

Roemer, J. (1998). *Equality of opportunity*. Cambrigde, Harvard University Press.

Smeets, P., R. Bauer, and U. Gneezy (2015). Giving behavior of millionaires. *Proceedings of the National Academy of Science,* 112, 10641–10644.

Starmans, C., M. Sheskin, and P. Bloom (2017). Why people prefer unequal societies. *Nature Human Behaviour*, *1*, 0082.

Stiglitz, J. (2012). *The Price of Inequality: How Today's Divided Society Endangers Our Future*. W. W. Norton and Company.

Tajfel, H. and J. Turner (1979). An Integrative Theory of Intergroup Conflict. In: Austin, W. G. and S. Worchel (Eds.), *The Social Psychology of Intergroup Relations*, 33–47, Monterey, Brooks/Cole.

Trautmann, S., G. van de Kuilen, and R. Zeckhauser (2013). Social Class and (Un)ethical Behavior: A Framework, With Evidence from a Large Population Sample. *Perspectives on Psychological Science,* 8, 487–497.

Verhaeghe, P. (2014). *What about me? The struggle for identity in a market-based society*. Scribe, London, UK.

Wilkinson, R., and Pickett, K. (2009). *The spirit level: Why more equal societies almost always do better*, London: Allen Lane.

Zheng, J. (2017). High Social Status Induces Pro-Social Behavior. *Working paper, Amsterdam*.