ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Fehr, Dietmar; Rau, Hannes; Trautmann, Stefan T.; Xu, Yilong

**Working Paper**

# Inequality, fairness and social capital

Discussion Paper Series, No. 650

**Provided in Cooperation with:**
Alfred Weber Institute, Department of Economics, University of Heidelberg

*Suggested Citation:* Fehr, Dietmar; Rau, Hannes; Trautmann, Stefan T.; Xu, Yilong (2018) : Inequality, fairness and social capital, Discussion Paper Series, No. 650, University of Heidelberg, Department of Economics, Heidelberg,
https://doi.org/10.11588/heidok.00025085

This Version is available at:
https://hdl.handle.net/10419/207626

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

University of Heidelberg

Department of Economics

Discussion Paper Series | No. 650

# Inequality, Fairness and Social Capital

Dietmar Fehr, Hannes Rau,
Stefan T. Trautmann, and Yilong Xu

July 2018

# Inequality, Fairness and Social Capital

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

[1]*Alfred-Weber-Institute,University of Heidelberg, Germany*

[2]*CentER, Tilburg University, Netherlands*

July 9, 2018

**Abstract**

We study the impact of unjust inequality on social trust and trustworthiness, and its separate effect on the economically successful and the unsuccessful, in a controlled economic experiment. We find evidence for a negative effect of unfair economic inequality on social interactions. Probing the boundaries of this effect, we document that this erosion of social capital critically depends on the context: if a well-off person is not directly responsible for the outcome of the worse-off person, then we observe no negative effects on trust and trustworthiness in the aggregate. Moreover, our data do not support the view that higher status or wealth leads to an erosion of pro-social attitudes: the successful are always more generous; groups of unsuccessful persons are least efficient and least generous in the trust game.

KEYWORDS: inequality, fairness, social capital

JEL CLASSIFICATIONS: C91, D31, D63

---

Corresponding author (Trautmann): Alfred-Weber-Institute for Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany, Phone: +49 6221 54 2952, Fax: +49 6221 54 3592, email: trautmann@uni-hd.de; and CentER, Tilburg University, P.O. BOX 90153, 5000LE Tilburg, The Netherlands. Dietmar Fehr thanks the German Science Foundation (DFG) for financial support through individual grant FE 1452/3-1.

## 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, 2016; 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 hypotheses can be derived from the literature in economics and the social sciences. The first hypothesis states that higher inequality, especially if perceived as unjust and caused by competition, hampers economic interaction (Alesina and Perotti, 1996; Bénabou, 1996; IPSP, 2017, Section 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; Fisman et al., 2015, Guinote et al., 2015; Nishi et al., 2015). That is, negative social consequences are caused by the behavior of the successful.

Both of these hypotheses are contested in the literature. However, empirical assessments of the effects of inequality and the role of the successful often suffer from an absence of counterfactuals and the endogeneity of status. Experimental methods offer an alternative approach for assessing the consequences of inequality as they make exogenous variation of inequality, the underlying causes of inequality, institutions and available information possible (e.g., Falk and Heckman, 2009; Charness and Fehr, 2015). While potentially having lower external validity, experiments thus provide a clear identification of causal effects and underlying processes.

This paper uses experimental methods to study the impact of unjust inequality on subsequent social interactions, differentiating between the behavior of the economically

successful and the unsuccessful. Our design thus aims to test both hypotheses within the same setting. We create income inequality in dyads, using a real-effort procedure with varying payment schemes. Subsequently, 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; Björnskov, 2017; Langer et al., 2017). 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.[1] 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.

We create exogenous variation in income inequality in the real-effort task by randomly assigning subjects to two different payment schemes. In our baseline condition subjects receive a piece-rate payment. This results in relatively low inequality and is typically not perceived as unjust. We compare the trust-game outcomes in this setting with an unjust high-inequality environment. To generate high inequality, we implement a relative-payment scheme that gives an undue advantage to one participant in the dyad.[2] This undermines equality of opportunity and the payment scheme can thus be seen as unfair from a normative perspective (e.g., Roemer, 1998). In a third condition, we employ the

[1] 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). According to this definition it is possible to measure social capital by eliciting values and beliefs with experimental tools such as the trust game (see e.g., Fehr (2009) for an extensive account of the measurement of trust and trust beliefs). 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).

[2] There is evidence documenting that (high) inequality is not per se seen as unfair (e.g., Bortolotti et al., 2017; Breza, Kaur, and Shamdasani, 2018; Fehr, 2018). For example, Fehr (2018) illustrates that an increase in inequality leads to more antisocial behavior but only if higher inequality cannot be clearly attributed to work effort and is possibly the result of immoral behavior.

same relative-payment scheme to generate unjust inequality as before, but randomly re-match participants in the trust-game stage (keeping earnings information constant across conditions). This eliminates the direct responsibility for each other's outcomes in the dyads and has the advantage of observing matches with equal and unequal outcomes.

Our results support the view that unjust inequality can negatively affect social interactions. That is, we document a significant decline in trust and trustworthiness when income inequality is the result of an income-generating process that is eminently perceived as unfair. However, we also find that this observed decline depends on a direct interaction in the first stage, i.e., when the well-off ("successful") player causing the poor outcome of the worse-off ("unsuccessful") player. If we take away the direct interaction by re-matching participants in the trust game, we find that especially the successful players maintain a high level of trust and trustworthiness, in particular when interacting among themselves. That is, the detrimental impact of inequality on social interactions critically depends on contextual factors.

We do not find evidence that the advantageous social position makes people more selfish: successful players are consistently more generous than the unsuccessful in absolute terms. However, holding the successful accountable to higher normative standards (such as sharing the trust-game pie equally), or evaluating generosity in terms of giving relative to someone's wealth position, we may well argue that they fall short on these standards.

In the next section, we introduce the experimental paradigm and design of our study followed by a description of how we induce unjust inequality. Section 3 shows that our experimental paradigm successfully induces inequality differences and a polarization of fairness perceptions. Clearly, neither inequality nor competitiveness have to be perceived negatively per se (e.g., Cappelen et al., 2007; Cappelen et al., 2013; Cappelen et al., 2014; Bartling et al., 2017, Bartling, Grieder, and Zehnder, 2017). Rather, it is the combination of inequality and unequal opportunity within a competitive environment that aims to trigger strong feelings of injustice in our experimental setup. While pooling these features hides their marginal contribution to the perception of unfairness, it guarantees a

powerful prime to reliably quantify the effects of unjust inequality on social interactions and, arguably, mirrors many settings outside the laboratory. [3] Competition, unequal opportunities, and inequality are inherent features of school education, universities, workplaces or labor markets more generally.[4] Section 4 discusses the effects of unjust inequality in fixed dyads and Section 5 discusses the effects when direct attributions of responsibility for others' outcomes cannot be made. We discuss these results in the context of the related literature in section 6.

## 2. Experimental Paradigm and Design

The current study employs an experimental paradigm in which dyads of participants interact in two stages. In the first stage, a repeated real-effort task involves either an individual piece-rate payment, or a competitive tournament with a favorable condition for the initial tournament winner (in a between-subjects design). While the piece-rate condition leads to modest inequality depending on individual performance, the tournaments amplify income differences in a way that is difficult to justify by the observed performance differences. In the second stage these same dyads then interact in a trust game. Consequently, we observe trust and trustworthiness depending on stage-1 conditions, and depending on stage-1 income. In a third treatment, the tournament-based real-effort stage is followed by a trust-game stage involving new matches of dyads, which have, however, exactly the same degree of information on each other's earnings as dyads in the fixed-pair tournament condition.

---

[3] Note that our paradigm can be extended to identify the marginal impact of the different features of the environment. However, the effects may not be additive making it impossible to disentangle them. See section 2.1 for a more thorough discussion of this issue.

[4] For example, Lemieux, MacLeod, and Parent (2009) document an economy-wide increase of performance-pay jobs in the U.S. labor market, along with a substantial increase in wage inequality. Features of competitive environments are innately linked to relative status concerns or relative-income comparisons, and it is long known that individuals care about their standing relative to others (e.g., Veblen, 1899). Several recent experimental studies suggest that such comparisons have, for example, detrimental effects on well-being (Card et al., 2012) or ethical behavior (e.g., Gill, Prowse and Vlassopoulos, 2013; John, Loewenstein and Rick, 2014).

In the following, we first describe the stage-1 income manipulation, and the elicitation of fairness judgments. We then provide details on the trust game stage with fixed dyads, and new dyads. Our three treatments are called *Piece Rate* (first-stage piece rate – fixed dyads); *Tournament* (first-stage tournament – fixed dyads); and *Tournament-New* (first-stage tournament – new partner in stage 2).

### 2.1. Stage 1: Inequality Manipulation

We implement a repeated real-effort slider task (Gill and Prowse, 2012) and vary the payment scheme to manipulate inequality, i.e., low inequality versus high and potentially unjust 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 Figure A.1 in the Appendix). 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.[5] 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 implemented payment scheme.

In the low-inequality condition (*Piece Rate*), 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 also salient in this setting.

In the high-inequality conditions (*Tournament* and *Tournament-New*), participants in a dyad also complete the slider task four times. In contrast to the *Piece Rate* condition,

[5] To avoid cheating, we used a keyboard locker to prevent students from using the arrow keys or the mouse wheel.

participants' payoffs in each round are determined through a relative performance scheme. 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 *Piece Rate*, participants receive information on the performance of each subject and the resulting payoffs after each round.

In addition to the high payoff, the subject with the higher performance receives a time bonus. More specifically, after an initial time budget of 120 second for both subjects, the winner of the first round obtains a time bonus of 8 seconds, and the winners of the second and third round get a time bonus of 6 and 4 seconds, respectively. The time bonus is subtracted from the time budget of the tournament loser in the respective round.

Tournament incentives are ubiquitous in economic life, and typically lead to a more spread pay distribution (and thus more inequality) than the underlying effort and ability justifies (Frank and Cook, 1995). We mimic this observation in our setup with a large difference in tournament prizes for winners and losers that hardly warrants the observed effort differences within dyads in a given round. This income difference magnifies over the rounds because of the substantial time gap (16 seconds) that arises after the first round and that makes it nearly impossible for the first-round loser to catch up in the subsequent rounds.[6] The condition thus induces inequality, caused by a competitive procedure that is difficult to justify on fairness grounds. In addition, this feature allows subjects to grow into their favorable or unfavorable economic positions over the course of the three remaining real-effort task rounds. This seems important in view of the conjecture that the successful are responsible for the erosion of the societal cooperation (e.g., Piff et al., 2010, 2012; Piff, 2013). For example, Piff (2013) observes that rich players in a rigged Monopoly Game experiment favoring their own economic status become increasingly imperious as inequality gets larger.

[6] Note that winning the first round depends on exerted effort and to a large degree on matching luck (i.e., the random assignment of the interaction partner).

Note that our *Tournament* design includes two components – competition and unjust procedure – that are absent in the *Piece Rate* condition and additionally results in higher income inequality than the *Piece Rate* condition. These three aspects arguably go often hand in hand in real-world settings, where initial advantages are amplified in competitive contexts, leading to enhanced inequality (e.g., Frank and Cook, 1995; Stiglitz, 2012). For example, if performance in or quality of primary school determines access to better secondary schools and subsequently to college, students end up with better jobs and higher earnings (see e.g., Chetty et al., 2011). At the same time, combining these three aspects provides a powerful instrument to probe the effects of (unjust) income inequality on social interactions. This is important as previous evidence suggests that inequality effects are subtle (see discussion in Section 6). As such, our focus is on maximizing the impact of inequality in the *Tournament* conditions in comparison to the inequality in the *Piece Rate* condition, and not on fully differentiating the marginal effects of the three ingredients (higher inequality, competition, unjust procedure).

### 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 game. At the beginning of the experiment, participants receive the 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, and field of study. After that 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.

### 2.3. Stage 2: Measurement of Social-Interaction Effects

In the second stage, we use a trust game to measure the effects of the exogenous income variation on social interactions. In this game there are two player roles, the first mover (trustor) and the second mover (trustee). The first mover has an endowment of €6.00 while the second mover has an endowment of €0.00. The first mover decides whether or not to transfer her endowment to the second mover. If she does not transfer, the game ends and the earnings will be €6.00 for the first mover and €0.00 for the second mover. In contrast, if she transfers her endowment, the experimenter triples the endowment such that the second mover receives €18.00 (and first mover has €0.00 now). The second mover then decides how much of the €18.00 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). We do not incentivize be-

liefs 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.

We implemented two variations of the trust game stage. In condition *Piece Rate* and *Tournament*, stage-1 dyads remain intact and proceed together to stage 2. We emphasized at the very beginning of the experiment that subjects will interact with the same partner throughout the whole experiment. At the start of stage 2, subjects are reminded of this fact. They also receive a reminder of their own and the other person's stage-1 earnings before making any choices in the trust game. In contrast, in condition *Tournament-New* the dyads are re-matched in stage 2, such that each person will play with a person with whom she did not interact in stage 1. Again, we made clear at the beginning of the experiment that they interact with different, randomly determined subjects in the two stages. At the beginning of stage 2, they 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 design precludes attributions of responsibility for each other's stage-1 outcomes. Moreover, as only earnings (but not effort) are communicated, is it not possible to attribute high or low stage-1 earnings to luck or effort.

### 2.4. 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*. While we conducted *Piece Rate* and *Tournament* in parallel, we added *Tournament-New* after completing the other conditions to scrutinize the generality of the results. 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 used the same subject pool and recruited 202 new subjects. In addition we ran sessions at the laboratory at the Technical University Berlin with a total of 140 subjects to increase power, given the larger number of subgroups in matching stage-1 winners and losers. Participants were undergraduate students from a wide range of differ-

ent majors, who were recruited with ORSEE (Greiner, 2015) in Berlin and Mannheim and with Hroot (Bock, Nicklisch, and Baetge, 2012) in Heidelberg.

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.

At the beginning of a session we matched participants in equal-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. The matching procedure was anonymous and in particular subjects were not aware of the exact matching procedure. We implemented this matching procedure 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).

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. In the presentation, we denote subjects as "successful" if stage 1 income equals €12.00 and as "unsuccessful" if stage-1 income equals €1.20. This definition reflects the typical payoff pattern for the successful and unsuccessful in condition *Tournament* (results are robust to alternative definitions). In our analysis using the successful-unsuccessful denomination, we drop observations with equal income (in *Piece Rate* and *Tournament*, N=12) and unclassified subjects with an income between €12.00 and €1.20 (in *Tournament-New*, N=54).

## 3. Results: Income Inequality Manipulation

We first provide evidence on effort levels, i.e., the number of correctly positioned sliders, in the different 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=0.795, two-sided t-test). Effort in *Tournament-New* was somewhat higher at 81 compared to *Tournament* (t=2.28, p=0.023). Importantly, 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 (3.93 in *Piece Rate*, 4.33 in *Tournament*, and 4.54 in *Tournament-New*, two-sided t-tests, all p>0.28).

**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. 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*.

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). 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 for all three conditions. In all three conditions, the unsuccessful perceive the task as less fair than the successful.

We conclude that the stage-1 manipulation succeeded in inducing strong differences in income inequality and fairness perceptions across piece rate and tournament 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.

**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, t-test; pairs with equal earnings excluded in analyses of successful and unsuccessful.

## 4. Results: Social Interaction Effects for Fixed Dyads

### 4.1. Main Effects

We now turn to the analysis of whether the strong differences in payoff inequality and fairness perception between *Piece Rate* and *Tournament* affect behavior in the stage-2 trust game. Figure 1 and Table 3 show 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 in either treatment.

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

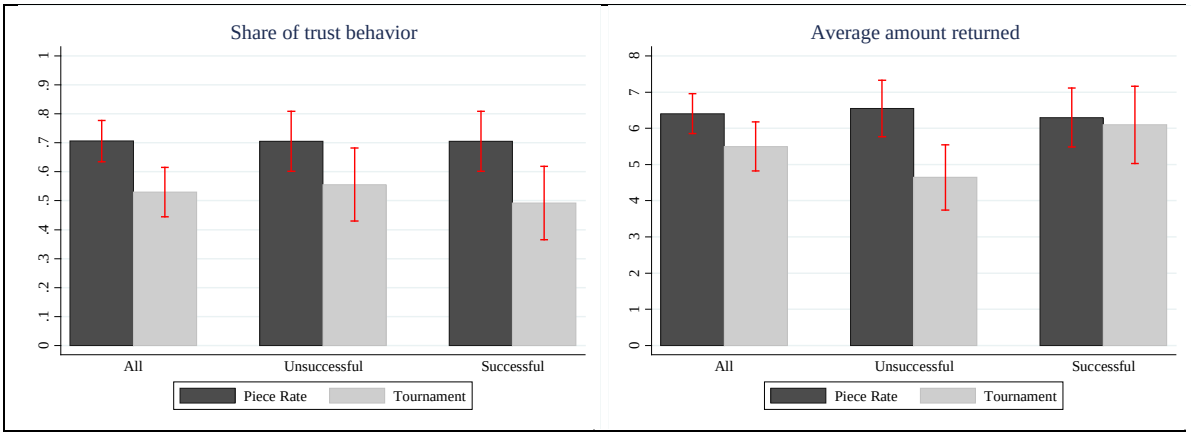


**Figure 1**: Trust and returns in *Piece Rate* and *Tournament*

**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=154) | €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, two-sided t-test for amounts returned, test of proportion for trust; pairs with equal earnings excluded in analyses of successful and unsuccessful.

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*.

Despite the higher amounts returned in *Tournament* by the successful, we may argue that they still fall strongly short of relevant normative benchmarks. First, they give less than the unsuccessful relative to their wealth. Second, in spite of having typically earned €12 in stage 1 (vs. €1.20 for their partner), they are far from sharing the stage-2 income (return €9), or overall income (return €15.60) equally. However, failure to meet such normative criteria is not restricted to the successful. In *Piece Rate*, stage-1 payoff differences are modest in most dyads, and both the successful and the unsuccessful fail to share their income equally (return €9). It seems that in general, stage-1 income is not taken into consideration when deciding about how much to return to the trustor. The observer's higher normative expectations towards the stage-1 winners make this behavior look less acceptable for the successful in *Tournament.*

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

***Result 3****:* In the low-inequality environment (*Piece Rate*) both the winners and the losers are equally generous; in the high-inequality environment (*Tournament*) the winners are more generous in absolute terms, and less generous relative to their wealth.

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. Expected trust game earnings are €1.08 lower in the *Tournament* condition (€7.26 vs. €6.18), a 15% loss compared to the *Piece Rate* condition.

### 4.1. Underlying Mechanism

The previous analysis has illustrated that there are substantial differences in trust and trustworthiness in the fixed-dyad design of the *Tournament* vs. *Piece Rate* condition. Our controlled laboratory context allows us to shed more light on the underlying mechanisms of this effect. We discuss the role of beliefs, the effect of pure inequality (not necessarily perceived as unjust), and the case of random losses in dyads with equal performance in *Tournament*.

*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 Appendix (Table A.1), we show that the *Tournament* condition induces more pessimistic beliefs regarding both trust and amounts returned. These effects are significant for the whole sample, but only significant for 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. The tables provide four specifications: Specifications 1 and 2 verify the raw comparisons discussed above including various controls. Specifications 3 and 4 include beliefs about trust and trustworthiness.

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. 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 if the other person were in the trustor's position). Results on trustworthiness support the reciprocal view as well. 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.

**Table 4**: Determinants of Trust

| *Dependent variable: Transfer (yes/no) to second mover* | | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Tournament* | -.178<br>(3.05)$^{***}$ | -.147<br>(1.74)$^{*}$ | -.133<br>(2.06)$^{**}$ | -.131<br>(1.45) |
| Successful | | .010<br>(.13) | | .070<br>(.81) |
| *Tournament* × Successful | | -.076<br>(.63) | | -.006<br>(.05) |
| Belief in trust by other | | | .428<br>(6.64)$^{***}$ | .411<br>(6.20)$^{***}$ |
| Belief in amount returned by other | | | .046<br>(3.27)$^{***}$ | .047<br>(3.33)$^{***}$ |
| Male | -.082<br>(1.42) | -.097<br>(1.61) | -.059<br>(.94) | -.071<br>(1.10) |
| N | 294 | 282 | 294 | 282 |
| Joint effect of tournament variable | | $\chi$=9.67, p<.01 | | $\chi$=4.05, p=.132 |

*Notes*: Marginal effects from probit regressions with robust z-statistics in parenthesis. All regressions control for session size and location. Linear regressions support the sign of the interaction terms in the probit regressions. Belief in amount returned by other scaled to 100 cents.

**Table 5**: Determinants of Amounts Returned

| *Dependent variable: Amount returned in cents* | | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Tournament* | -101<br>(1.9)$^{*}$ | -198<br>(2.70)$^{***}$ | -40<br>(.85) | -166<br>(2.68)$^{***}$ |
| Successful | | -8<br>(.13) | | 17<br>(.31) |
| *Tournament* × Successful | | 167<br>(1.53) | | 253<br>(2.80)$^{***}$ |
| Belief in trust by other | | | 13<br>(.26) | 22<br>(.44) |
| Belief in amount returned by other | | | 77<br>(7.42)$^{***}$ | 84<br>(8.62)$^{***}$ |
| Male | -215<br>(4.05)$^{***}$ | -207<br>(3.77)$^{***}$ | -170<br>(3.63)$^{***}$ | -157<br>(3.33)$^{***}$ |
| 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 t-statistics in parenthesis. All regressions control for session size and location. 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.

*Pure inequality*. While our design does not aim at disentangling the different aspects of unjust inequality and the subsequent erosion of trust and trustworthiness, we can use within-treatment variation in stage-1 earnings differences to obtain some insights into the effects of pure inequality, i.e., inequality that is not necessarily perceived as unjust. We define the earnings difference as the difference between a participant's own and the partner's stage-1 earnings. We use the same specification as in the regressions in Tables 4 and 5 and include the earnings difference, or alternatively its absolute value. We do this in the *Piece Rate* and *Tournament* conditions separately, and in the combined set of observations. Note that in the *Piece Rate* condition, we can study the effect of pure inequality absent the unjust and competitive allocation mode in *Tournament*. Although inequality is less severe than in *Tournament*, in *Piece Rate* there were still 78 dyads with a nonzero earnings difference, ranging from €0.05 to €4.10.

**Table 6**: Effect of Pure Inequality

| | *Piece Rate* | | *Tournament* | | All | |
|---|---|---|---|---|---|---|
| *Trust* | | | | | | |
| Earnings difference | .01 | (.31) | -.003 | (.78) | -.003 | (.75) |
| Earnings difference (absolute value) | .1 | (2.11)** | -.013 | (.82) | -.017 | (2.83)*** |
| *Amounts returned* | | | | | | |
| Earnings difference | -5 | (.19) | 8 | (1.79)* | 7 | (1.76)* |
| Earnings difference (absolute value) | 39 | (.91) | -14 | (1.19) | -12 | (2.09)** |

*Notes*: Each cell reports the coefficient from a separate regression. Marginal effects from probit regressions for Trust with robust z-stats in parenthesis. Tobit regressions for Amounts Returned with robust t-stats in parenthesis. Amounts are in cents and Earnings and Earnings differences are scaled to 100 cents. All regressions control for session size, location and gender. *,**,*** indicates significant difference from zero at 10%, 5% and 1% level.

Table 6 shows the coefficients for the earnings difference variables (each entry refers to one separate regression). We do not find evidence of any *negative effects* of inequality on stage-2 behavior within either the *Piece Rate* or the *Tournament* conditions. When combining the observations from the two treatments, the absolute value of the earnings difference becomes significant and negative for both trust and amounts returned, captur-

ing the treatment effects between *Piece Rate* and *Tournament*. In sum, there is no evidence that pure inequality is driving the observed negative social-interaction effects.

*Equality of outcome versus equality of opportunity*. In the *Tournament* condition, 8 dyads ended up with an equal performance in the first round of the slider task. In this case, a random draw determined the player who received the high payoff and the time bonus (vs. low payoff and time penalty). Comparing random winners and losers, we find that random winners tend to trust less but return more money, albeit the differences are insignificant possibly due to low number of observations. Controlling for a random loss or win in the regressions in tables 4 and 5 by including a dummy for bad and good luck, we find that all results are qualitatively unaffected. There are no significant effects for random winners and losers compared to other successful and unsuccessful. That is, the treatment effects are not merely driven by a potential perception of the random draw (equality of opportunity) being unfair compared to, for example, an equal split of the payment (equality of outcome).

## 5. Results: Social-Interaction Effects in New Dyads

The comparison between *Piece Rate* and *Tournament* has revealed strong detrimental effects on social interactions. In this section, we test the boundaries of this effect by rematching subjects into new dyads in stage 2. While the experience and perception of competition and unjust inequality is identical to the *Tournament* condition (see Section 3 results), a direct attribution of "responsibility" for the mutual stage-1 outcomes is absent in this condition. 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.[7] The rematching of dyads allows us to distinguish between the role of a play-

[7] König-Kersting et al. (2017) find that outcome information biases the perception of the underlying process ("outcome bias"). They find that the bias 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.

er's own income and the income of the matched partner: this was impossible in *Tournament* because these incomes were perfectly correlated.

We first run simple probit/tobit regressions with treatment dummies (and controls) to compare average behavior over all groups in *Tournament-New* (trust = 65%; amount returned = €6.61) to *Piece Rate* (trust = 71%; amount returned = €6.41) and *Tournament* (trust = 53%; amount returned = €5.50). The results show that *Tournament-New* does not differ significantly from *Piece Rate*, but leads to significantly larger trust and generosity than *Tournament* ( $\chi^2 = 4.82$, p=.028 and $\chi^2 = 9.52$, p=0.02).

**Table 7**: Social Interaction Effects – *Tournament-New*

| | Participants | vs. all (1) | vs. successful (2) | vs. unsuccessful (3) |
|---|---|---|---|---|
| Trusting | all | 65% (n=342) | 64% (n=144) | 64% (n=144) |
| | Successful | 68% (n=144) | 71% (n=56) | 69% (n=67) |
| | Unsuccessful | 62% (n=144) | 61% (n=67) | 55% (n=56) |
| Amount returned | all | €6.61 (n=342) | €6.49 (n=144) | €6.51 (n=144) |
| | Successful | €7.37 (n=144) | €7.96 (n=56) | €6.98# (n=67) |
| | Unsuccessful | €5.74 (n=144)*** | €5.48*** (n=67) | €5.52** (n=56) |

*Notes*: *,**,*** indicates significant difference between successful and unsuccessful; #,##,### indicates significant difference between successful partner and unsuccessful partner; at the 10%, 5%, 1% level, test of proportion for trust, and two-sided t-test for amounts returned. Unclassified participants (*n=54*, i.e., those with an income between €12.00 and €1.20) 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 7 shows detailed results for Trust and for Amounts Returned, separately for successful and unsuccessful decision makers, and successful and unsuccessful partners in the dyad. The upper panel of Table 7 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 stage-1 losers less and also for the losers to trust less. Accordingly, trust within dyads of unsuccessful participants is lower than trust within dyads of successful participants (55% vs. 71%, z=1.77,

p=0.0775). Regressions reveal that winners are 12.4 percentage points more likely to trust others than losers, which is a significant effect (see Table 8).

The lower panel of Table 7 shows that stage-1 winners are significantly more generous as second movers than stage-1 losers are. This holds for interactions with other winners and for interactions with losers. In fact, the successful in *Tournament-New* behave more generously on average than the successful under the Piece-rate condition (7.37 vs 6.3, p=0.02, two-sided t-test). When matched with another stage-1 winner, winners give even more to the partners in the dyad than when matched with a stage-1 loser (€7.96 vs. €6.97). As in the case of trust, these effects lead to an overall large difference of generosity within the group of unsuccessful people versus the group of successful people (€5.52 vs. €7.96, t=3.56, p<0.001).[8]

**Table 8**: Determinants of Trust and Amounts Returned – *Tournament-New*

| | Trust | Trust | Amounts Returned | Amounts Returned |
|---|---|---|---|---|
| Successful | .124<br>(2.00)** | .141<br>(1.83)* | 234<br>(4.18)*** | 160<br>(3.53)*** |
| Successful Partner | .050<br>(.79) | -.062<br>(.82) | 56<br>(1.00) | -35<br>(.83) |
| Belief in trust by other | | .581<br>(7.42)*** | | 122<br>(2.08)** |
| Belief in amount returned by other | | .062<br>(4.51)*** | | 76<br>(6.08)*** |
| Male | -.127<br>(2.00)** | -.130<br>(1.75)* | -88<br>(1.52) | -82<br>(1.79)* |
| N | 246 | 246 | 246 | 246 |

*Notes*: Marginal effects from probit regressions for Trust with robust z-statistics in parenthesis. Tobit regressions for Amounts Returned with robust t-stats in parenthesis. Amounts are coded in cents. All regressions control for session size and location. Belief in amount returned by other scaled to 100 cents.

[8] We can compare behavior in mixed dyads of successful matched with unsuccessful in Tournament-New to the respective group in Tournament. We find that the successful are more trusting in Tournament-New than in Tournament (p=0.02) and equally trustworthy. The behavior of the unsuccessful does not differ significantly between Tournament and Tournament-New.

The result that dyads of stage-1 losers subjects perform worst 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 reduced trust and trustworthiness within the group of dyads of stage-1 losers, stage-2 inequality is larger, and stage-2 welfare is lower in this group compared to the winner dyads. The expected welfare loss of the loser dyads amounts to €0.96, a 13% loss compared to the winner dyads. As in the case of trust, a regression analysis shows that the winners return significantly higher amounts in the trust game (Table 8).

***Result 4***: 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 the participants' beliefs explains the differences in trust game behavior between *Tournament* and *Tournament-New*. Table 8 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 winning partners, there are no such negative effects in *Tournament-New* (see Appendix A.2). Moreover, in *Tournament-New* the successful stage-1 players hold more positive views than the unsuccessful ones, especially when paired with another successful person.

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 loser 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*, where unsuccessful players were always matched with successful ones, there were no differences between the two groups. In contrast, recent literature suggests that, if relative position is salient, inequality may lead the poor to take higher levels of risk (see Payne et al., 2017, and references therein). 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

Our experiment investigates the potential negative effects of unjust economic inequality on social interactions and focuses, in particular, on the role of the economically successful in harming the social fabric. Our finding that unjust inequality arising in a competitive environment 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).[9] Increased pessimism about others' willingness to cooperate and thus a lower willingness to take the social risk of trusting a stranger is also indicative for a decline in social capital. Indeed, we not only find that beliefs are correlated with behavior but also that they are significantly more pessimistic if inequality is unjust. As a consequence, a vicious cycle of decreasing trust and cooperation may result, leading to a substantial loss of social capital.

Importantly, we find that the decline in trust and trustworthiness is mostly driven by the less well-off. Thus, we find no evidence for the hypothesis that the behavior of the successful is mainly responsible for the erosion of the social fabric. This is consistent with recent findings of Camera et al. (2017). They report that the worse-off subjects discriminate against better-offs by cooperating less with them in a repeated helping game, even when wealth is determined by chance, leading to an overall efficiency loss in the long run. Zheng (2017) similarly reports a higher degree of selfish behavior in a team production setting for low status subjects, where status is endowed in non-monetary terms (public praise). In Table A3 in the appendix, we summarize a larger set of experimental studies that relate to the question of the impact of inequality and competition on cooperation and trust. Although these studies greatly differ in terms of design, the overall picture is consistent with negative social capital effects being more likely. However, the

[9] Besides negative economic consequences, limited social interaction between the poor and rich may also increase the cultural gap between them. New evidence by Bertrand and Kamenica (2017) suggests that media consumption, consumer behavior, and time use of the rich and poor in the US have not diverged much since the 1960s despite the tremendous increase in income inequality, while social attitudes did diverge.

table shows a rather mixed picture about which social status group may drive the observed effects. That is, differences in implementation of inequality may be important for the relevant channel driving social capital effects.

If the, arguably modest, degree of competition and unjust inequality in a lab setting can induce strong effects on social behavior, we may expect the consequences to be even more severe in more significant situations outside the lab. However, our results also hint to the boundaries of such effects. Negative effects on trust and trustworthiness are overall reduced 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 tournament as equally unfair in the two *Tournament* conditions. At a first glance, this result contradicts results in Buser and Dreber (2016) who report negative effects of competition on cooperation even in newly assembled groups. In contrast to Buser and Dreber, however, subjects in our new-dyads condition were aware of their own and the other player's income situation. It seems likely that the apparent uncertainty about outcomes in Buser and Dreber induces a behavior closer to our condition of fixed dyads. Indeed, positive trust game effects emerge in the new-dyads condition especially in interactions between two stage-1 winners, i.e., in a situation with high income and income equality. If information about other's income is absent, positive effects on trust (and trustworthiness) may not be easily realized.

The observed differences between the fixed dyads and the newly assembled dyads hint at the volatility of the subtle psychological effects caused by inequality or fairness cues. Moreover, our manipulation combined strong inequality with a competitive and perceived unjust payment scheme. 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 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 from which unequal outcomes derive thus seems to constitute an essential aspect. Our results lend support to Starmans et al. (2017), who argue that it is not inequality per

se that bothers people in life, but economic unfairness. 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 inequality per se.

The finding of low social capital among the poor 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). Our results suggest that these field data may not simply caused by selection of people in or out of certain neighborhoods. Nevertheless, selection and upbringing may be important in the field. For example, in contrast with our and with Nettle et al.'s finding, Martinsson et al. (2015) report that Colombian university students from a wealthy university are less cooperative among each other than those at a lower social status university. The differences in upbringing and life experiences seem to have an opposite effect in this sample compared to Nettle et al.'s UK data.

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). In particular, this literature makes the causal claim that increasing wealth 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), 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; Smeets et al., 2015), and also relative to their wealth position (Korndörfer et al., 2015). A negative causal effect of increased wealth and status on prosociality may still exist, dampening an otherwise positive correlation between wealth and prosocial behavior through a selection effect if the prosocial are economically more successful.

In contrast to the results found in the above cited psychological literature, in our experiment the better-off stage-1 winners are always more generous than the worse-off in the second stage of the trust game. Arguably, stage-1 losers should thus be more trusting

than the winners, expecting higher returns from trust. Yet, this is not the case. Moreover, in the *Tournament-New* condition we observe that unsuccessful when matched with another unsuccessful subject are less trusting and less trustworthy than the successful when matched with another successful. That is, overall welfare is reduced and a higher degree of inequality emerges within their group of stage-1 losers. These results suggest that negative effects of unjust inequality are driven by the behavior of the poor, rather than the behavior of the rich.

Some qualifications need to be made with respect to the last point. Despite their higher degree of generosity in absolute terms, the successful players still fall substantially short of obvious normative benchmarks for second movers, such as equal sharing of the stage-2 payoffs, or even equal sharing of total experimental payoffs; they give a lower share of their income compared to the poor. That is, while the successful are more prosocial, they fall short of the potential normative expectations we may hold with respect to their behavior (in contrast to the empirical expectations as measured in the experiment, which may turn out more consistent with actual behavior). This is not the case for the poor, for whom no such expectations exist in the current setup. The same is probably true in larger contexts outside the lab. 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 by Piff et al. (2012) and others supporting the view of the selfish elite.

## Appendix

### A.1. Instructions and Screen Shots

An English translation of the original instructions can be found online at https://www.dropbox.com/s/21c7unjcko336ck/Merged%20Instructions%20%28English%29.pdf?dl=0 (to prevent the current document from becoming excessively large). The instructions also contain relevant screen shots with explanations. Here we present the screenshot of the real-effort task as referred to in the main text.

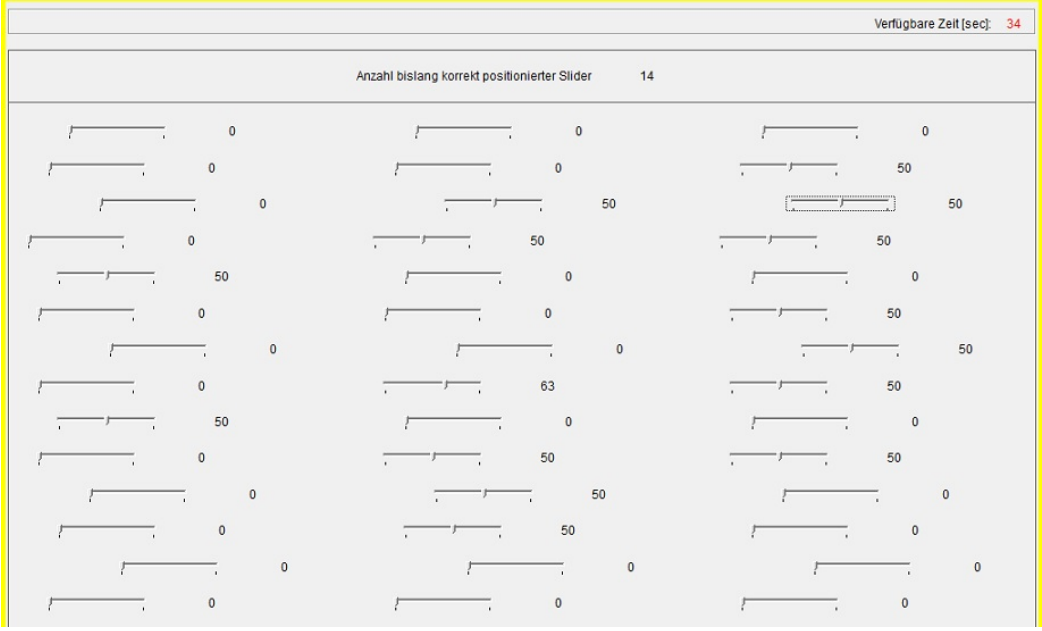


**Figure A1:** Screen Shot: Slider Task (42 sliders per round)

## A.2. Effects of Stage-1 Condition on Beliefs

Tables A1 and A2 show beliefs in treatments *Piece rate* and *Tournament*, and *Tournament-New*, respectively. Treatment comparisons find no significant differences between *Piece rate* and *Tournament-New* beliefs about trust (63% vs. 59%, p=0.386) and about amounts returned (€5.85 vs. €5.74, p=0.676).

**Table A1**: Effects of Stage-1 Condition on Trust Game Beliefs

| | Participants | *Piece rate* | *Tournament* |
|---|---|---|---|
| Belief in trust by other | all | 63% (n=160) | 50%$^{**}$ (n=134) |
| | successful | 58% (n=75) | 43%$^{*}$ (n=63) |
| | unsuccessful | 68% (n=75) | 56% (n=63) |
| Expected amount returned by other | all | €5.85 (n=160) | €5.08$^{**}$ (n=134) |
| | successful | €5.69 (n=78) | €4.36$^{***}$ (n=63) |
| | unsuccessful | €5.96 (n=78) | €5.60$^{\#\#}$ (n=63) |

*Notes*: *,**,*** indicates significant difference between treatment; #,##,### indicates significant difference between successful and unsuccessful; at the 10%, 5%,1% level, two-sided t-test for amounts returned, test of proportion for trust; pairs with equal earnings excluded in analyses of successful and unsuccessful.

**Table A2**: Effects of Stage-1 Condition on Trust Game Beliefs – *Tournament-New*

| | Participants | | vs. successful | vs. unsuccessful |
|---|---|---|---|---|
| Belief in Trust by Other | all | 59% (n=342) | 63% (n=144) | 54% (n=144) |
| | successful | 56% (n=144) | 66% (n=56) | 51%$^{\#}$ (n=67) |
| | unsuccessful | 61% (n=144) | 63% (n=67) | 59% (n=56) |
| Belief in Amount Returned by Other | all | €5.74 (n=342) | €5.90 (n=144) | €5.35 (n=144) |
| | successful | €6.04 (n=144) | €6.91 (n=56) | €5.40$^{\#\#\#}$ (n=67) |
| | unsuccessful | €5.33 (n=144)$^{**}$ | €5.44$^{***}$(n=67) | €4.93 (n=56) |

*Notes*: *,**,*** indicates significant difference between successful and unsuccessful; #,##,### indicates significant difference between successful and unsuccessful partner; at the 10%, 5%,1% level, two-sided t-test for amounts returned, test of proportion for trust; number of unclassified participants differs in cells conditioning on successful vs. unsuccessful or successful vs. unsuccessful partner.

### A.3. Experimental literature on inequality and competition

Table A3 presents laboratory experiments that study questions regarding the effect of competition and inequality on social interaction. We concisely summarize the key study aspects and the social interaction effect. If there exist any such effects, we indicate whether they are driven by the behavior of the successful/rich or the unsuccessful/poor.

**Table A3:** Overview of experimental studies

| | **Treatments** | **Stage 1** | **Stage 2** | **Social interaction effect** |
|---|---|---|---|---|
| Anderson et al. (2006) | Public / private show up fee | High / low show-up fees as inequality "*priming*" | Trust game | Private: Trust **(−)**, driven by the *successful*; Public: Trust **(=)**. |
| Sadrieh and Verbon (2006) | High skewed treatment | Randomly endowed "*earnings*" to create inequality | Public goods game | Contribution **(+)**, driven by the *unsuccessful*. |
| Brandts et al. (2009) | Rivalry/non-rivalry treatment | Prisoner's dilemma game (with a competitive setting in rivalry treatment) that creates inequality as "*priming*" | The circle test (similar to a dictator game) | "Generosity" towards others who they interacted before **(−)**, driven by *unsuccessful*. |
| Harbring (2010) | Competition game | Inequality as "priming" via a competitive game. | Trust game | Trust **(−)**, unclear who drives the results. |
| Heap et al. (2013) | High / low inequality | Randomly endowed "*earnings*" to create inequality. | Trust game (standard trust game or a labor market setting) | Trust and trustworthiness **(−)**, driven by *both* the successful and the unsuccessful. |
| Smith (2011) | Inequality | Randomly endowed "*earnings*" to create inequality | Trust game | Trust and trustworthiness **(=)** because the *successful* trust less but return more, while the *unsuccessful* do the opposite. |

**Table A3:** Overview of experimental studies (*continued*)

| | **Treatments** | **Stage 1** | **Stage 2** | **Social inter-action effect** |
|---|---|---|---|---|
| Greiner et al. (2012) | High / low inequality | Randomly endowed "*earnings*" to create inequality | Modified trust game (1st period only) | Trust **(−)**, driven by *both* |
| Nishi et al. (2015) | Visible/ non-visible wealth difference under three levels of inequality. | Randomly endowed "*earnings*" to create inequality | Cooperation game | Cooperation **(−)**, driven by the *successful* when inequality is visible. Inequality itself is not sufficient to drive this result, visibility is the key driver. |
| Buser and Dreber (2016) | Feedback on slider task | Competitive sliders task tournament, as "*priming*" | Public goods game | Contribution **(−)**, driven by *both* but more so by *the unsuccessful*. |
| Camera et al. (2017) | Info on wealth and role vs. no info | Helping game, payoffs as "*earnings*" to create inequality | Helping game (cont'd) | The act of helping others **(−)**, driven by *both*. |
| Brandts and Riedl (2017) | Direct, indirect, and no competition | Competitive double-auction market, payoffs as "*earnings*" to create inequality. In the absence of competition, randomly endow subjects. | Public goods game | Contribution **(+)** when no direct competition in stage 1, driven by the *successful*. Contribution (−) with direct competition, driven by *both*. |
| Falk (2017) | High vs. low social status. | Relative status info revealed, as "*priming*" of social status. | Electric shocks to others for personal gain | Incidence of shocking others is higher when high/low status group interact, driven by *both*. |

**Table A3:** Overview of experimental studies (*continued*)

| | **Treatments** | **Stage 1** | **Stage 2** | **Social inter-action effect** |
|---|---|---|---|---|
| Friedrichsen (2017) | Inequality | Randomly endowed "*earnings*" to create inequality. | Consumers with different initial wealth choose between socially responsible product and a cheaper alternative. | The unsuccessful choose the socially responsible products significantly more than the successful; no baseline available to compare overall effect due to inequality. |
| Lotito et al. (2017) | High vs. low inequality | Competitive real-effort task (admin tasks), as inequality "*priming*" | Public goods game | Contribution **(−)**: partial info on income / performance. Contribution (+) if full info. No competition effect, results driven by information about inequality. |
| Zheng (2018) | Baseline / random / true status | Painting evaluation, as status "*priming*" | Similar to a dictator game over losses | Less selfish behavior **(+)**, driven by the *successful*, those who earned their high status. |

Notes: In column Stage 1, "priming" indicates that stage-1 game payoffs either prime a winner/successful or loser/unsuccessful mindset and that they are not used as an endowment for the stage-2 game (i.e., payoffs in the two games are independent). "Earnings" indicate that the amount of money earned/randomly assigned to the subjects in the stage-1 game is used as the endowment of the stage-2 game. In column Social interaction effect, "(−)", "(=)", and "(+)" denotes a decrease, no effect, and increase of socially desirable interaction such as trust, cooperation, and contribution to public goods. Entries with n/a indicates not applicable because no relevant information is available.

## References


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

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

Anderson, L. R., J. M. Mellor, and J. Milyo (2006). Induced heterogeneity in trust experiments. *Experimental Economics*, *9*(3), 223–235.

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.

Bertrand, M. and E. Kamenica (2017). Coming Apart? Lives of the Rich and Poor Over Time in the United States, unpublished presentation, https://www.povertyactionlab.org/event/coming-apart-lives-rich-and-poor-over-time-united-states.

Björnskov, C. (2017). The Political Economy of Trust. In: R. Congleton, B. Grofman and S. Voigt (eds.), *Oxford Handbook of Public Choice*, forthcoming.

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*.

Brandts, J., A. Riedl, and F. van Winden (2009). Competitive rivalry, social disposition, and subjective well-being: An experiment. *Journal of Public Economics*, *93*(11), 1158–1167.

Brandts, J. and A. Riedl (2017). Market Interaction and Efficient Cooperation. *Working paper*, Barcelona GSE Working Paper Series, nr. 868.

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? *Working paper, Chapman University*.

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.

Cappelen, A., T. Eichele, K. Hugdahl, K. Specht, E. Sorensen, and B. Tungodden (2014). Equity Theory and Fair Inequality: A Neuroeconomic Study. *Proceedings of the National Academy of Science*, *111(43)*, 15368–15372.

Card, D., A. Mas, E. Moretti, and E. Saez. 2012. Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review*, 102(6), 2981–3003.

Charness, G., and E. Fehr (2015). From the Lab to the Real World, *Science* 350, 512-513.

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.

Falk, A. and J. Heckman (2009). Lab Experiments Are a Major Source of Knowledge in the Social Sciences, *Science* 326, 535-538.

Falk, A. (2017). Status Inequality, Moral Disengagement and Violence. *Working paper*, DIW discussion paper.

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

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.

Fisman, R., P. Jakiela, S. Kariv, and D. Markovits (2015): The distributional preferences of an elite, *Science*, 349 (6254).

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.

Friedrichsen, J. (2017). Is Socially Responsible Production a Normal Good? *Discussion Papers of DIW Berlin* 1644, DIW Berlin, German Institute for Economic Research.

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.

Gill, D., V. Prowse, and M. Vlassopoulos (2013). Cheating in the workplace: An experimental study of the impact of bonuses and productivity, *Journal of Economic Behavior & Organization*, 96, 120–134.

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

Greiner, B., A. Ockenfels, and P. Werner (2012). The dynamic interplay of inequality and trust—An experimental study. *Journal of Economic Behavior & Organization*, *81*(2), 355–365.

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

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.

Harbring, C. (2010). On the Effect of Incentive Schemes on Trust and Trustworthiness. *Journal of Institutional and Theoretical Economics*, 166, 690–713.

Heap, S. P. H., J. H. W. Tan, and D. J. Zizzo (2013). Trust, inequality and the market. *Theory and Decision*, *74*(3), 311–333.

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

John, L., G. Loewenstein, and S. Rick (2014). Cheating more for less: Upward social comparisons motivate the poorly compensated to cheat." *Organizational Behavior and Human Decision Processes*, 123(2), 101–109.

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.

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.

Lotito, G., M. Migheli, and G. Ortona (2017). Competition, Information and Cooperation. Working paper, Torino.

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.

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

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.

Piff, P. (2013, October). Does money make you mean? TedxMarin. TEDtalks. Video retrieved from http://www.ted.com/talks/paul_piff_does_money_make_you_mean

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 (2016). Distributional national accounts: Methods and estimates for the United States. *NBER Working paper No. 22956.*

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

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

Sadrieh, A., and H. A. A. Verbon (2006). Inequality, cooperation, and growth: An experimental study. *European Economic Review*, *50*(5), 1197–1222.

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

Smith, A. (2011). Income inequality in the trust game. *Economics Letters*, *111*(1), 54–56.

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.

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.

Veblen, T. **(**1899). *The Theory of the Leisure Class*. New York: Macmillan Company.

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.