Running head: INVISIBLE INEQUALITY

Invisible Inequality Leads to

Punishing the Poor and Rewarding the Rich

Oliver P. Hauser*
University of Exeter Business School

Gordon T. Kraft-Todd
Yale University

David G. Rand
Massachusetts Institute of Technology

Martin A. Nowak
Harvard University

Michael I. Norton
Harvard Business School

* To whom correspondence should be addressed. E-mail: o.hauser@exeter.ac.uk

Acknowledgements: We are grateful to constructive feedback from our colleagues at Harvard Business School, the Program for Evolutionary Program at Harvard University, and the Human Cooperation Lab at Yale University and MIT. We appreciate funding from the Harvard Foundations of Human Behavior Initiative and the Harvard Business School.

Invisible Inequality Leads to

Punishing the Poor and Rewarding the Rich

Abstract

Four experiments examine how lack of awareness of inequality affect behaviour towards the rich and poor. In experiment 1, participants who became aware that wealthy individuals donated a smaller percentage of their income switched from rewarding the wealthy to rewarding the poor. In experiments 2 and 3, participants who played a public goods game—and were assigned incomes reflective of the U.S. income distribution either at random or on merit—punished the poor (for small absolute contributions) and rewarded the rich (for large absolute contributions) when incomes were unknown; when incomes were revealed, participants punished the rich (for the low percentage of income contributed) and rewarded the poor (for their high percentage). In experiment 4, participants provided with public education contributions for five New York school districts levied additional taxes on mostly poorer school districts when incomes were unknown, but targeted wealthier districts when incomes were revealed. These results shed light on how income transparency shapes preferences for equity and redistribution. We discuss implications for policy-makers.

Keywords: inequality; transparency; cooperation; punishment; reward

The Norwegian government operates an online database which contains detailed information about all citizens' income, wealth, and tax contributions – which any Norwegian citizen can access (Norwegian Tax Administration 2015). While social scientists' first reaction to such radical transparency and opportunity for social comparison may be fear of the tearing of the social fabric, Norwegians seem to have survived the openness – and, notably, even have high tax morale (I. Lago-Peñas & S. Lago-Peñas 2010). While anecdotal, this example suggests that revealing incomes can be associated with increased support for contributing to the public good. Such preferences for spending on public goods such as social programs or health care are based, at least in part, on beliefs about income and wealth inequality (Oishi et al. 2011; Alesina & Angeletos 2005; Kuziemko et al. 2015; Brown-Iannuzzi et al. 2015). Survey evidence suggests, however, that many people may not be aware of the true extent of inequality in their country (Davidai & Gilovich 2015; Norton & Ariely 2011; Kiatpongsan & Norton 2014; Hauser & Norton 2017). A lack of transparency of incomes could lead individuals to hold different preferences than they would have if income information were available to them, and this misinformation might, in turn, have downstream consequences.

We examine how invisible, or hidden, income inequality affects group-level outcomes and individual people's behaviours towards the richest and the poorest group members – relative to when inequality is revealed. We hypothesised that revealing information about inequality might exert a substantial effect on behaviour: if people do not realise how little the poor have, and how much the rich have, they may be less sympathetic to low contributions from those who cannot afford to give more, and less likely to punish the rich for not contributing their "fair share."

Previous research on cooperation has explored the determinants and consequences of such sanctioning behaviour (Fehr & Gächter 2002; Rand et al. 2009; Hauser et al. 2016; Crockett et al. 2014; Krasnow et al. 2016; Jordan et al. 2016; Hauser, Nowak, et al. 2014). Results typically reveal that low contributors are punished, while high contributors are rewarded. In these studies, all players typically receive identical endowments in each round, and this equality is common knowledge to all players; thus the majority of these experiments, while highly informative regarding the maintenance of cooperation, shed little light on perceptions of, and reactions to, inequality. Indeed, only recently, scholars have been investigating inequality in the provisioning of public goods and social dynamics more generally (Nishi et al. 2015; Gächter et al. 2017; Anderson et al. 2008; Hauser, Traulsen, et al. 2014; for a full literature review, see Supplemental Online Material (SOM) Section 1.2.)

Building on this previous research, we introduce three novel features to explore the impact of people's recently demonstrated lack of knowledge (Kiatpongsan & Norton 2014; Hauser & Norton 2017) on public goods provisioning: (*i*) we experimentally vary whether the income distribution is hidden or revealed to explore the causal effect of knowledge of inequality on behaviour toward the rich and poor; (*ii*) we use income distributions that are extremely unequal (e.g., the actual United States distribution) to explore behaviour toward the rich and poor under conditions reflective of real-world inequality; and (*iii*) we allow participants to either punish, reward, or both punish and reward the poor and the rich to explore how these sanctions are utilised to address perceived inequity.

In Experiment 1, we randomly assign participants to one of two conditions, in which they are either aware of donors' incomes (the *revealed* condition) or unaware (the *hidden* condition). Based on representative of real-world income and donation distributions, we study participants' reward behaviour towards a real donor. In Experiment 2, we randomly assign participants to different incomes—reflective of the income distribution of the United States—and then further randomly assign them to either the *revealed* or *hidden* condition; we examine sanctioning behaviour—both punishment and reward—towards other players in a repeated public goods game. In Experiment 3, we examine whether merit might play a moderating role: whereas in Experiment 2 incomes are assigned randomly, in the third experiment incomes are assigned based on task performance. We again assign participants to either the *revealed* or *hidden* condition, and examine their decisions to punish and reward other players. Finally, in Experiment 4, we explore potential policy consequences, examining preferences for taxation policies: we use actual data on charitable contributions to public schools in New York to examine which school districts—wealthy or poor—people believe should be taxed more highly, as a function of whether the incomes of those districts are hidden or revealed.

Across all four experiment, our results can be summarized as follows: when incomes were hidden, participants rewarded the richest group members for their seemingly high contributions, while punishing the poorest for contributing apparently little. When incomes were revealed, however, participants reversed this behavioural pattern, such that they rewarded the poorest (for contributing a high relative amount of their small endowment) and punished the richest (for contributing a low relative amount of their large endowment).

Experiment 1

Methods

Participants ($N = 315$) were recruited on Amazon Mechanical Turk from the United States. Participants were told that they had to decide to which out of five donors—all of them previous participants in another study on Amazon Mechanical Turk—they would assign a $1.00 bonus payment. The donors were chosen so that their donation behaviour represented the actual distribution of U.S. donors across five income ranges (Maryland CPA 2016): on average, households donated $1,874 (income under $25,000), $2,594 ($25,000–$50,000), $2,970 ($50,000–75,000), $3,356 ($75,000–$100,000), and $4,130 ($100,000–$200,000) in 2014, the year with the latest available data. Participants were asked which one of the five donors should receive a $1.00 bonus payment. The decision was incentive-compatible: one participant was drawn at random and their decision implemented to pay the donor.

Participants were randomly assigned to one of two conditions. In the *revealed* condition, participants were told the donors' average donation as well as their income range. Conversely, in the *hidden* condition, they saw the donation amount only but not the income range. We expected participants in the *hidden* condition to reward the (unbeknownst to them richest) donor who donated the largest absolute amount of money in the past year, while we predicted a reversal of reward behaviour in the *revealed* condition, such that participants would reward the poorest donor who gave the highest percentage of their income to charity. See Supplemental Online Material (SOM) Section 2.2.1 for detailed methods.

Results

As predicted, we find that the distributions of donors rewarded is significantly different between the two conditions (Figure 1; using linear regression: coeff = -2.250, $p < 0.001$, Table S1; qualitatively similar results are obtained with rank-sum test: $Z = -10.935$, $p < 0.001$). This shift in reward behaviour occurs only for the top and bottom income classes: when income ranges are revealed, participants are more likely to reward the poorest donor (coeff = 2.681, $p < 0.001$, Table S2) and less likely to reward the richest donor (coeff = -3.72, $p < 0.001$, Table S2). There is no difference in likelihood to reward donors in the middle of the distribution (all ps > 0.1).

Experiment 2

Methods

Participants ($N = 855$) were recruited on Amazon Mechanical Turk, read the instructions and answered comprehension questions, and were assigned to groups of five to play a two-stage economic game over 10 rounds.

We used a standard paradigm in experimental economics; an incentive-compatible, repeated public goods game (PGG) in groups of 5 players. In each of 10 rounds, every player was assigned an "income" and chose how much of that income to contribute to a common pool; all contributions were doubled and divided equally among the five players (see SOM Section 2.2.1 for more details about experiment design). We then showed each player the contributions of all other players, and gave each player the opportunity for costly sanctioning of all other players. In the *punishment* condition, participants could pay 1 unit to decrease any other participant's payoff by 3 units; in the *reward* condition, participants could pay 1 unit to increase any other participant's payoff by 3 units. Participants could spend up to 2 units on each of the other participants.

Before the start of the game, participants were randomly assigned to receive an income, and were told they would receive this same income in each round of the game. We used the United States pre-tax incomes by quintile to create incomes for the five players (Congressional Budget Office 2007): The top quintile participant received 55 units out of 100 units in the group (or 55% of all income), the next 19 units, the next 13 units, the next 9 units, and the bottom quintile participant 4 units (Fig. 2A). Once assigned to an income level, participants received the same income each round for 10 rounds.

Across all conditions, participants played two stages in each round. In Stage 1, participants could contribute any amount of their income to a common project. Any units contributed were doubled and split equally among all five group members. In Stage 2, participants could see everyone's contributions and could either punish or reward their group members (depending on the condition). Participants could not spend more in Stage 2 than they had earned in Stage 1. At the end of each round, participants saw their group members' decision to reward or punish them in Stage 2, and a summary of their payoff in this round. To ensure that participants could not identify one another across multiple rounds and to avoid retaliation (Nikiforakis et al. 2012), each player was only known by a series of random letters that changed at the beginning of every round.

The experimental design was a 2 (*punishment* versus *reward*) X 2 (*hidden* versus *revealed*) between-participants design ($N = 600$; for details about the experimental design, see SOM Section 2.2.2.). In the *hidden* condition, players had no information about the incomes of the others in their group (Fig. 2B): they made contributions, viewed others' contributions, and decided to punish or reward based only on the total amounts contributed by other players. In the *revealed* condition, in contrast, participants were shown the income of each player as they made their decisions to punish or reward – allowing them to base their decisions not only on the total amount contributed, but also the *percentage* of available income that each player chose to contribute (Fig. 2C). For example, a player who contributed just three units in the *hidden* condition may look stingy; learning that this player had only four total units in the *revealed* condition may dramatically alter perceptions of their contribution.

We expected that in the *hidden* condition, participants would generally view the (low total) contributions of bottom quintile players unfavourably, inducing punishment, and the (high total) contribution of the top quintile players favourably, inducing reward. In contrast, we expected that in the *revealed* condition, participants would generally view the (high percentage) contributions of bottom quintile players favourably, inducing reward, and the (low percentage) contribution of the top quintile players unfavourably, inducing punishment.

Results

We find that, indeed, participants in the *hidden* condition rewarded richer participants more (coeff = 0.636, $p < 0.001$), whereas those in the *revealed* condition rewarded poorer participants more (coeff = -0.720, $p < 0.001$; interaction between income and *revealed* dummy, coeff = -1.356, $p < 0.001$; Figure 3 and Table S3). We observe a mirror image of these results for decisions to punish: participants in the *hidden* condition punished poorer participants more (coeff = -0.282, $p = 0.042$), whereas those in the *revealed* condition punished richer subjects more (coeff = 0.692, $p < 0.001$; interaction between income and *revealed* dummy, coeff = 0.974, $p < 0.001$; Figure 3 and Table S5). Thus, knowledge about economic inequality had a profound effect on sanctioning.

Why did players sanction so differently in the *hidden* and *revealed* conditions? Across both conditions, richer players contributed larger total amounts (*hidden*: coeff = 3.172, $p < 0.001$; *revealed*: coeff = 4.734, $p < 0.001$; Table S7), but lower percentages of their income (*hidden*: coeff = -0.098, $p < 0.001$; *revealed*: coeff = -0.058, $p < 0.001$; Table S8) (Figure 4). Collapsing across conditions, top quintile participants contributed 20.49 out of 55 units (or 37% of their income) whereas bottom quintile participants contributed 2.83 out of 4 units (or 71% of their income). The pattern of sanctioning we observe therefore follows naturally if sanctions were assigned based on percentage of income contributed in the *revealed* condition but total amount contributed in the *hidden* condition.

Supporting this logic, in the *revealed* condition, participants conditioned their sanctioning decisions on the percentage of the target's income that was contributed (using percentage contributed as the independent variable; predicting punishment: coeff = -4.664, $p < 0.001$, Figure 5A; predicting reward: coeff = 6.320, $p < 0.001$, Figure 5B; Table S12), more so than on the absolute amount contributed. In the *hidden* condition, conversely, where only total contribution amounts were known, sanctioning was based on total amount contributed (predicting punishment: coeff = -1.863, $p = 0.019$, Figure 5C; predicting reward: coeff = 4.700, $p < 0.001$, Figure 5D; Table S11), but not on percentage of income contributed (predicting punishment: coeff = 0.030, $p = 0.954$; predicting reward: coeff = -0.216, $p = 0.677$; Table S11).

We next consider the consequences of income transparency on total contributions and final payoff inequality. Overall, significantly more units were contributed in the *revealed* condition compared to the *hidden* condition (coeff = 1.745, $p = 0.002$; Table S17). However, these overall greater contributions in the *revealed* condition were not distributed equally: the richest participant earned significantly less per round (predicting round payoff of top quintile only: coeff = -5.430, $p < 0.001$), the fourth quintile was unchanged (coeff = -0.474, $p = 0.715$), but all other participants earned significantly more (first quintile: coeff = 6.261, $p < 0.001$; second quintile: coeff = 3.950, $p = 0.003$; third quintile: 3.399, $p = 0.010$).

In addition, revealing income not only affected contributions and payoffs; it also resulted in less inequality by the end of the game. The Gini coefficient—a common summary measure of inequality—after the final round of the game was significantly higher in the *hidden* condition (average 0.238) compared to *revealed* (average 0.169; rank-sum, $p < 0.001$). Notably, participants in the bottom (poorest) through fourth (second richest) quintiles maintained (or even increased) their contribution levels over the ten rounds in both the *hidden* and *revealed* conditions; in contrast, although participants in the top (richest) quintile in the *revealed* condition also continued to contribute steadily over time (coeff = -0.382, $p = 1.000$ bonferroni-corrected), top quintile players in the *hidden* condition *decreased* their contributions over the ten rounds (coeff = -1.077, $p < 0.001$ corrected) (Tables S19 and S20). In other words, in the *hidden* condition, sanctions were less effective at maintaining contributions among those with the greatest ability to contribute to the public good.

Experiment 3

Participants in our second experiment were assigned their income randomly. However, incomes in the real world are not just the product of chance, but also of effort. Earned incomes could justify inequalities and thus reduce a desire for redistribution through punishment or reward (Cappelen et al. 2013). In Experiment 3, we thus assigned income based on participants' performance in an individual effort task before playing the public goods game. While the general setup of the game was similar to Experiment 2, we also made several additional changes to the design, which are described in detail in SOM Section 2.2.3; below is a shortened summary.

Methods

Participants (*N* = 440) who were recruited on Amazon Mechanical Turk played a two-phase experiment. In Phase 1, participants completed an individual effort task that affected their income level in the second phase. In Phase 2, participants played the same repeated two-stage economic game with sanctions (unlike before, both reward and punishment options were available simultaneously). As before, participants were assigned to one of two conditions: participants in the *revealed* conditions saw their own income, the income of the other participants and the sum of all incomes, while participants in the *hidden* condition only saw their own income and the sum of all incomes in the group.

Unlike in the previous experiment, participants were not randomly assigned their income at the start of the economic game, but earned their position in the income distribution beforehand through an effort task (Abeler et al. 2011) and this was common knowledge to all players. The best-performing participant in a group earned the highest income, the second-best performing participant earned the second-highest income, and so on.

Furthermore, in this experiment, we increased external validity by ensuring that participants in *both* conditions were aware of some degree of inequality, as they are in the real world (Kiatpongsan & Norton 2014): we informed participants (who had been recruited exclusively from the U.S.) in both the *hidden* and *revealed* conditions in the second experiment that the income distribution used in the game was derived from the United States income distribution by quintile (U.S. Census Bureau 2013). However, based on previous research (Norton & Ariely 2011; Hauser & Norton 2017), we anticipated that participants would misjudge the actual extent of U.S. inequality (and thus inequality in the game) and would therefore not adjust their sanctioning behaviour sufficiently, thereby showing a similar sanctioning pattern to the previous experiment.

Results

In our third experiment, we found qualitatively similar results as before: participants continued to reward the rich and punish the poor in the *hidden* condition (predicting number of units received, where positive values imply receiving on average reward and negative values punishment: coeff = 0.053, $p = 0.042$), while this trend reversed completely in the *revealed* condition (coeff = -0.171, $p < 0.001$; interaction between income and *revealed* dummy: coeff = -0.225, $p < 0.001$; Table S22). Across conditions, richer participants contributed more in absolute terms (coeff = 3.979, $p < 0.001$; Table S22), but less as a percentage of their income (coeff = -0.078, $p < 0.001$; Table S25), than poorer participants. As before, this sanctioning pattern was linked to reward and punishment decisions: higher absolute contributions received more reward in the *hidden* condition (coeff = 0.571, p < 0.001; Table S28), but higher *percentage* of income contributed was more rewarded in the *revealed* condition (coeff = 1.775, p < 0.001; Table S28).

Groups in the *revealed* condition (average Gini index = 0.124) again ended up with more equal payoffs than those in the *hidden* condition (Gini 0.255). Furthermore, overall contributions to the public good were higher in the *revealed* condition than in the *hidden* condition (coeff = 3.134, $p < 0.001$; Table S30) but they did not benefit everyone equally: the richest participant earned significantly less per round (coeff = -4.923, $p = 0.011$), the fourth-quintile participant's payoff did not change (coeff = 2.864, $p = 0.105$), but all other participants earned significantly more (first quintile: coeff = 11.990, $p < 0.001$; second quintile: coeff = 8.269, $p < 0.001$; third quintile: coeff = 6.218, $p = 0.001$).

In our third experiment, even when incomes were earned *and* when participants were informed that the income distribution was reflective of their own country's distribution, participants continued to punish the poor and reward the rich when the income distribution was hidden, but reward the poor and punish the rich when incomes were revealed.

Experiment 4

In our final experiment, we turn to the question of how revealing inequality might influence relevant policy outcomes. In particular, we examine taxation preferences with regards to school funding. We used actual data from donations to Parent-Teacher Associations (PTA) in five New York school districts, which vary on both average income and average donation amounts. In particular, districts with higher incomes tend to contribute more to PTAs, resulting in inequality in educational funding. We showed participants either the contributions to the PTA from each district (*hidden* condition) or both the contributions and the average income (*revealed* condition), and asked which district they believed should be responsible for an additional tax which helps all schools across districts. We expected that in the *hidden* condition, poor school districts would be "punished" with the additional tax, whereas in the *revealed* condition, punishment would switch toward wealthier districts.

Methods

Participants (N = 313) were recruited on Amazon Mechanical Turk and randomly assigned to two conditions (*hidden* versus *revealed*). Participants in all conditions read about the annual fundraisers that the PTA organises across American schools, which helps schools afford non-state funded initiates such as a science lab, teachers' aides, and additional equipment; as in the public goods game in Experiments 2 and 3, these funds were pooled across the districts and then distributed back to each district. Participants were then informed that the city government wants to raise $2,000 additional funding for each school by raising taxes, and were asked: "Which parents do you think should pay the additional tax to cover the $2,000 per school in all five schools?" The choice of the school parents that bear the additional tax is our measure of "punishment" in this study.

In the *revealed* condition, participants saw a list of five schools (identified by a string of two random letters), the average PTA donation from parents at this school, as well as the average household income in that area. Conversely, in the *hidden* condition, participants saw the same five schools but only the average PTA donation with no income information. In both conditions, the five schools are modelled after a real dataset. We obtained data from average PTA donations (Sullivan & Felton 2014) and median household incomes (Weissman Center for International Business 2016) from five New York City school districts (Bronx, Brooklyn, Manhattan, Queens, Staten Island). Participants saw the following five schools which appeared in random order: $353 average PTA donation (median income: $35,176), $1,227 ($51,141), $4,249 ($60,422), $9,759 ($71,622), and $1,486 ($75,575). See SI Section 2.2.4 for more details on the experimental design.

Note that the absolute donation averages do not perfectly track with household incomes: parents in Manhattan—the district with the highest median household income—give less to PTA fundraisers ($1,486) than parents in Queens ($4,249) and Staten Island ($9,759). Our predictions remain qualitatively unchanged, however: we predict that participants' punishment preferences will shift away from raising taxes on the poorest school parents in the *hidden* condition to raising them on the richest parents when incomes are *revealed*.

Results

As predicted, we found that participants' preference shifted from punishing poor school parents in the *hidden* condition to punishing richer school parents in the *revealed* condition (using linear regression predicting 1 = poorer to 5 = richer school areas by *revealed* dummy: coeff = 1.826, $p < 0.001$, Table S34; qualitatively similar results are obtained with rank-sum test: $Z = -11.883$, $p < 0.001$). Most changes in preferences occurred in a shift to the richest school parents. Specifically, participants were significantly less likely to raise taxes from poorer parents (M = 41.5%) in the *revealed* condition relative to the *hidden* condition (M = 7.1%; using logit regression predicting likelihood of choosing the poorest school parents to pay taxes by *revealed* dummy: coeff = -2.222, $p < 0.001$, Table S35). Conversely, participants were more likely to raise taxes on the richest parents when incomes were revealed (M = 77.3%) compared to when incomes were hidden (M = 4.4%; predicting choosing the richest school parents: coeff = 4.302, $p < 0.001$, Table S35).

Participants were also less likely to want taxes raised from the second-richest school districts in the *revealed* condition than in the *hidden* condition (predicting choosing second-richest school: coeff = -1.985, $p < 0.001$, Table S35); there were no significant differences in punishment behaviour in remaining two districts (both ps > 0.1).

In sum, Experiment 4 offers an example of a real-world public goods dilemma with similar dynamics to our laboratory paradigms: people have an opportunity to contribute to a valued public good (in this case, education for their children) via voluntary contributions (in the form of charitable gifts), which are then pooled across groups of people with different incomes (in this case, schools districts with higher and lower annual incomes). The decision for participants is to choose which school district should be the target of "punishment" – in the form of an additional tax levied by the government. Consistent with the results from the previous studies, awareness of inequality shifted people's preferences away from punishing the poor (in the form of additional taxation) and toward rewarding them.

General Discussion

In sum, revealing inequality had substantial effects on people's decisions to reward or punish others, and on total contributions to the public good. Participants were more likely to punish poorer participants and reward richer participants when inequality was hidden; when income was revealed, participants became more sensitive to people's *ability* to contribute – leading them to punish the rich and reward the poor. These general patterns held true across charitable donations (Experiment 1), contributions to public goods in interactive group-based studies (Experiments 2 and 3), and with regards to taxation to support public education (Experiment 4).

To some, these results may not come as a surprise – indeed, one might argue that revealing the income distribution to participants would necessarily change their behaviour. However, we believe that this is only obvious in hindsight. First, people in the real world *are* aware of inequality in their communities and lives, but at the same time they underestimate the extent to which incomes (and wealth) are so drastically different between people (Norton & Ariely 2011; Hauser & Norton 2017) – and consequently they do not take these different into account sufficiently when they evaluate contributions to a public good. Thus, we contribute to the literature by demonstrating that implicit knowledge of inequality in a country (such as the *hidden* conditions in Experiments 3 and 4) is not sufficient to make people realise that they could account for large differences in contribution amounts. People do not seem to spontaneously consider the background wealth of others when evaluating public contributions. Conversely, once incomes were revealed, the extent to which participants reserve their sanctioning patterns is also quite remarkable: we find that people are very responsive to this type of information, such that they punish the rich and reward the poor consistently in all our experiments.

Participants' motivation to punish the rich and reward the poor in the *revealed* condition could arise from seeking equity in contributions or simply aiming to reduce inequality in the group, or both (Rawls 1971; Frohlich et al. 1987; Van Dijk & Wilke 1994). Though equity concerns were present in our sample, we additionally show in the SOM that participants in the *revealed* condition punished the top quintile more than any other player, even when the rich had contributed the same percentage of their income. In other words, when inequality was revealed, participants in our experiments desired not just equity but that the wealthy were slightly less well off, suggesting that spite may play a role in sanctioning under transparency. The observation that our participants were unaccepting of inequality adds to a growing literature on social preferences, egalitarianism and libertarianism (Konow 2000; Cappelen et al. 2007). Conversely, what levels of inequality are acceptable remains an open question but some recent work has started to shed light on this question (Norton 2014; Kiatpongsan & Norton 2014). Certainly, more research is need to explore what shapes belief formation of perceived and ideal inequality (Sheehy-Skeffington et al. 2016); one fruitful area to investigate is the role of normative second-order beliefs (Jachimowicz, Hauser, et al. 2018; Kraft-Todd et al. 2018)—one's beliefs of what others believe—which might, in turn, shape one's own attitudes towards an issue, including inequality acceptance.

Our participants' unwillingness to tolerate inequality of income persisted even when incomes were assigned by performance. Literature on pay dispersion in organisations has found that being able to attribute unequal rewards to differences in production and performance can help reduce feelings of unfairness (Bloom & Michel 2002; Shaw 2014; Breza et al. 2017). Furthermore, plausible justifications typically make people more accepting of inequalities (Cappelen et al. 2013). However, in Experiment 3, we did not find evidence that earning incomes moderated the effects in either the *hidden* or *revealed* conditions. One explanation might be that most previous research has focused on moderate, not extremely high, levels of inequality, which are likely to map more closely onto people's (inaccurate) beliefs about the distribution of wealth and income; varying the extent to which distributions reflect reality versus perception offers an important area for future research. Another explanation might be that the act of earning incomes was less salient in our experiments than the incomes themselves. Future research is needed to investigate other potential moderators for the observed sanctioning behaviour, including individual traits such as subjective status (Kraus & Mendes 2014; Akinola & Mendes 2013; Brown-Iannuzzi et al. 2015), sense of control (Kraus et al. 2009), or risk preferences (Payne et al. 2017). Furthermore, group size and the extent to which (lack of) sense of control is affected by inequality (Chou et al. 2016; Jachimowicz, To, et al. 2018) could help explain why and how participants are using reward and punishment in our experiments.

Visibility of income will of course not always have positive effects. For example, without the opportunity to sanction others, revealed inequality can lead to more segregation of the rich and poor, and even further inequality (Nishi et al. 2015). Furthermore, making incomes between co-workers public can reduce satisfaction and productivity of low earners (Card et al. 2012; Pfeffer & Langton 1993). Most studies have usually focused on individuals who conduct similar work and might thus feel unfairly treated if their salaries were vastly different. Much less research has looked at the consequences of CEO salary visibility, and it remains an open question to how mandatory reporting policies (such as the U.S. Dodd Frank Act) that will require CEO salaries to be disclosed publicly will affect productivity, retention and satisfaction of workers.

While our income distributions were drawn from the real world, our paradigms necessarily offer stylised examinations of the impact of inequality on the public good and some limitations must be acknowledged. First, our experiments were conducted on Amazon Mechanical Turk (AMT), an online labour market, often used for research purposes (Buhrmester et al. 2011; Rand 2012); however, this labor market is not nationally representative and thus results should be interpreted with caution. Of particular relevance to our own investigation is the demographic skew towards younger, more educated and more technologically savvy people in a typical AMT sample compared to the general U.S. population. For example, in exploratory analysis in our fourth experiment, we found that, while the overall shift to levying a tax on the richest household (versus the poor) was statistically significant and economically meaningful across all age groups in the *revealed* condition, there was more variation across age groups in the *hidden* condition, such that older individuals in our sample were more likely than younger participants to assign the tax to the poorest households. This suggests that older participants were (likely unbeknownst to them) more willing to punish the poor and not consider their unobservable circumstances when incomes were hidden. Whether this suggestive evidence is more generalisable remains to be understood in follow up work; however, these results do point to the importance of conducting research with a wide range of participants from different and representative backgrounds. Future research should thus consider demographic variation in lab and online experiments or ideally, wherever possible, conduct research on inequality using field experiments. The latter is particularly policy relevant: for example, policy-makers and scholars interested in behavioural science and choice architecture (or "nudges", Thaler & Sunstein (2008)) might want to run an audit or procedural field experiment (Harrison & List 2004; List 2006; Ludwig et al. 2011; Hauser et al. 2017; Kraft-Todd et al. 2015) to study and potentially shift incorrect beliefs about inequality ("budging" versus "nudging", see Hauser et al. (2019)) and its effects on redistribution; recent examples of large-scale, policy-relevant field experiments in the context of inequality include Jachimowicz et al. (2017) and Sands (2017).

Second, we had to make some design choice in our experiments that may differ from the real world. For example, in Experiments 2 and 3, we restricted the amount that all participants could pay to punish or reward other players to the same fixed number of units per player (as in Rockenbach & Milinski (2006); Rand et al. (2009)). We observe a weak, marginally significant relationship between income and sanctioning behaviour across both the *revealed* and *hidden* conditions in Experiment 2 (linear regression predicting units spent on reward or punishment by income quintile: coeff = 0.070, p = 0.053), though not in Experiment 3 ($p > 0.1$), suggesting that placing restrictions on sanctioning may have some effect. Future research is needed to delineate the effects of varying limitations on sanctioning. Of course, the real world may not always provide an upper bound: given their greater resources, the rich have much greater ability to inflict harm or bestow benefits on others. Still, there are some real-world situations in which all decisions count equally: for instance, casting a vote in democratic elections carries equal weight despite differences in income.

Conversely, our experiment also did not vary the degree to which punishment and reward affected individuals across the income spectrum differently. Future research should model and study potentially interesting policies that have varying effects on the rich and poor. For example, rich and poor participants in Experiments 2 and 3 could face the same maximum, absolute number of punishment (or reward) points, but being punish with the maximum would likely have different effects on the person depending on their income. A punishment of 4 points is less than 10% of the rich participants' income, but it equals the *total* income each round for the poorest participants. It would thus be interesting to ask whether behaviour would be affected differently if punishment were the same fraction of one's income in both cases. There exists some precedent for a proportional fine structure: in the UK, for example, speeding fines can be as high as 150% of the offender's weekly income – however, the caveat is that the maximum fine cannot exceed £2,500 (BBC 2017). While the proportionality with income has comparable effects on individuals across the income spectrum, the choice of a ceiling likely limits the punishment the richest could be exposed to in extreme circumstances. Future research should experimentally vary the degree to which policy institutions assign punishment proportional to income, or impose an upper bound.

In conclusion, a range of experiments and methodologies—from incentive-compatible economic games to psychological vignettes—demonstrate that preferences for who to sanction changed when incomes were made transparent, and that revealing incomes decreased inequality and increased total contributions. To highlight, some of our results speak to the concerns of policymakers: our final experiment shows that revealing information about inequality and wealth is an important factor in people's perceptions of where resources ought to be allocated to sustain a public good. We return to our introductory example to speculate what revealing inequality might look like in the real world: while revealing all citizens' incomes may seem challenging to implement or even hard to imagine in some countries, it is common practice in others – as in the Norway example cited earlier. In a world of income transparency, the "haves" may become less rewarded and the "have-nots" less punished, with implications for the common good.

References

Abeler, J. et al., 2011. Reference Points and Effort Provision. *American Economic Review*, 101(2), pp.470–492.

Akinola, M. & Mendes, W.B., 2013. It's Good to Be the King: Neurobiological Benefits of Higher Social Standing. *Social Psychological and Personality Science*, 5(1), pp.43–51.

Alesina, A. & Angeletos, G.-M., 2005. Fairness and Redistribution. *American Economic Review*, 95(4), pp.960–980.

Anderson, L.R., Mellor, J.M. & Milyo, J., 2008. Inequality and public good provision: An experimental analysis. *The Journal of Socio-Economics*, 37(3), pp.1010–1028.

BBC, 2017. Speeding fines rise for most serious cases. *BBC.co.uk*. Available at: https://www.bbc.co.uk/news/uk-38724301.

Bloom, M. & Michel, J.G., 2002. The relationships among organizational context, pay dispersion, and among managerial turnover. *Academy of Management Journal*.

Breza, E., Kaur, S. & Shamdasani, Y., 2017. The Morale Effects of Pay Inequality. *The Quarterly Journal of Economics*, 133(2), pp.611–663.

Brown-Iannuzzi, J.L. et al., 2015. Subjective Status Shapes Political Preferences. *Psychological Science*, 26(1), pp.15–26.

Buhrmester, M., Kwang, T. & Gosling, S.D., 2011. Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, 6(1), pp.3–5.

Cappelen, A.W. et al., 2013. Just Luck: An Experimental Study of Risk-Taking and Fairness. *American Economic Review*, 103(4), pp.1398–1413.

Cappelen, A.W. et al., 2007. The Pluralism of Fairness Ideals: An Experimental Approach. *American Economic Review*, 97(3), pp.818–827.

Card, D. et al., 2012. Inequality at Work: The Effect of Peer Salaries on Job Satisfaction. *American Economic Review*, 102(6), pp.2981–3003.

Chou, E.Y., Parmar, B.L. & Galinsky, A.D., 2016. Economic Insecurity Increases Physical Pain. *Psychological Science*, 27(4), pp.443–454.

Congressional Budget Office, 2007. Average Federal Tax Rates. Available at: http://www.cbo.gov/publication/42870.

Crockett, M.J. et al., 2014. Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences*, 111(48), pp.17320–17325.

Davidai, S. & Gilovich, T., 2015. Building a More Mobile America—One Income Quintile at a Time. *Perspectives on Psychological Science*, 10(1), pp.60–71.

Fehr, E. & Gächter, S., 2002. Altruistic punishment in humans. *Nature*, 415(6868), pp.137–140.

Frohlich, N., Oppenheimer, J.A. & Eavey, C.L., 1987. Choices of principles of distributive justice in experimental groups. *American Political Science Review*, 31(3).

Gächter, S. et al., 2017. Growth and Inequality in Public Good Provision. *Journal of Public Economics*.

Harrison, G.W. & List, J.A., 2004. Field Experiments. *Journal of Economic Literature*, 42(4), pp.1009–1055.

Hauser, O.P. & Norton, M.I., 2017. (Mis)perceptions of inequality. *Current Opinion in Psychology*, 18, pp.21–25.

Hauser, O.P. et al., 2016. Think Global, Act Local: Preserving the Global Commons. *Scientific Reports*, 6.

Hauser, O.P., Gino, F. & Norton, M.I., 2019. *Budging beliefs, nudging behaviour*,

Hauser, O.P., Linos, E. & Rogers, T., 2017. Innovation with field experiments: Studying organizational behaviors in actual organizations. *Research in Organizational Behavior*, 37, pp.185–198.

Hauser, O.P., Nowak, M.A. & Rand, D.G., 2014. Punishment does not promote cooperation under exploration dynamics when anti-social punishment is possible. *Journal of Theoretical Biology*, 360(C), pp.163–171.

Hauser, O.P., Traulsen, A. & Nowak, M.A., 2014. Heterogeneity in background fitness acts as a suppressor of selection. *Journal of Theoretical Biology*, 343, pp.178–185.

Jachimowicz, J., To, C. & Hauser, O., 2018. Team Pay Inequality: Pay Disparity In Large Teams Increases Employee Turnover By Reducing Job Control. *Working Paper*.

Jachimowicz, J.M. et al., 2017. Community trust reduces myopic decisions of low-income individuals. *Proceedings of the National Academy of Sciences*.

Jachimowicz, J.M., Hauser, O.P., et al., 2018. The critical role of second-order normative beliefs in predicting energy conservation. *Nature Human Behaviour*, 2, pp.757–764.

Jordan, J.J. et al., 2016. Third-party punishment as a costly signal of trustworthiness. *Nature*, 530(7591), pp.473–476.

Kiatpongsan, S. & Norton, M.I., 2014. How Much (More) Should CEOs Make? A Universal Desire for More Equal Pay. *Perspectives on Psychological Science*, 9(6), pp.587–593.

Konow, J., 2000. Fair shares: Accountability and cognitive dissonance in allocation decisions. *The American Economic Review*, 90(4), pp.1072–1091.

Kraft-Todd, G.T., Bollinger, B., Gillingham, K., Lamp, S. and Rand, D.G., 2018. Credibility-enhancing displays promote the provision of non-normative public goods. *Nature*, 563(7730).

Kraft-Todd, G., Yoeli, E., Bhanot, S. and Rand, D., 2015. Promoting cooperation in the field. Current Opinion in Behavioral Sciences, 3, pp.96-101.

Krasnow, M.M. et al., 2016. Looking Under the Hood of Third-Party Punishment Reveals Design for Personal Benefit. *Psychological Science*, pp.1–14.

Kraus, M.W. & Mendes, W.B., 2014. Sartorial symbols of social class elicit class-consistent behavioral and physiological responses: A dyadic approach. *Journal of Experimental Psychology: General*.

Kraus, M.W., Piff, P.K. & Keltner, D., 2009. Social class, sense of control, and social explanation. *Journal of Personality and Social Psychology*, 97(6), pp.992–1004.

Kuziemko, I. et al., 2015. How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments. *American Economic Review*, 105(4), pp.1478–1508.

Lago-Peñas, I. & Lago-Peñas, S., 2010. The determinants of tax morale in comparative perspective: Evidence from European countries. *European Journal of Political Economy*, 26(4), pp.441–453.

List, J., 2006. Field Experiments: A Bridge between Lab and Naturally Occurring Data. *Advances in Economic Analysis & Policy*, 6(2).

Ludwig, J., Kling, J.R. & Mullainathan, S., 2011. Mechanism experiments and policy evaluations. *Journal of Economic Perspectives*, 25(3).

Maryland CPA, 2016. Charitable deduction averages 2014. *Marylands Premier Group of Certified Public Accountants*. Available at: http://www.maryland-cpa.com/news/charitable-deduction-averages-2014/ [Accessed April 2017].

Nikiforakis, N., Noussair, C.N. & Wilkening, T., 2012. Normative conflict and feuds: The limits of self-enforcement. *Journal of Public Economics*, 96(9-10), pp.797–807.

Nishi, A. et al., 2015. Inequality and visibility of wealth in experimental social networks. *Nature*, pp.1–14.

Norton, M.I., 2014. Unequality: Who Gets What and Why It Matters. *Policy Insights from the Behavioral and Brain Sciences*, 1(1), pp.151–155.

Norton, M.I. & Ariely, D., 2011. Building a Better America—One Wealth Quintile at a Time. *Perspectives on Psychological Science*, 6(1), pp.9–12.

Norwegian Tax Administration, 2015. Norway Income Database. Available at: http://www.skatteetaten.no/nn/Person/Skatteoppgjer/Sok-i-skattelistene/Hva-star-i-skattelistene/.

Oishi, S., Kesebir, S. & Diener, E., 2011. Income Inequality and Happiness. *Psychological Science*, 22(9), pp.1095–1100.

Payne, B.K., Brown-Iannuzzi, J.L. & Hannay, J.W., 2017. Economic inequality increases risk taking. *Proceedings of the National Academy of Sciences*, 114(18), pp.4643–4648.

Pfeffer, J. & Langton, N., 1993. The Effect of Wage Dispersion on Satisfaction, Productivity, and Working Collaboratively: Evidence from College and University Faculty. *Administrative Science Quarterly*, 38(3), pp.382–407.

Rand, D.G., 2012. The promise of Mechanical Turk How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology*, 299(C), pp.172–179.

Rand, D.G. et al., 2009. Positive interactions promote public cooperation. *Science*, 325, pp.1272–1275.

Rawls, J., 1971. *A theory of justice*, Belknap.

Rockenbach, B. & Milinski, M., 2006. The efficient interaction of indirect reciprocity and costly punishment. *Nature*, 444(7120), pp.718–723.

Sands, M.L., 2017. Exposure to inequality affects support for redistribution. *Proceedings of the National Academy of Sciences of the United States of America*.

Shaw, J.D., 2014. Pay Dispersion. *Annual Review of Organizational Psychology and Organizational Behavior*, 1, pp.521–544.

Sheehy-Skeffington, J., Kteily, N. & Hauser, O.P., 2016. Antecedents and consequences of perceptions of economic inequality. *Tobin Project Conference on Inequality Decision-Making, Paper Presentation*.

Sullivan, M. & Felton, E., 2014. PTA fundraising data shows massive gap between haves and have-nots. *The New York World*. Available at: http://www.thenewyorkworld.com/2014/08/06/pta-2/ [Accessed April 29, 2017].

Thaler, R.H. & Sunstein, C.R., 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*, New Haven: Yale University Press.

U.S. Census Bureau, 2013. *Income and Poverty in the United States*, Available at: http://www.census.gov/hhes/www/income/data/incpovhlth/2013/table4.pdf.

Van Dijk, E. & Wilke, H., 1994. Asymmetry of wealth and public good provision. *Social Psychology Quarterly*.

Weissman Center for International Business, 2016. New York City Median Household Income. *Baruch College CUNY*. Available at: http://www.baruch.cuny.edu/nycdata/income-taxes/med_hhold_income.htm [Accessed April 29, 2017].

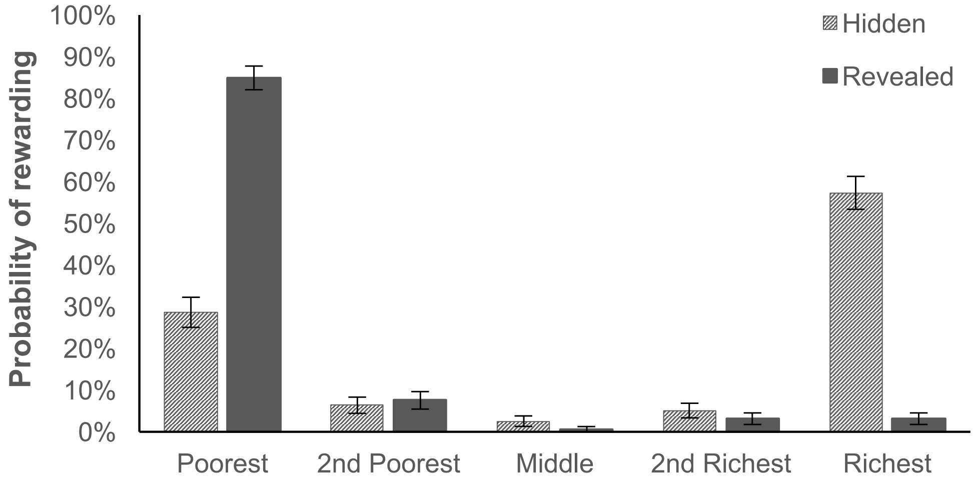


Figure 1. Reward behaviour shifts from rewarding the richest donor who donates the highest absolute amount in the* hidden *condition to the poorest donor who donates the largest relative amount in the* revealed *condition.

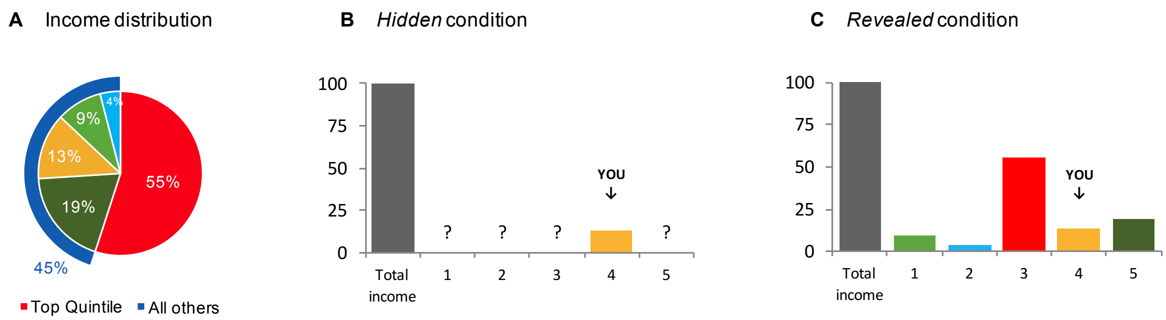


Figure 2. The income distribution in our game and the main experimental manipulation between the* hidden *and* revealed *conditions. ***A*** *Each player in a group of five was randomly assigned to a position in an income distribution. In the first experiment, we used the 2007 U.S. pre-tax income distribution* (Congressional Budget Office 2007)*: in each of ten rounds, the top quintile participant received 55 units, while the bottom quintile player received 4 units.* ***B*** *When making decisions to punish and reward, participants in the* hidden *condition saw their own income and the sum of all incomes.* ***C*** *In the* revealed *condition, participants viewed all players' incomes.*

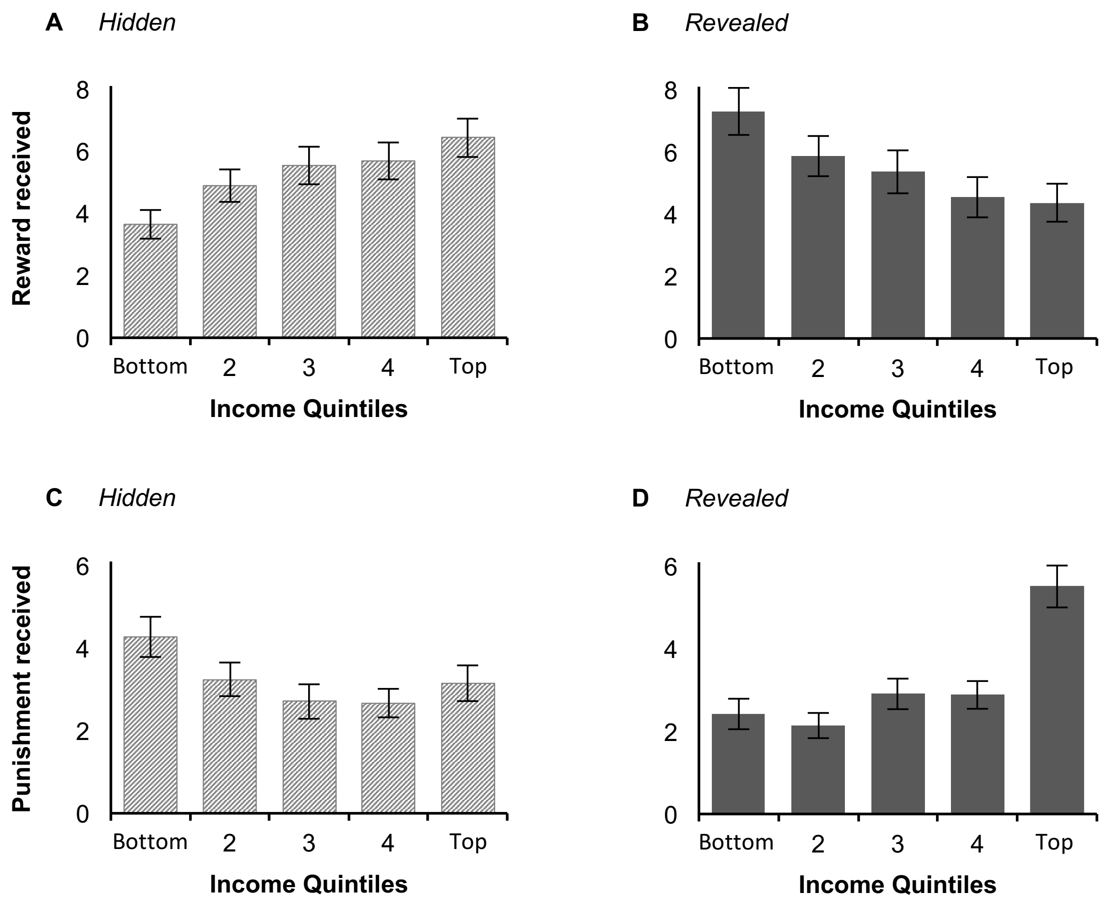


Figure 3. Amount of received reward (top panels) and punishment (bottom panels) depends on income quintile and whether income was* hidden *(left panels) or* revealed *(right panels). A *Participants rewarded higher income participants more in the* hidden *condition, but* ***B*** *less in the* revealed *condition.* ***C*** *Punishment behaviour is a mirror image of reward: participants punished poorer participants more in the* hidden *condition, while* ***D*** *punishing richer participants more in the* revealed *condition.*

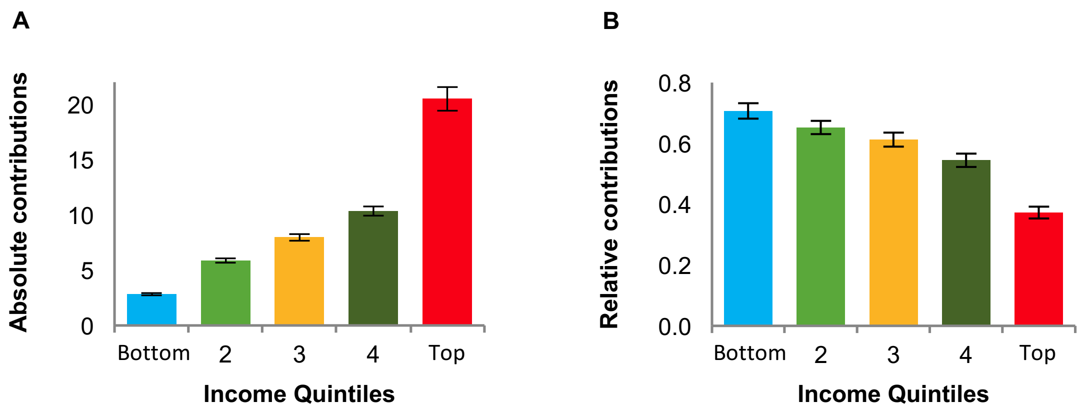


Figure 4. Who contributes more? A *In the* hidden *condition, only absolute contributions could be assessed, such that richer participants appeared to contribute more.* ***B*** *In the* revealed *condition, where participants could view contributions relative to income, it became apparent that lower income participants contributed a larger fraction of their income.*

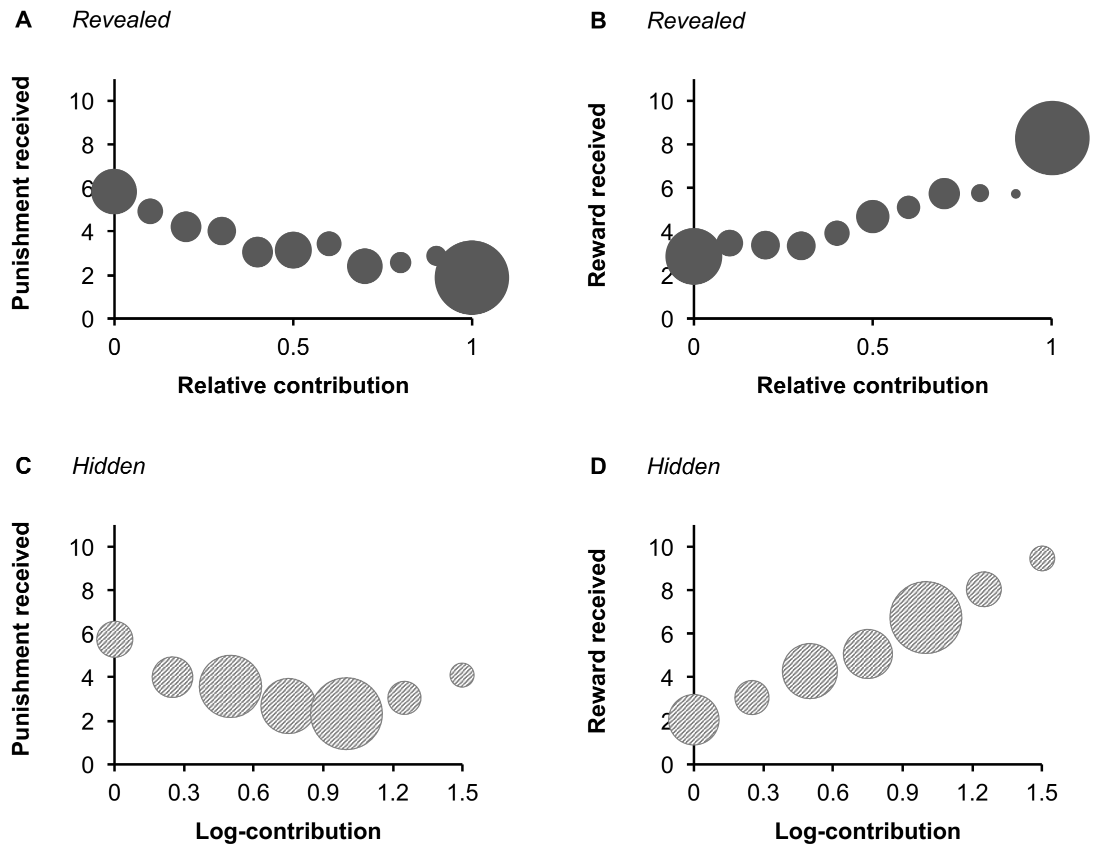


Figure 5. Received punishment and reward depends on percentage of income contributed in the* revealed *condition (top panels) and on absolute income contributed in the* hidden *condition (bottom panels). ***A,B*** *When incomes were revealed, participants who contributed a higher percentage of their income were* ***A*** *punished less and* ***B*** *rewarded more.* ***C,D*** *Conversely, when incomes were hidden, participants who contributed a higher absolute amount were* ***C*** *punished less and* ***D*** *rewarded more. Bubble size is proportional to the fraction of corresponding participants.*

Supplementary Material

for

Invisible Inequality Leads to Punishing the Poor and Rewarding the Rich

Table of Contents:

1. Motivation and relation to previous work 2
   1.1 Research motivation 2
   1.2 Previous research 3
2. Methods 5
   2.1 Data collection on Amazon Mechanical Turk 5
   2.2 Basic flow of the experiments 5
3. Statistical details 12
   3.1 Experiment 1 12
   3.2 Experiment 2 13
   3.3 Experiment 3 29
   3.4 Experiment 4 39
4. Instructions 42
   4.1 Experiment 1 42
   4.2 Experiment 2 44
   4.3 Experiment 3 52
   4.4 Experiment 4 58
References 60

1. Motivation and relation to previous work

1.1 Research motivation

The central observation motivating the present work is that people, both in the U.S. and in 39 other countries, systematically underestimate the extent of inequality in their country (Kiatpongsan & Norton, 2014; Norton & Ariely, 2011). We hypothesised that this misinformation regarding inequality can have major negative impacts on societal well-being: if people do not realise how little the poor have, and how much the rich have, they may be less sympathetic to low contributions from those who cannot afford to give more, and less able to hold the rich to account for not contributing their "fair share".

To explore these societal impacts experimentally, we used the standard paradigm from experimental economics for studying group social interactions: the public goods game (PGG). In particular, we built on the large body of prior work demonstrating that people tend to punish players who do not contribute in the PGG, and reward players who do (Fehr & Gächter, 2000; 2002; Gächter, Renner, & Sefton, 2008; Herrmann, Thöni, & Gächter, 2008; Rand, Dreber, Ellingsen, Fudenberg, & Nowak, 2009; Sefton, Shupp, & Walker, 2007; Sutter, Haigner, & Kocher, 2010). (There are, however, cultural differences across countries: while punishment of low contributors is the norm in most Western countries, so-called 'anti-social' punishment aimed at high contributors is observed at high frequency in countries with a weak rule of law (Herrmann et al., 2008).) While most prior PGG studies have focused on groups where incomes were equally distributed, several recent studies have begun to explore how inequality affects contribution and sanctioning behaviour (Buckley & Croson, 2006; Reuben & Riedl, 2013) (see Section 1.2 for more details).

In the current paper, we add to our understanding of inequality by incorporating three key features of inequality that have received little prior attention. First, in most prior work, the income distribution was common knowledge among all PGG groups members. Thus, little is known about our central question of the consequences of the empirical observation that people do not have an accurate understanding of the level of inequality (Kiatpongsan & Norton, 2014; Norton & Ariely, 2011). To that end, we experimentally manipulate when the income distribution is known or hidden.

Secondly, most prior studies focus on the impact of levels of inequality that were much lower than what is observed outside the laboratory. The Gini index is the most a common measure of inequality (Allison, 1978). Using the most recent country-level data from the World Bank (World Bank, n.d.), we found that globally, the Gini index ranges between 0.25 and 0.66. Almost no prior studies used endowments that reflected that level of inequality: while 90% of all countries had Gini indices higher than 0.29, 91% of previous PGG experiments we surveyed had a Gini index below 0.29. To better reflect the reality of income inequality, we used PGG endowments that match the actual U.S. income quintiles from 2007 having a Gini index of 0.440 (Experiment 1), and 2013 having a Gini index of 0.444 (Experiment 2).

Third, most prior studies have randomly assigned subjects to receive higher or lower incomes. Yet in reality, variation in income is (at least in part) determined by non-random factors such as effort. Thus we also explore the impact of earned vs random inequality, and how this interacts with knowledge about the distribution of incomes.

Finally, while most prior experiments have either not allowed for sanctioning, or have focused exclusively on punishment, we examine both costly peer punishment and costly peer rewarding through different operationalisations. Rewards (e.g. "positive" sanctions) play a key role in much of human social life, and thus understanding whether inequality impacts rewards differently from punishment is important.

1.2 Previous research

The vast majority of prior literature using public goods game has focused on equally-endowed participants. In recent years, however, economists and psychologists have begun to study the effects of endowment (or income) inequality on cooperation. Here we provide a brief survey of prior literature on inequality and public goods. (While we focus on *income* inequality, we note that inequality can also be induced by varying the marginal per capita return (Cardenas, Stranlund, & Willis, 2002; Fisher, Isaac, Schatzberg, & Walker, 1995; Reuben & Riedl, 2008), show-up fees for participants (Anderson, Mellor, & Milyo, 2008), or marginal abatement costs (Brick & Visser, 2012), or by taking advantage of endogenous variation in participants' real-world wealth (Cardenas, 2007; 2003).)

One of the most consistent findings has been that exogenous inequality in incomes leads to lower levels of overall contributions in a group (Buckley & Croson, 2006; Cherry, Kroll, & Shogren, 2005; Isaac & Walker, 1988; Keser, Markstädter, Schmidt, & Schnitzler, 2014b). The reduction in contributions is driven primarily by richer participants contributing less so as to match the level of contributions of the poorer participants (Buckley & Croson, 2006). Consequently, wealthier participants contribute less relative to their income than do poorer participants (Buckley & Croson, 2006; Chan, Mestelman, Moir, & Muller, 1996; Keser et al., 2014b).

Several researchers have aimed to introduce interventions to increase cooperation between unequal group members. Institutional fines for low contributors (Brick & Visser, 2012), minimum contribution requirements (Keser, Markstädter, & Schmidt, 2014a), communication (Brick & Visser, 2012; Chan, Mestelman, Moir, & Muller, 1999; Hackett, Schlager, & Walker, 1994) and punishment (Antinyan, Corazzini, & Neururer, 2015; Bornstein & Weisel, 2010; Reuben & Riedl, 2013) can play an important role in sustaining contributions.

Of particular relevance for the current paper, Reuben and Riedl (2013) demonstrated that punishment can stabilise contributions from participants with different incomes. In their experiments, participants used punishment in order to enforce both efficiency norms (everyone contributing their entire income) as well as relative contribution norms (everyone contributing the same fraction of their income); Carpenter and Matthews

(2009) found similar results among equally-endowed players (Carpenter & Matthews, 2009). Similarly, Antinyan et al. (2013) and Bornstein and Weisel (2010) showed that punishment was effective in sustaining contributions in unequal groups under fully informed condition. (Note that while Bornstein and Weisel (2010) also investigate a situation in which players' incomes in each round were not observable, the *distribution* of incomes is always known in their experiment, and each subject typically has experience receiving each possible level of income.)

Importantly, research using PGGs without inequality shows that incomplete information about contributions can lead to increased spending on punishment (Ambrus & Ben Greiner, 2012) and lower average payoffs (Grechenig & Nicklisch, 2010). Thus, it seems likely that prior conclusions about punishment and inequality may change when the income distribution is unknown.

Furthermore, some research indicates that the origin of incomes can matter. Most studies randomly assign incomes to participants. However, various theories of fairness suggest that earning incomes by exerting individual effort can lead to more acceptance of inequality. The experimental literature has led to mixed results on this conjecture: Van Dijk and Wilke (1994) found that participants who were made to believe that their group members earned more money by exerting more effort were more likely to contribute more to the public good, and vice versa (Van Dijk & Wilke, 1994).

In contrast, Cherry et al. (2005), Hofmeyr et al. (2007) and Antinyan et al. (2015) found that public goods contributions were not significantly different when endowments were earned or received as a 'windfall' (Antinyan et al., 2015; Cherry et al., 2005; Hofmeyr, Burns, & Visser, 2007). In some cases, punishment towards low contributors can be reduced when income was earned (Antinyan et al., 2015). Thus, understanding how earned incomes affects cooperation and sanctioning, and how it might interact with (lack of) knowledge of the income distribution, remains an open question.

2. Methods

2.1 Data collection on Amazon Mechanical Turk

We recruited U.S. residents to participate using the online labour market Amazon Mechanical Turk (AMT). AMT is an online market place in which employers can pay users for completing short tasks (generally about 10 minutes) – usually referred to as Human Intelligence Tasks (HITs) – for a relatively small payment (generally less than a $1). Workers who have been recruited on AMT receive a baseline payment and can be paid a bonus depending on their performance in the task. This setup lends itself to incentivised economic experiments: the baseline payment acts as the 'show up' fee and the bonus payment may derive from the workers' behaviour in the economic game and/or other tasks throughout the experiment.

The sample of recruited participants on AMT has been shown to be more diverse and more nationally representative than subject pools at most research universities (Buhrmester, Kwang, & Gosling, 2011). Numerous studies have been carried out to validate results collected using AMT (Berinsky, Huber, & Lenz, 2012; Crump, McDonnell, & Gureckis, 2013; Paolacci & Chandler, 2014). Of particular relevance are studies showing quantitative agreement between play in economic games conducted on AMT and in the physical laboratory (Amir, Rand, Gal, & Gal, 2012; Horton, Rand, & Zeckhauser, 2011; Mason & Suri, 2011; Rand, Greene, & Nowak, 2012).

All data was collected using Software Platform for Human Interaction Experiments (SoPHIE) (Hendriks, 2012). Experiment 1 was carried out in summer 2014, while Experiment 2 was conducted in summer 2015. SoPHIE is a novel experimental platform that enables participants to interact with one another in real time. Participants were recruited on the AMT website, were grouped together, and then made decisions simultaneously; their decisions were exchanged through an external server provided by SoPHIE Labs (www.sophielabs.net).

The experiments were approved by the Harvard University Committee on the Use of Human Subjects in Research.

2.2 Basic flow of the experiments

2.2.1 Experiment 1: Charity survey experiment

2.2.1.1 Donor Survey. Before launching the main experiment, we recruited participants who fit the following demographic information. First, participants were asked whether they earned either "under $25,000", "$25,000–$50,000", "$50,000–$75,000", "$75,000–$100,000" or "$100,000–$200,000". Depending on their income selection, they were then asked whether last year they donated approximately $1,874, $2,594, $2,970, $3,356 or $4,130, respectively. The five donation amounts were chosen to represent the average U.S. donations made in 2014 within those income ranges (Maryland CPA, 2016). We selected at random five donors who fit these criteria. These participants made up the recipient pool eligible for a bonus payment. The bonus was paid in accordance with a decision by another participant, chosen at random from the main experiment.

2.2.1.2 Main Experiment. We recruited a (non-overlapping) sample of 315 participants on AMT for an academic survey. We chose 150 observations per condition as our sample size to ensure we would be able to detect relatively small effect sizes. The final sample size was slightly above this target ($N = 315$).

All participants earned \$0.40 for participating. Participants agreed to a consent form and read the instructions of the study. The instructions informed the participant that we had previously recruited five donors on AMT and they would now choose to whom of the five participants they would award a \$1.00 bonus payment. We varied the information that participants received about those five donors between two conditions.

In the *hidden* condition, participants saw only the donation amount that the donors gave to charity last year. Each donor was identified by a random string of two letters which we fully randomized across the five donors. Conversely, participants in the *revealed* condition saw the two-letter identifier, donation amount and the income range for each donor. Thus, participants in the *revealed* condition could assess approximately how much of their income a donor had given to a charity, whereas they were not able to do so in the *hidden* condition.

As described above (and in the instructions to participants), the decision was incentive compatible: one participant was chosen at random and their decision implemented as described.

2.2.2 Experiment 2: PGG with 'windfall' incomes

All participants earned a \$1.50 showup fee and had the opportunity to earn additional bonus payments between \$0.00 and \$3.88, depending on the outcome of the game. Participants took part in the experiment through an online survey provided by SoPHIE Labs. After participants had read the experimental instructions, they had to pass a comprehension quiz about the rules of the game in order to partake in the actual experiment. Participants who did not pass the comprehension quiz on the first attempt were given the chance to try again; no participants were thus removed from the experiment.

After participants passed all comprehension questions, they waited for up to 10 minutes in a designated online 'waiting room.' As soon as five participants had arrived in the waiting room, the public goods game (PGG) started automatically.

We recruited at least 30 groups of five participants who completed the experiment in each condition (N = 600; including drop-out groups, N = 855). Our sample size of 150 participants per condition was based on previous studies and the feasibility of collecting interactive group decisions online. At the beginning of the game, participants were randomly assigned to a position in the income distribution. This implied that participants earned a 'windfall' endowment, which can in some cases affect contribution behaviour in public goods games (Van Dijk & Wilke, 1994). In experiment 3, we show, however, that assigning incomes based on performance, instead of randomly assigning incomes, does not alter our reward and punishment results.

The income distribution was common knowledge to all participants in the *revealed* condition, but not in the *hidden* condition. Across conditions, the actual distribution (Congressional Budget Office, 2007) is the same: the participant with the highest income (referred to as top quintile player) earns 55 units per round, the second-highest earner 19 units, the middle participant 13 units, the second-lowest 9 units and the bottom quintile player 4 units. That is, the sum of income units distributed each round is 100 units, and this total is also known to all players in all conditions. Once incomes had been assigned, participants received the same income in each of the 10 rounds.

The game lasted 10 rounds and each round consisted of two stages. Participants were given no information about the length of the game to avoid any end-round effects, as in prior work (Rand et al., 2009; Rand, Nowak, Fowler, & Christakis, 2014). In stage 1, participants were asked how many units they wanted to contribute to the public good. All contributed units were doubled and then split equally among all five players. In stage 2, participants were either participating in a punishment-game or a reward-game. In the punishment game, each player could pay 1 unit to take away 3 units from another player. In the reward game, conversely, players could pay 1 units to increase another player's payoff by 3 units.

Across conditions, we limited the number of units that could be spent on punishment and reward to 2 per target player. In addition, no participant could spend more than they had accumulated with their income and their earnings from stage 1. In other words, participants could spend up to 8 units per round on the other 4 players but if they had less than 8 units in their account, the upper bound of spending was their remaining endowment. To examine whether this upper bound could affect the decisions participants were able to make (e.g., consistently prevented them from punishing or rewarding), we examined all players' payoffs after stage 1 across all conditions. We find that only 0.84% of the time did a participant have less than 8 units available, and the number of affected participants does not vary by condition ($\chi^2(21) = 23.16$, $p = 0.34$). Given the small number of incidences in which participants were constrained in their decisions to punish and reward, it is unlikely that they affect our results.

At the end of each round, participants were informed about the payoff they earned. At the end of 10 rounds, participants filled out a short demographics survey. If at any time a participant became unresponsive (because they quit the game or lost their Internet connection), the remaining participants in the group were automatically moved to an 'early exit' screen. They were informed that a participant had left the game unexpectedly and thus the experiment could not be continued. All remaining participants were asked to fill out a questionnaire and earn a bonus of $1.00 to compensate them for their time spent on the study.

Across all conditions, in 29.8% of groups, a participant dropped out before the end of the game (which is consistent with previous studies carried out on AMT, e.g. (Rand et al., 2012). Dropout rates did not vary between conditions ($\chi^2(3) = 2.74$, $p = 0.433$). In half of the groups that dropped out, one or more participants did not respond within the time available to them (up to two minutes per decision stage). In the other half, one of the participants in each dropout group quit their browser or tab (either by choice or due to a failed Internet connection). Dropout rates did not differ significantly between quintiles ($\chi^2(12) = 7.89$, $p = 0.793$). The majority of dropouts (73.9%) occurred in the first half of the game and the time of dropout did not differ by quintile ($\chi^2(21) = 23.123$, $p = 0.337$).

We further analysed participants' likelihood of dropping out based on the average sum of contributions in the group or the rewards and punishment a participant received. Across conditions, the sum of contributions did not affect whether a participant finished the game (logistic regression using sum of a group's contributions to predict finishing the game: coeff = 0.007, $p = 0.340$). There was not interaction between the sum of contributions and income visibility (logistic regression using sum of contributions interacted with *revealed* dummy: coeff = 0.008, $p = 0.614$).

The amount of reward or punishment received did not affect a participant's likelihood of finishing the game (logistic regression using sanctions received to predict finishing the game; reward: coeff = 0.023, $p = 0.618$; punishment: coeff = -0.048, $p = 0.093$). Income visibility did not affect dropout rates: the amount a participant was rewarded or punished did not significantly affect their chances of finishing the game (logistic regression using sanctions received interacted with *revealed* dummy; reward: coeff = 0.135, $p = 0.101$; punishment: coeff = -0.028, $p = 0.600$).

In our main analysis, we include groups that did not finish the game. However, we found qualitatively similar results when we dropout groups are not included in the analysis.

2.2.3 Experiment 3: Earned incomes

All participants earned a \$2.00 show-up fee and had the opportunity to earn an additional in bonus payments between \$0.00 and \$4.26 depending on the outcome of the game. Participants took part in the experiment through SoPHIE. The experiment consists of two phases.

In phase 1, participants completed an individual task (Abeler, Falk, Goette, & Huffman, 2011). They were asked to count the number of 0s in a matrix randomly made up of 0s and 1s. The goal was to complete as many such matrices as possible within 2 minutes. Each time, after they submitted a solution, a new matrix was presented to them. Participants were not informed of how many matrices they had solved correctly. However, they were told that their performance mattered for their income in the upcoming group task. The best performing participant in the individual task was assigned the highest income level; the second-best performing participant received the second-highest level of income, and so on. Participants were fully informed in both conditions about the assignment procedure of incomes.

In phase 2, participants then read instructions about an economic game. After reading the instructions, the were asked to answer several comprehension questions. Once participants had correctly answered the comprehension questions, they waited for four other participants to arrive in the 'waiting room'. As soon as five participants were ready, the began the repeated two-stage economic game comprising of a public goods stage and a sanctioning stage.

We collected 30 groups of five participants who completed the game in each condition ($N = 300$; including dropout groups, $N = 440$). We excluded participants from experiment 2 from participating in experiment 3. While conceptually similar, the economic game in experiment 2 differed from the economic game in experiment 1 in several ways. First, the income distribution used in experiment 3 was updated to reflect the latest U.S. census data (U.S. Census Bureau, 2013): the top quintile player received 51 units, the player in the next quintile 23 units, the middle-quintile player 15 units, the second-lowest player 8 units and the lowest earner 3 units.

In addition, participants in this experiment knew that the income distribution used in this game represented the current income distribution of the United States. Unlike in experiment 3, participants thus had an implicit reference point, which they could use to make informed reward and punishment decisions assuming that their estimate of U.S. income inequality is accurate. Participants in experiment 3 were not given any hint about the distribution used and only knew that their incomes would be different from one another. Based on previous research (Norton & Ariely, 2011), our prediction was that participants in experiment 3 would, however, not hold accurate beliefs about the extent of U.S. inequality and it would thus qualitatively not alter the way in which people punished or rewarded.

Furthermore, the reward and punishment conditions in experiment 3 were collapsed into one combined treatment: it was thus possible for participants to punish or reward in all conditions in experiment 3. Each player could choose not to pay any units and thus leave another player's payoff unaffected; or they could choose to pay 2 units to reduce another player's by 6 units or pay 2 units to increase another player's by 6 units. Participants could spend up to 8 units on all four other players or up to as much as they had earned in stage 1. Only 2.33% of the time a participant was restricted in punishing or rewarding their group members because they had less than 8 units available in stage 2, which did not vary by condition ($\chi^2(9) = 12.532$, $p = 0.185$).

Finally, we adopted the standard infinite-game paradigm used in economics (Dal Bo, 2005): participants were told that the game would last at least 8 rounds and each additional round would occur with a probability of 50% to avoid end-game effects (Dal Bo, 2005; Rand et al., 2009). Due to a programming error, we did not collect demographic information from participants in experiments 3. However, since participants were randomly assigned to the *hidden* or *revealed* condition, there should not be any systematic variation in demographics across conditions. Furthermore, when we controlled for demographics in experiment 2, we found that it did not affect our results.

Dropout rates in experiment 3 were comparable to those in experiment 2: 31.82% of groups did not finish the game in experiment 3. Dropout rates did not differ by condition ($\chi^2(1) = 0.098$, $p = 0.755$). Across quintiles, there was no variation in the rate of dropout ($\chi^2(4) = 3.280$, $p = 0.512$). The majority of dropouts (71.4%) occurred in the first five rounds of the game; and the time of dropout did not differ by condition ($\chi^2(4) = 4.278$, $p = 0.370$).

The behavioural history of participants did, however, matter to whether a group finished the game in experiment 3. While the sum of contributions across both condition did not affect a participant's likelihood of finishing (logistic regression using sum of group contributions to predict finishing the game: coeff = 0.013, $p = 0.112$), there was an interaction between the sum of contributions and income visibility: participants in the *revealed* condition were more likely to finish the game, the more the group contributed to the common pool, while there was no effect of sum of contributions in the *hidden* condition (logistic regression using sum of contributions as IV to predict finishing the game: coeff = -0.014, $p = 0.186$; sum of contributions interacted with *revealed* dummy: coeff = 0.043, $p = 0.004$).

The amount of reward or punishment participants received also affected their dropout rates in experiment 3. Across both conditions, participants were not significantly affected by rewards and punishment received (logistic regression using number of units received to predict finishing the game: coeff = 0.180, $p = 0.094$). However, a difference emerged by condition: whereas groups in the *revealed* condition were more likely to finish the game when they received more rewards, whereas there was no effect of units received in the *hidden* condition (logistic regression using number of units received to predict finishing the game: coeff = -0.168, $p = 0.287$; sum of contributions interacted with *revealed* dummy: coeff = 0.66, $p < 0.001$).

Since neither sum of contributions nor received reward or punishment predicted dropout rates in the *hidden* condition, we can be reasonably confident that selection effects are not driving our results in the *hidden* condition. However, in the *revealed* condition, groups in which fewer contributions were made altogether and groups in which fewer reward units were distributed were more likely to drop out. Thus, to mitigate concerns about potential selection effects, we include all groups in our main analyses, regardless of whether they completed the game or dropped out. However, we found qualitatively similar results when we did not include dropout groups.

2.2.4 Experiment 4: school donations and tax choice

We recruited $N = 313$ participants on Amazon Mechanical Turk to take part in an academic survey. After consenting to participating in the research, participants read a vignette about the Parent-Teacher Associations (PTA). The PTA is an annual fundraiser organized voluntarily across the United States, soliciting donations from households in the area (in particular, parents who have children in school) to contribute to an educational fund. All donations are typically pooled at the school district level and donations are distributed across schools to pay for educational expenses that the schools do not receive state or federal funding for, such as a science lab, teachers' aides, and additional equipment.

Participants were told in the experiment that the city government was planning to introduce an additional tax for educational purposes (in addition to the existing funding and PTA donations) in the amount of $2,000 per school across the five schools in the city. Participants had to choose which school parents of the five schools would have to bear the additional tax to cover this educational spending.

The donations and incomes presented in this experiment was inspired by real-world data sources: we obtained data from average PTA donations (Sullivan & Felton, 2014) as well as median incomes (Weissman Center for International Business, 2016) from five New York school districts (Bronx, Brooklyn, Manhattan, Queens, Staten Island). Because individual school level data was not available, we used the available data from the five school districts which we labelled as five schools for ease of understanding. Participants saw the five schools, which were identified by a random two-letter string and appeared in random order: $353 average PTA donation (median income: $35,176), $1,227 ($51,141), $4,249 ($60,422), $9,759 ($71,622), and $1,486 ($75,575).

Note that donation amounts do not perfectly track with household incomes: parents in Manhattan—the district with the highest median household income—give less to PTA fundraisers ($1,486) than parents in Queens ($4,249) and Staten Island ($9,759). Our predictions remain qualitatively unchanged, however: we predict that participants' punishment preferences will shift away from raising taxes on the poorest school parents in the *hidden* condition to raising them on the richest parents when incomes are revealed.

3. Statistical details

3.1 Experiment 1

We first looked at whether the average choice for which donor should receive the $1.00 bonus differed between the *hidden* and *revealed* conditions. Specifically, we coded donors in order of their income range (1 = poorest donor, 5 = richest donor). We predicted that participants would shift their choice of who to reward towards lower-income donors when incomes were revealed.

Indeed, we found that participants rewarded, on average, donors in lower income ranges more in the *revealed* condition than in the *hidden* condition (coeff = -2.250, $p < 0.001$; Table S1 col. 1). Results were qualitatively similar when control variables were included (Table S1 col. 2). A rank-sum test revealed a qualitatively similar result ($Z = -10.935$, $p < 0.001$).

Furthermore, we investigated which donors, in particular, were chosen to be rewarded between the two conditions. We found that the poorest donors were rewarded significantly more when incomes were revealed (using logit regression predicting choice of poorest donor by *revealed* dummy: coeff = 2.681, $p < 0.001$; Table S2 col. 1). Conversely, the richest donor was significantly less likely to be chosen for the reward payment when incomes were revealed (coeff = -3.716, $p < 0.001$; Table S2 col. 5). There were not significant difference for the other three donors (Table S2 cols. 2-4).

Table S1: Linear regression estimating the choice of donor (where 1 = poorest donor to 5 = richest donor) who received a reward payment by treatment.

VARIABLES	(1) Donor choice	(2) Donor choice
1=Revealed	-2.250***	-2.311***
	(0.161)	(0.160)
1=Female		0.110
		(0.163)
Age		-0.013
		(0.007)
Constant	3.561***	4.015***
	(0.144)	(0.293)
Observations	315	306
R-squared	0.386	0.402

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S2: Logit regression estimating the likelihood that the poorest donor (col. 1), the 2nd poorest donor (col. 2), the middle donor (col. 3), the 2nd richest donor (col. 4) or the richest donor (col. 5) was chosen for the reward payment.

VARIABLES	(1) Poorest	(2) 2nd Poorest	(3) Middle	(4) 2nd Richest	(5) Richest
1.Revealed	2.682***	0.189	-1.412	-0.496	-3.716***
	(0.287)	(0.445)	(1.126)	(0.583)	(0.483)
Constant	-0.912***	-2.688***	-3.644***	-2.925***	0.295
	(0.177)	(0.327)	(0.507)	(0.363)	(0.162)
Observations	315	315	315	315	315

Robust standard errors in parentheses
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

3.2 Experiment 2

Unless otherwise noted, all statistics are linear regressions with income quintile as a continuous independent variable. Because participants' decisions are not independent within each group over time, we cluster standard errors on the group. To check the robustness of our results, we also use log-transformed absolute income as continuous IV. We use log-transformed income, rather than absolute income, because the distribution of incomes is highly right skewed (the income of the rich participant is an outlier relative to the other 4 income levels).

3.2.1 Received reward and punishment

We first assessed whether income visibility affects participants' reward and punishment behaviour. We examined which player(s) participants rewarded and punished when not informed about the income distribution (*hidden* condition) compared to when they were informed (*revealed* condition). The independent variable in our main analysis was the income quintile of the recipient of the sanctions (1 to 5). The higher the participant's quintile, the higher her income: the participant in the first quintile was the poorest player (1 = bottom quintile) whereas the participant in the fifth quintile was the richest player (5 = top quintile).

We found qualitatively similar results when we used amount of income of the recipient as the independent variable, rather than quintile. To account for the fact that the distribution of incomes is highly right skewed (the income of the top quintile player is an outlier), we used log-transformed income amounts. We included the regression table of the log-transformed income models below each of the corresponding quintile models.

3.2.1.1 Reward

In the *hidden* condition, participants could not take another player's ability to contribute to the public good into account, since this information is not available to them. We thus expected participants to view the (high total) contributions of the top quintile participants favourably—leading to more reward targeted toward them. In the *revealed* condition, conversely, we predicted that participants would view the (high percentage) contributions of the bottom quintile participants favourably, inducing higher reward.

To test these hypotheses, we estimated the amount of reward that a participant receives as a function of their income quintile and whether the income distribution was hidden or revealed (Table S3). In the *hidden* condition, we found that higher income participants were rewarded significantly more (coeff = 0.636, $p < 0.001$, Table S3 col. 1), whereas in the *revealed* condition, lower income players were rewarded significantly more (coeff = -0.720, $p < 0.001$, Table S3 col. 2). Furthermore, a regression including data from both *hidden* and *revealed* conditions together showed that this difference was itself significant (interaction between income and *revealed* dummy, coeff = -1.356, $p < 0.001$; Table S3).
We also found qualitatively similar results when we included demographic information: higher income participants were rewarded more in the *hidden* condition (coeff = 0.629, $p < 0.001$, Table S3 col. 4) while, conversely, they were rewarded less in the *revealed* condition (interaction between income and *revealed* dummy: coeff = -1.320, $p < 0.001$, Table S3 col. 4).

Table S3: Linear regression model estimating the effect of a target's income quintile (i.e., their position in the income distribution) on the amount of reward they received. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Revealed*	(3) Interaction	(4) Interaction
Quintile	0.636***	-0.720***	0.636***	0.629***
	(0.141)	(0.134)	(0.141)	(0.143)
1=Revealed			4.316***	4.180***
			(0.957)	(0.961)
Quintile X Revealed			-1.356***	-1.320***
			(0.193)	(0.199)
1=Female				0.561
				(0.468)
Age				0.0368
				(0.0279)
Location				-0.809
				(0.490)
Constant	3.332***	7.648***	3.332***	2.119*
	(0.534)	(0.801)	(0.531)	(0.917)
Observations	1,735	1,690	3,425	3,381
R-squared	0.038	0.038	0.039	0.051

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Finally, we repeated the same analysis with log-transformed income as independent variable. We found qualitatively similar results: higher income participants were rewarded more in *hidden* (coeff = 1.043, $p < 0.001$, Table S4 col. 1) but, conversely, they were rewarded less in *revealed* (coeff = -1.140, $p < 0.001$, Table S4 col. 2), and the interaction between condition and income was significant (interaction between log-transformed income and *revealed* dummy, coeff = -2.183, $p < 0.001$, Table S4 col. 3). The results were qualitatively similar when demographics were included (Table S4 col. 4).

Table S4: Linear regression model estimating the effect of a target's log-transformed income on the amount of reward they received. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Revealed*	(3) Interaction	(4) Interaction
Log(income)	1.043***	-1.140***	1.043***	1.014***
	(0.223)	(0.226)	(0.221)	(0.228)
1=Revealed			5.966***	5.749***
			(1.130)	(1.142)
Log(income) X Revealed			-2.183***	-2.111***
			(0.315)	(0.328)
1=Female				0.553
				(0.470)
Age				0.0370
				(0.0276)
Location				-0.799
				(0.489)
Constant	2.509***	8.475***	2.509***	1.345
	(0.632)	(0.945)	(0.628)	(0.976)
Observations	1,735	1,690	3,425	3,381
R-squared	0.038	0.036	0.037	0.049

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.2.1.2 Punishment

We followed the same analysis procedure as above for punishment. We expected a mirror image of the results compared to reward. In the *hidden* condition, we predicted that participants would view the (low total) contributions of the bottom quintile participants unfavourably, leading them to punish them more. In the *revealed* condition, participants were able to take the participant's ability to contribute into account and thus we expected that the (low percentage) contributions of the top quintile participants would be viewed disapprovingly, leading to higher punishment.

In our analysis, the dependent variable was the amount of punishment that a participant received. The independent variables are the income quintile and whether the income distribution was hidden or revealed (Table S5). In the *hidden* condition, we found that higher income participants were punished less (coeff = -0.282, p = 0.042, Table S5 col. 1), whereas in the *revealed* condition, in contrast, higher income participants were punished more (coeff = 0.692, p < 0.001, Table S5 col. 2). Furthermore, a regression including data from both *hidden* and *revealed* conditions together showed that this difference was itself significant (interaction between income and *revealed* dummy, coeff = 0.974, p < 0.001; Table S5 col 3).

Qualitatively similar results were obtained when we included demographics in the regression: higher income participant in the *hidden* condition were punished marginally less (coeff = -0.258, p = 0.057), while, in the *revealed* condition, higher income participants were punished more (interaction between quintile and *revealed* dummy: coeff = 1.005, p < 0.001).

We found qualitatively similar results with log-transformed income: higher income participants were punished marginally less in *hidden* (coeff = -0.417, p = 0.057, Table S6 col. 1) but were punished more heavily in *revealed* (coeff = 1.212, p < 0.001, Table S6 col. 2); a difference which was itself significant (interaction between log-transformed income and *revealed* dummy, coeff = 1.629, p < 0.001, Table S6 col. 3). Results are qualitatively similar when demographics were included (Table S6 col. 4).

Table S5: Linear regression model estimating the effect of a target's income quintile on the amount of punishment they received. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Revealed*	(3) Interaction	(4) Interaction
Quintile	-0.282*	0.692***	-0.282*	-0.258
	(0.133)	(0.0982)	(0.132)	(0.134)
1=Revealed			-2.956***	-3.045***
			(0.640)	(0.665)
Quintile X Revealed			0.974***	1.005***
			(0.164)	(0.170)
1=Female				-0.185
				(0.273)
Age				0.0129
				(0.0153)
Location				0.185
				(0.307)
Constant	4.056***	1.099**	4.056***	3.614***
	(0.513)	(0.391)	(0.509)	(0.687)
Observations	1,590	1,819	3,409	3,230
R-squared	0.011	0.061	0.038	0.045

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S6: Linear regression model estimating the effect of a target's log-transformed income on the amount of punishment they received. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Revealed*	(3) Interaction	(4) Interaction
Log(income)	-0.417	1.212***	-0.417	-0.376
	(0.212)	(0.168)	(0.210)	(0.212)
1=Revealed			-4.303***	-4.459***
			(0.798)	(0.828)
Log(income) X Revealed			1.629***	1.690***
			(0.269)	(0.277)
1=Female				-0.182
				(0.276)
Age				0.014
				(0.015)
Location				0.181
				(0.306)
Constant	4.302***	-0.001	4.302***	3.788***
	(0.635)	(0.493)	(0.630)	(0.796)
Observations	1,590	1,819	3,409	3,230
R-squared	0.009	0.070	0.042	0.050

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.2.2 Absolute vs. relative contribution

What caused participants to punish and reward other players so differently in the *hidden* and *revealed* conditions? Contribution behaviour provides a potential answer. We hypothesised that the *hidden* and *revealed* conditions enabled participants to view contributions differently: in the *hidden* condition, participants could not take another player's ability to contribute into account since they did not know the income distribution. In the *revealed* condition, on the other hand, participants could evaluate the amount contributed relative to the player's income before choosing whom to punish or reward – in other words, participants could differentiate between absolute and relative contributions.

The hypothesis that relative contributions were driving the difference in sanctions between *hidden* and *revealed* generated several predictions. First, we expected that absolute contributions would be higher for higher income participants while relative contributions (as a percentage of income) would be higher for lower income participants.

Second, we expected absolute contributions to predict received reward and punishment when income is hidden, but relative contributions to predict received reward and punishment when income is revealed: participants would punish (reward) those who give little (a lot) in absolute terms more in *hidden*, while their sanctions would be driven by relative contributions in *revealed*.

3.2.2.1 Absolute vs. relative contribution by quintile

We first examined absolute contributions by quintile. We expected that higher income participants contributed a larger amount of units to the public good but a smaller fraction of their total income – such that lower income participants would contribute a larger percentage of their income.

As predicted, we found that higher income participants in both the *hidden* (coeff = 3.172, $p < 0.001$, Table S7 col. 1) and *revealed* (coeff = 4.738, $p < 0.001$, Table S7 col. 3) conditions contributed a larger number of units. Conversely, we found that higher income participants made smaller relative contributions (percentage of income contributed) in both *hidden* (coeff = -0.098, $p < 0.001$, Table S8 col. 1) and *revealed* (coeff = -0.058, $p < 0.001$, Table S8 col. 3). All results are robust to inclusion of demographic variables (Tables S5 and S8 cols. 2 and 4).

This variation in contribution across incomes can also be illustrated with an example: across all conditions, top quintile participants contributed 20.49 out of 55 units (or 37% of their income) to the public good. In contrast, bottom quintile participants contributed 2.83 out of 4 units (or 71% of their income).

In Tables S9 and S10, we repeated the same analysis with log-transformed income as the independent variable; results were qualitatively similar.

Table S7: Linear regression model estimating the effect of income on absolute contribution. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Hidden*	(3) *Revealed*	(4) *Revealed*
Quintile	3.172***	3.195***	4.734***	4.644***
	(0.257)	(0.260)	(0.341)	(0.349)
1=Female		-1.285*		0.367
		(0.645)		(0.830)
Age		0.065		-0.025
		(0.038)		(0.035)
Location		0.381		0.304
		(0.571)		(0.752)
Constant	-0.923	-2.550	-3.853***	-3.097*
	(0.476)	(1.338)	(0.637)	(1.314)
Observations	3,412	3,343	3,655	3,461
R-squared	0.233	0.241	0.316	0.312

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S8: Linear regression model estimating the effect of income on *percentage* of income contributed (relative contribution). Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Hidden*	(3) *Revealed*	(4) *Revealed*
Quintile	-0.098***	-0.097***	-0.058***	-0.061***
	(0.009)	(0.009)	(0.010)	(0.011)
1=Female		-0.050		0.025
		(0.027)		(0.029)
Age		0.002		-0.000
		(0.001)		(0.001)
Location		0.053		0.030
		(0.032)		(0.036)
Constant	0.850***	0.788***	0.773***	0.762***
	(0.032)	(0.055)	(0.036)	(0.059)
Observations	3,412	3,343	3,655	3,461
R-squared	0.141	0.153	0.047	0.053

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S9: Linear regression model estimating the effect of log-transformed income on absolute contribution. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Hidden*	(3) *Revealed*	(4) *Revealed*
Log(income)	5.416***	5.455***	8.082***	7.937***
	(0.464)	(0.467)	(0.611)	(0.626)
1=Female		-1.292*		0.472
		(0.617)		(0.785)
Age		0.060		-0.023
		(0.037)		(0.034)
Location		0.531		0.218
		(0.551)		(0.734)
Constant	-5.598***	-7.155***	-10.827***	-10.068***
	(0.918)	(1.539)	(1.211)	(1.639)
Observations	3,412	3,343	3,655	3,461
R-squared	0.253	0.261	0.344	0.339

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S10: Linear regression model estimating the effect of log-transformed income on *percentage* of income contributed. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Hidden*	(3) *Revealed*	(4) *Revealed*
Log(income)	-0.163***	-0.160***	-0.101***	-0.105***
	(0.014)	(0.015)	(0.016)	(0.017)
1=Female		-0.050		0.023
		(0.027)		(0.030)
Age		0.002		-0.000
		(0.001)		(0.001)
Location		0.049		0.031
		(0.032)		(0.036)
Constant	0.981***	0.914***	0.863***	0.857***
	(0.042)	(0.063)	(0.047)	(0.067)
Observations	3,412	3,343	3,655	3,461
R-squared	0.144	0.156	0.052	0.059

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.2.2.2 Absolute vs. relative contribution predicted sanctioning behaviour

Our hypothesis was that sanction behaviour followed from the contribution pattern. In the *hidden* condition, participants could only consider absolute contribution: thus, because richer participants contributed more units, they would be rewarded more and punished less. In the *revealed* condition, conversely, participants could consider the amount contributed relative to the amount players earn—that is, the percentage of income participants contributed. Because poorer participants contributed more relative to their income in the in the *revealed* condition, they would receive more reward and less punishment

To test these hypotheses, we examined the effect on sanctioning of the target's relative contribution (percentage of income contributed) and absolute contribution. We log-transformed absolute contribution because of the same right skew that also underlies absolute income, and we added 1 to all contributions prior to log-transforming as the log(0) is undefined (McDonald, 2014; Rand et al., 2012).

In the *hidden* condition, as predicted, we found that higher absolute contributions led to less punishment (coeff = -1.863, $p = 0.019$, Table S11 col. 1) and more reward (coeff = 4.700, $p < 0.001$, Table S11 col. 3) received. Because richer participants contributed a larger absolute number of units, they received less punishment and more reward in *hidden*. Relative contributions, in contrast, predicted neither punishment ($p = 0.954$, Table S11 col. 1) nor reward ($p = 0.677$, Table S11 col. 4) received in the *hidden* condition; this is unsurprising, given that relative contributions were not observable in the *hidden* condition. We found qualitatively similar results when including demographics (Table S11 cols. 2 and 4).

In the *revealed* condition, participants could assess both relative and absolute contributions, and we expected them to primarily pay attention to relative contribution. Indeed, we found that a higher percentage of income contributed led to less punishment (coeff = -4.664, $p < 0.001$, Table S12 col. 1) and more reward (coeff = 6.782, $p < 0.001$, Table S12 col. 4) received. This is in line with our prediction: as we have shown before, poor participants contributed a larger percentage of their income and are thus punished less and rewarded more in the *revealed* condition.

We also observed an effect of absolute contribution in the *revealed* condition, in the opposite direction of the effect in the *hidden* condition: higher absolute contributions led to more punishment (coeff = 1.365, $p = 0.003$, Table S11 col. 1) and marginally less reward (coeff = -0.980, $p = 0.080$, Table S11 col. 4). Although richer participants contributed a larger amount of units, they made low relative contributions; because those larger absolute contributions were correlated with the lowest relative contributions, larger absolute contributions were punished more and rewarded less.

Table S11: Linear regression model estimating the effect of absolute log-transformed contribution and relative contribution on received punishment (cols. 1 and 2) and reward (cols. 3 and 4) in the *hidden* condition. To deal with zero-contributions, a constant of 1 was added to all contributions before applying the log-transformation. Standard errors clustered on group.

VARIABLES	(1) Punishment received	(2) Punishment received	(3) Reward received	(4) Reward received
Log(contribution+1)	-1.863*	-1.730*	4.700***	4.592***
	(0.755)	(0.739)	(0.659)	(0.641)
Relative contribution	-0.030	-0.256	-0.216	-0.112
	(0.515)	(0.522)	(0.514)	(0.518)
1=Female		-0.243		0.711
		(0.306)		(0.366)
Age		0.021		0.018
		(0.021)		(0.023)
Location		0.272		-0.719
		(0.402)		(0.400)
Constant	4.774***	4.154***	1.624**	0.960
	(0.565)	(0.835)	(0.475)	(0.818)
Observations	1,590	1,553	1,735	1,713
R-squared	0.033	0.037	0.188	0.198

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S12: Linear regression model estimating the effect of relative contribution and absolute log-transformed contribution on received punishment (cols. 1 and 2) and reward (cols. 3 and 4) in the *revealed* condition. To deal with zero-contributions, a constant of 1 was added to all contributions before applying the log-transformation. Standard errors clustered on group.

VARIABLES	(1) Punishment received	(2) Punishment received	(3) Reward received	(4) Reward received
Relative contribution	-4.664***	-4.782***	6.320***	6.420***
	(0.489)	(0.532)	(0.993)	(1.034)
Log(contribution+1)	1.365**	1.465**	-0.980	-0.987
	(0.434)	(0.471)	(0.545)	(0.601)
1=Female		-0.213		0.036
		(0.401)		(0.692)
Age		0.011		0.050
		(0.019)		(0.036)
Location		0.073		-1.542*
		(0.403)		(0.759)
Constant	4.889***	4.636***	2.641***	1.351
	(0.363)	(0.613)	(0.466)	(0.994)
Observations	1,819	1,677	1,690	1,668
R-squared	0.122	0.128	0.182	0.205

Robust standard errors in parentheses
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

3.2.2.3 Social preference or a desire to take from the rich?

We hypothesised that the poor (rich) would be rewarded (punished) in the *revealed* condition because of their high (low) relative contributions. Next, we were interested in whether reward or punishment behaviour in the *revealed* condition was motivated by more than just eliciting higher relative contributions, such as potentially a desire to target and reduce the income of the rich.

To evaluate this prediction, we tested whether both relative contribution and income quintile both predicted punishment and reward received in the *revealed* condition. Indeed, holding constant the fraction of income contributed, richer participants were rewarded less (coeff = -0.309, p = 0.031, Table S13 col. 1) and punished more (coeff = 0.563, p < 0.001, Table S14 col. 1), indicating that participants were not only concerned with the higher income participant's relative contribution but generally more willing to take units from, and less willing to give units to, the rich.

Results were qualitatively similar when demographics are included. Holding constant relative contribution, higher quintiles were rewarded marginally less (coeff = -0.298, p = 0.052, Table S13 col. 2) and punished more (coeff = 0.604, $p < 0.001$, Table S14 col. 2). Similar results are obtained with log-transformed income as IV (Table S15 and S16).

While our results shed light on the combined effect of relative contribution and income quintile, it remains an important question for future research to explore to what extent income rank alone motivates the targeting of sanctions.

Table S13: Linear regression model estimating the effect of relative contribution and quintile on received reward in the *revealed* condition. Standard errors clustered on group.

VARIABLES	(1) Reward received	(2) Reward received
Relative contribution	5.209***	5.318***
	(1.038)	(1.031)
Quintile	-0.309*	-0.298
	(0.138)	(0.148)
1=Female		-0.016
		(0.695)
Age		0.049
		(0.036)
Location		-1.549*
		(0.757)
Constant	3.426***	2.146
	(0.727)	(1.102)
Observations	1,690	1,668
R-squared	0.184	0.207

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S14: Linear regression model estimating the effect of relative contribution and income quintile on received punishment in *revealed*. Standard errors clustered on group.

VARIABLES	(1) Punishment received	(2) Punishment received
Relative contribution	-3.224***	-3.212***
	(0.377)	(0.394)
Quintile	0.563***	0.604***
	(0.096)	(0.103)
1=Female		-0.181
		(0.402)
Age		0.010
		(0.019)
Location		0.133
		(0.414)
Constant	3.501***	3.107***
	(0.455)	(0.593)
Observations	1,819	1,677
R-squared	0.147	0.156

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S15: Linear regression model estimating the effect of relative contribution and income quintile on received reward in *revealed*. Standard errors clustered on group.

VARIABLES	(1) Reward received	(2) Reward received
Relative contribution	5.239***	5.346***
	(1.042)	(1.038)
Log(income)	-0.441	-0.428
	(0.234)	(0.253)
1=Female		-0.011
		(0.697)
Age		0.050
		(0.036)
Location		-1.547*
		(0.757)
Constant	3.639***	2.350
	(0.907)	(1.193)
Observations	1,690	1,668
R-squared	0.182	0.205

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S16: Linear regression model estimating the effect of relative contribution and income quintile on received punishment in *revealed*. Standard errors clustered on group.

VARIABLES	(1) Punishment received	(2) Punishment received
Relative contribution	-3.162***	-3.136***
	(0.380)	(0.393)
Log(income)	0.983***	1.060***
	(0.162)	(0.172)
1=Female		-0.173
		(0.402)
Age		0.012
		(0.019)
Location		0.117
		(0.408)
Constant	2.577***	2.045**
	(0.553)	(0.620)
Observations	1,819	1,677
R-squared	0.152	0.162

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.2.3 Public good provisioning and inequality

We next explored the effect of revealed and hidden incomes on public good provisioning as well as subsequent inequality. We find that revealing incomes had a positive effect on total contributions to the public good that was provided – and from whom these contributions came.

3.2.3.1 Revealing incomes increased contributions

We assessed the effect of revealing incomes on total contributions and whether certain players in the income distribution were affected more than others.

Overall, contributions were higher in the *revealed* than in the *hidden* condition (coeff = 1.745, p = 0.002, Table S17 col. 1). Examining how income and condition interact, we saw that in the *hidden* condition, higher income participants contributed more than lower income participants (coeff = 3.172, p < 0.001, Table S17 col. 2); and that this difference became significantly larger when incomes were revealed (interaction between income and *revealed* dummy, coeff = 1.562, p < 0.001, Table S17 col. 2).

We found qualitatively equivalent results when demographics (Table S17 col. 3) are included as well as when log-transformed income is used as the independent variable (Table S18).

Table S17: Linear regression model estimating the effect of income visibility (*revealed* dummy) and income on average contribution to the public good. Standard errors clustered on group.

VARIABLES	(1) Contribution	(2) Contribution	(3) Contribution
1=Revealed	1.745**	-2.930***	-2.651**
	(0.562)	(0.793)	(0.817)
Quintile		3.172***	3.190***
		(0.256)	(0.257)
Quintile X Revealed		1.562***	1.434**
		(0.426)	(0.437)
1=Female			-0.412
			(0.533)
Age			0.016
			(0.026)
Location			0.272
			(0.473)
Constant	8.592***	-0.923	-1.354
	(0.353)	(0.474)	(1.009)
Observations	7,067	7,067	6,804
R-squared	0.007	0.291	0.286

Robust standard errors in parentheses
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Table S18: Linear regression model estimating the effect of income visibility (*revealed* dummy) and log-transformed income on average contribution to the public good. Standard errors clustered on group.

VARIABLES	(1) Contribution	(2) Contribution	(3) Contribution
1=Revealed	1.745**	-5.229***	-4.790**
	(0.562)	(1.515)	(1.556)
Log(income)		5.416***	5.442***
		(0.463)	(0.464)
Log(income) X Revealed		2.666***	2.462**
		(0.765)	(0.784)
1=Female			-0.362
			(0.507)
Age			0.015
			(0.025)
Location			0.307
			(0.460)
Constant	8.592***	-5.598***	-6.043***
	(0.353)	(0.915)	(1.270)
Observations	7,067	7,067	6,804
R-squared	0.007	0.316	0.310

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.2.3.2 Revealing incomes reduced inequality

Finally, we assessed the effect of revealing incomes on the level of inequality and the distribution of participant payoffs at the end of the game relative to when incomes were hidden. We computed the Gini index—a commonly used measure of inequality—of the final payoffs of each group. The Gini index is defined as (Allison, 1978):

$$G = \frac{1}{2\mu n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j| \tag{1}$$

where n is the number of players with mean income μ over incomes x. The Gini takes a value between 0 and 1: the higher the value, the more unequal the set of incomes.

We found that the Gini index at the end of the game is lower in the *revealed* condition (average 0.169) than in the *hidden* condition (average 0.238; rank-sum, $p < 0.001$). Thus, revealing incomes decreased inequality relative to keeping incomes hidden.

To explain this difference in inequality, we examined contributions over time. To account for multiple testing in these regressions, we report Bonferroni-corrected *p*-values. Participants in quintiles 1 (poorest) through 4 never decreased their contributions in either the *hidden* or *revealed* conditions. In fact, the poorest player in the *hidden* condition actually increased their contributions over time (coeff = 0.047, p = 0.04 bonferroni-corrected; all other bottom-to-4th quintiles in both conditions: *p*s = 1.000 corrected; Tables S19 and S20 cols. 1-4; all regressions were also robust to including demographics).

The contributions of the top quintile participants did, however, differ substantially over time between the *hidden* and *revealed* conditions. In the *revealed* condition, rich participants maintained their contributions over time (coeff = -0.382, p = 1.000 corrected, Table S20 col. 5), whereas they decreased their contributions over time in the *hidden* condition (coeff = -1.077, p < 0.001 corrected, Table S19 col. 5). In other words, when incomes were revealed, sanctions were effective in maintaining cooperation from all players, including those with the greatest ability to contribute.

Table S19: Linear regression model estimating the effect of round on contribution to the public good in the *hidden* condition. Standard errors clustered on group.

VARIABLES	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
Round	0.047*	0.087	0.092	-0.203	-1.077***
	(0.017)	(0.043)	(0.056)	(0.092)	(0.240)
Constant	2.620***	5.387***	7.549***	10.173***	22.829***
	(0.137)	(0.318)	(0.482)	(0.736)	(1.781)
Observations	683	683	681	682	683
R-squared	0.008	0.006	0.003	0.009	0.039

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05 [Bonferroni corrected]

Table S20: Linear regression model estimating the effect of round on contribution to the public good in the *revealed* condition. Standard errors clustered on group.

VARIABLES	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
Round	0.024	0.037	0.032	0.024	-0.382
	(0.021)	(0.061)	(0.073)	(0.103)	(0.266)
Constant	2.663***	5.700***	7.721***	11.385***	25.626***
	(0.157)	(0.369)	(0.434)	(0.587)	(1.723)
Observations	733	733	729	730	730
R-squared	0.002	0.001	0.000	0.000	0.003

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05 [Bonferroni corrected]

3.3 Experiment 3

Following the procedures of the second experiment, we repeat the same statistical analysis for experiment 3. Unless otherwise noted, all statistics are linear regressions with income quintile as a continuous independent variable and standard errors are clustered on the group. To check the robustness of our results, we also use log-transformed absolute income as continuous IV.

3.3.1 Individual performance

In experiment 3, participants first completed an individual effort task before they were assigned to groups and received an income level based on their performance in the individual task. Their rank among their group members determined which income level they were assigned to: the best-performing participant in the individual task was allocated the highest income level; the second-best performing participant was assigned the second-highest income level; and so on.

In the individual task, participants had to count the number of 0s in a random matrix of 0s and 1s (Abeler et al., 2011). The more matrices they solved correctly, the higher their performance score. There was no difference in the mean number of correctly solved matrices between the *hidden* (mean = 4.489, s.d. = 1.869) and *revealed* (mean = 4.544, s.d. = 2.029) conditions; $t(438) = -0.298$, $p = 0.766$. There were no significant differences in performance for any quintile between conditions (Table S21).

Table S21: The number of correctly solved matrices did not differ between conditions for any quintile. Mean values on top, standard deviation in parentheses.

	Hidden condition	*Revealed* condition	Two-tailed t-test
Top quintile	6.578 (1.530)	6.953 (1.252)	$p = 0.212$
2nd highest quintile	5.267 (1.074)	5.627 (0.976)	$p = 0.103$
Middle quintile	4.589 (1.125)	4.674 (1.063)	$p = 0.951$
2nd lowest quintile	3.667 (1.000)	3.441 (1.120)	$p = 0.323$
Bottom quintile	2.244 (1.026)	2.023 (1.080)	$p = 0.327$

3.3.2 Received reward and punishment

In our second experiment, reward and punishment are available in all conditions. Thus, the dependent variable is the number of units that a participant received: negative units represent punishment received, while positive units represent reward received. The independent variable is the income quintile of the recipient of the sanctions (1 to 5). We found qualitatively similar results when we used log-transformed income of the recipient as the independent variable.

In the *hidden* condition, participants could not assess to what extent another player can contribute. Thus, we expected participants to punish the poor for their (low total) contributions and to reward the rich for their (high total) contributions. In the *revealed* condition, conversely, we predicted the mirror image: poorer participants would be rewarded more units for their (high percentage) contribution than wealthier ones.

We estimated the number of units that a participant received as a function of their income quintile and whether the income distribution was hidden or revealed (Table S22). In the *hidden* condition, we found that higher income participants received more units (coeff = 0.053, p = 0.042, Table S22 col. 1), whereas in the *revealed* condition, higher income players received fewer units (coeff = -0.171, $p < 0.001$, Table S22 col. 2). Furthermore, a regression including data from both *hidden* and *revealed* conditions together showed that this difference was itself significant (interaction between income and *revealed* dummy, coeff = -0.225, $p < 0.001$, Table S22 col. 3).

We found qualitatively similar results with log-transformed income as independent variable: lower income participants received more units in *hidden* (coeff = 0.188, p = 0.037, Table S23 col. 1) but, conversely, they received fewer units in *revealed* (coeff = -0.582, $p < 0.001$, Table S23 col. 2), and the interaction between condition and income was significant (interaction between log-transformed income and *revealed* dummy, coeff = -0.771, $p < 0.001$, Table S23 col. 3).

Table S22: Linear regression model estimating the effect of a target's income quintile (i.e., their position in the income distribution) on the number of units they received. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Revealed*	(3) Interaction
Quintile	0.054*	-0.171***	0.054*
	(0.026)	(0.030)	(0.025)
1=Revealed			0.949***
			(0.193)
Quintile X Revealed			-0.225***
			(0.039)
Constant	-0.054	0.895***	-0.054
	(0.121)	(0.151)	(0.120)
Observations	1,970	1,935	3,905
R-squared	0.006	0.044	0.043

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S23: Linear regression model estimating the effect of a target's log-transformed income on the number of units they received. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Revealed*	(3) Interaction
Log(income)	0.188*	-0.582***	0.188*
	(0.087)	(0.097)	(0.087)
1=Revealed			1.143***
			(0.217)
Log(income) X Revealed			-0.771***
			(0.130)
Constant	-0.105	1.038***	-0.105
	(0.139)	(0.169)	(0.138)
Observations	1,970	1,935	3,905
R-squared	0.006	0.044	0.043

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.3.3 Absolute vs. relative contribution

Participants in the *hidden* condition could not take another player's ability to contribute into account since they did not know the income distribution. In the *revealed* condition, on the other hand, participants could evaluate the amount contributed relative to the player's income before choosing whom to punish or reward – in other words, participants could differentiate between absolute and relative contributions.

We hypothesised that absolute contributions would predict the number of units received when income is hidden, but that relative contributions would predict units received when income is revealed. Thus we expected participants to take (give) more units those who give little (a lot) in absolute terms in *hidden*, while their sanctions would be driven by relative contributions in *revealed*.

3.3.3.1 Absolute vs. relative contribution by quintile

As predicted, higher income participants contribute more in absolute terms in both the *hidden* (coeff = 3.465, $p < 0.001$, Table S24 col. 1) and *revealed* (coeff = 5.573, $p < 0.001$, Table S24 col. 2) conditions. Conversely, we found that higher income participants contributed a smaller percentage of their income in the *hidden* condition (coeff = -0.064, $p < 0.001$, Table S25 col. 1).

Surprisingly, in the *revealed* condition, there was only a weak trend of higher income participants contributing a lower percentage of their income (coeff = -0.0263, $p = 0.108$, Table S25 col. 2). This is a slight departure from our previous results in experiment 1: while higher income participants in both experiments contributed more after being punished more and rewarded less in the *revealed* condition, it appears that sanctions were more effective in the *revealed* condition in experiment 2 to encourage richer participants to contribute a higher fraction of their income.

These results were qualitatively similar when log-transformed income is used as the independent variable (Tables S26 and S27).

Table S24: Linear regression model estimating the effect of income on absolute contribution. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Revealed*
Quintile	3.465***	5.573***
	(0.346)	(0.535)
Constant	-2.446**	-5.640***
	(0.698)	(0.921)
Observations	1,581	1,598
R-squared	0.236	0.381

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S25: Linear regression model estimating the effect of income on percentage of income contributed (relative contribution). Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Revealed*
Quintile	-0.064***	-0.026
	(0.011)	(0.016)
Constant	0.664***	0.671***
	(0.050)	(0.054)
Observations	1,581	1,598
R-squared	0.055	0.009

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S26: Linear regression model estimating the effect of log-transformed income on absolute contribution. Standard errors clustered on group.

	(1)	(2)
VARIABLES	*Hidden*	*Revealed*
Log(income)	11.742***	18.733***
	(1.151)	(1.761)
Constant	-5.264***	-9.996***
	(0.937)	(1.288)
Observations	1,581	1,598
R-squared	0.235	0.373

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S27: Linear regression model estimating the effect of log-transformed income on percentage of income contributed. Standard errors clustered on group.

	(1)	(2)
VARIABLES	*Hidden*	*Revealed*
Log(income)	-0.210***	-0.084
	(0.037)	(0.054)
Constant	0.710***	0.687***
	(0.058)	(0.065)
Observations	1,581	1,598
R-squared	0.052	0.008

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.3.3.2 Absolute vs. relative contribution predicted sanctioning behaviour

We hypothesised that participants in the *hidden* condition would reward those who give more in absolute terms (the rich) and punish those who give less in absolute terms (the poor). Conversely, participants in the *revealed* condition would reward those who give a high percentage of their income (poorer participants) but punish those who give a smaller percentage of their income (richer participants).

As predicted, higher absolute contributions in the *hidden* condition led to receiving more reward units and fewer punishment units (coeff = 0.571, $p < 0.001$, Table S28 col. 1). This helped mostly richer participants because they contributed a larger absolute number of units. Relative contributions, in contrast, did not predict the number of units received ($p = 0.098$, Table S28 col. 1) in the *hidden* condition since relative contributions were not observable.

In the *revealed* condition, we found that higher relative contribution led to receiving more units (coeff = -1.775, $p < 0.001$, Table S28 col. 2). Since poorer participants contributed a larger percentage of their income, they were punished less and rewarded more in the *revealed* condition. We also observed that absolute contribution had an effect in the *revealed* condition, in the opposite direction of the effect in the *hidden* condition: higher absolute contributions led to fewer units received (coeff = -0.461, $p = 0.001$, Table S28 col. 2).

Table S28: Linear regression model estimating the effect of absolute log-transformed contribution and relative contribution on the number of units received in the *hidden* and *revealed* condition. To deal with zero-contributions, a constant of 1 was added to all contributions before applying the log-transformation. Standard errors clustered on group.

VARIABLES	(1) *Hidden*	(2) *Revealed*
Log(contribution+1)	0.571***	-0.461***
	(0.109)	(0.123)
Relative contribution	0.195	1.775***
	(0.115)	(0.212)
Constant	-0.365**	-0.211*
	(0.117)	(0.097)
Observations	1,581	1,598
R-squared	0.084	0.206

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.3.3.3 Social preference or a desire to take from the rich?

We hypothesised that the poor (rich) would be rewarded (punished) in the *revealed* condition because of their high (low) relative contributions, for which we found evidence presented above. Next we investigated to what extent participants were motivated by more than just relative contributions, such as a desire to reduce the income of the rich regardless of their relative contribution.

Holding constant the fraction of income contributed, we found that richer participants indeed received fewer units (coeff = -0.173, $p < 0.001$, Table S29 col. 1). Results were qualitatively similar when we used log-transformed income as the independent variable (coeff = -0.595, $p < 0.001$, Table S29 col. 2). Thus, participants gave fewer units to the rich, even when they contributed the same relative amount of their income.

Table S29: Linear regression model estimating the effect of relative contribution and income quintile on units received in the *revealed* condition. Standard errors clustered on group.

VARIABLES	(1) Units received	(2) Units received
Relative contribution	1.291***	1.294***
	(0.173)	(0.172)
Quintile	-0.173***	
	(0.032)	
Log(income)		-0.595***
		(0.109)
Constant	0.217	0.367*
	(0.143)	(0.161)
Observations	1,598	1,598
R-squared	0.225	0.226

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.3.4 Public good provisioning and inequality

3.3.4.1 Revealing incomes increased contributions

Contributions were higher in the *revealed* than in the *hidden* condition (coeff = 3.134, $p < 0.001$, Table S30 col. 1). We observed that higher income participants in the *hidden* condition contributed more than lower income participants (coeff = 3.194, $p = 0.007$, Table S30 col. 2); a difference that became significantly larger when incomes were revealed (interaction between income and *revealed* dummy, coeff = 2.109, $p = 0.001$, Table S30 col. 2). We found qualitatively equivalent results when using log-transformed income as the independent variable (Table S31).

Table S30: Linear regression model estimating the effect of income visibility (*revealed* dummy) and income on average contribution to the public good. Standard errors clustered on group.

VARIABLES	(1) Contribution	(2) Contribution
1=Revealed	3.134***	-3.194**
	(0.877)	(1.149)
Quintile		3.465***
		(0.344)
Quintile X Revealed		2.109**
		(0.633)
Constant	7.946***	-2.446***
	(0.467)	(0.695)
Observations	3,179	3,179
R-squared	0.018	0.338

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S31: Linear regression model estimating the effect of income visibility (*revealed* dummy) and log-transformed income on average contribution to the public good. Standard errors clustered on group.

	(1)	(2)
VARIABLES	Contribution	Contribution
1=Revealed	3.134***	-4.732**
	(0.877)	(1.583)
Log(income)		11.742***
		(1.145)
Log(income) X Revealed		6.992**
		(2.092)
Constant	7.946***	-5.264***
	(0.467)	(0.932)
Observations	3,179	3,179
R-squared	0.018	0.333

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

3.3.4.2 Revealing incomes reduced inequality

Finally, we assessed the effect that revealing incomes had on inequality. Using (1), we computed the Gini index of the final payoffs of each group. We found that the Gini index at the end of the game was lower in the *revealed* condition (average 0.124) than in the *hidden* condition (average 0.255; Rank-sum, $p < 0.001$). Revealing incomes decreased inequality relative to keeping incomes hidden.

What led to lower inequality in the *revealed* condition? We examined contributions over time. To account for multiple testing in these regressions, we report Bonferroni-corrected *p*-values. Participants in quintiles 1 (poorest) through 4 never decreased their contributions in either the *hidden* or *revealed* conditions (all *p*s > 0.5 corrected; Tables S32 and S33 cols. 1-4).

Contributions of the highest earners, however, did marginally differ over time between the *hidden* and *revealed* conditions. In the *revealed* condition, rich participants maintained their contributions over time (coeff = -0.056, $p = 1.000$ corrected, Table S32 col. 5), whereas they marginally decreased their contributions over time in the *hidden* condition (coeff = -1.105, $p = 0.080$ corrected, Table S33 col. 5).

Table S32: Linear regression model estimating the effect of round on contribution to the public good in the *hidden* condition. Standard errors clustered on group.

VARIABLES	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
Round	0.029	0.075	0.069	-0.240	-1.105
	(0.021)	(0.067)	(0.203)	(0.190)	(0.442)
Total # rounds	0.036	-0.450*	-0.831***	-0.600	0.957
	(0.105)	(0.144)	(0.211)	(0.302)	(1.591)
Constant	1.287	7.873***	15.421***	15.727***	13.124
	(0.940)	(1.330)	(2.016)	(3.120)	(13.069)
Observations	316	317	316	316	316
R-squared	0.010	0.051	0.061	0.036	0.033

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05 [Bonferroni corrected]

Table S33: Linear regression model estimating the effect of round on contribution to the public good in the *revealed* condition. Standard errors clustered on group.

VARIABLES	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
Round	0.025	0.112	0.347	0.192	0.056
	(0.034)	(0.088)	(0.147)	(0.219)	(0.476)
Total # rounds	-0.032	-0.143	-0.360	-0.688	-0.714
	(0.142)	(0.254)	(0.665)	(0.813)	(1.159)
Constant	2.046	5.444*	11.876*	18.442*	31.846**
	(1.266)	(2.354)	(5.782)	(7.564)	(11.144)
Observations	320	320	318	320	320
R-squared	0.003	0.011	0.030	0.012	0.003

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05 [Bonferroni corrected]

3.4 Experiment 4

The analysis of experiment 4 was similar to that of our first experiment: in this study, participants had to choose which school parents would be the target of a “tax increase” – our measure of punishment in this study. We repeated the same analysis as in experiment 1, first looking at the average and median choice of target, followed by a look at the distribution to identify which parents (sorted by their income) were targeted specifically.

We predicted that revealing incomes would lead participants to punish richer households more.

In line with this prediction, we found that participants' choice of the target of a tax increase shifted from the poorest school parents to richer ones when incomes were revealed (using linear regression predicting the choice of target by *revealed* dummy: coeff = 1.826, $p < 0.001$, Table S34 col. 1). Results were qualitatively similar when control variables were included (Table S34 col. 2) and when estimated with a rank-sum test instead ($Z = -11.883$, $p < 0.001$).

Finally, we investigated which school parents in particular are chosen to pay the additional tax between the *revealed* and *hidden* conditions. We observed that participants were significantly less likely to choose the poorest school parents for the tax payment when incomes were revealed (using logit regression predicting choosing the poorest school parents by *revealed* dummy: coeff = -2.222, $p < 0.001$, Table S35 col. 1) while they were significantly more likely to choose the richest school parents (coeff = 4.302, $p < 0.001$, Table S35 col. 5). Moreover, when incomes were revealed, parents in the second-richest district were also less likely to be chosen to pay the additional tax (coeff = -1.985, $p < 0.001$, Table S35 col. 4). School parents from other incomes were not chosen significantly differently between the two conditions (ps > 0.05, Table S35 cols. 2-3).

Table S34: Linear regression estimating the choice of school parents (where 1 = poorest school parents to 5 = richest school parents) who would be "punished" with an additional tax to pay for more educational funding across schools.

	(1)	(2)
VARIABLES	Target school	Target school
1=Revealed	1.826***	1.810***
	(0.149)	(0.153)
1=Female		0.093
		(0.155)
Age		-0.012
		(0.007)
Constant	2.648***	3.046***
	(0.117)	(0.290)
Observations	313	304
R-squared	0.323	0.328

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S35: Logit regression estimating the likelihood that the poorest (col. 1), the second poorest (col. 2), the middle (col. 3), the second richest (col. 4) or the richest school parents (col. 5) were chosen for the additional tax payment.

VARIABLES	(1) Poorest	(2) 2nd Poorest	(3) Middle	(4) 2nd Richest	(5) Richest
1.Revealed	-2.222***	-0.491	-0.707	-1.985***	4.302***
	(0.352)	(0.740)	(0.478)	(0.324)	(0.432)
Constant	-0.343*	-3.428***	-2.338***	-0.317*	-3.078***
	(0.161)	(0.455)	(0.280)	(0.161)	(0.387)
Observations	313	313	313	313	313

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

4. Instructions

In this section, we include the main instructions from each of the experiments. All experiments were run on Amazon Mechanical Turk. Experiments 1 and 4 were built with Qualtrics and Experiments 2 and 3 were programmed on SoPHIE, an interactive platform for decision-making experiments (Hauser, Hendriks, Rand, & Nowak, 2016). To obtain all screenshots and instructions, please email the authors.

4.1 Experiment 1

4.1.1 Hidden condition

We asked a sample of MTurkers how much they gave to charity last year. Below are the responses of five MTurkers.

You will now choose to give a $1.00 bonus to one of these MTurkers. We will pay this MTurker the $1.00 bonus; the other MTurkers will receive no bonus.

(Please note that your choice of who gets the bonus will be implement if your decision is drawn at random at the end of the study.)

Which MTurker do you want to give a $1.00 bonus?

MTurkers:	Charitable contributions last year	Your decision:
MTurker "RW"	$2,594	This MTurker should receive the $1.00 bonus
MTurker "PV"	$1,874	This MTurker should receive the $1.00 bonus
MTurker "QM"	$4,130	This MTurker should receive the $1.00 bonus
MTurker "NG"	$2,970	This MTurker should receive the $1.00 bonus
MTurker "HT"	$3,356	This MTurker should receive the $1.00 bonus

4.1.2 Revealed condition

We asked a sample of MTurkers how much they gave to charity last year. Below are the responses of five MTurkers.

You will now choose to give a $1.00 bonus to one of these MTurkers. We will pay this MTurker the $1.00 bonus; the other MTurkers will receive no bonus.

(Please note that your choice of who gets the bonus will be implement if your decision is drawn at random at the end of the study.)

Which MTurker do you want to give a $1.00 bonus?

MTurkers:	Charitable contributions last year	Income range last year	Your decision:
MTurker "HT"	$4,130	$100,000–$200,000	This MTurker should receive the $1.00 bonus
MTurker "PV"	$2,594	$25,000–$50,000	This MTurker should receive the $1.00 bonus
MTurker "QM"	$1,874	Under $25,000	This MTurker should receive the $1.00 bonus
MTurker "NG"	$3,356	$75,000–$100,000	This MTurker should receive the $1.00 bonus
MTurker "RW"	$2,970	$50,000–$75,000	This MTurker should receive the $1.00 bonus

4.2 Experiment 2

4.2.1 Hidden condition

Detailed Instructions:

At the beginning of every round:

Each person in your group receives an income (in units) per round. In total, there are 100 units across all players per round. On the next pages you will be told your income in each round. You receive the same income every round.

Note that you and every participant in your group will only be told your own income and not the income of others. In other words, you are the only one who knows your income.

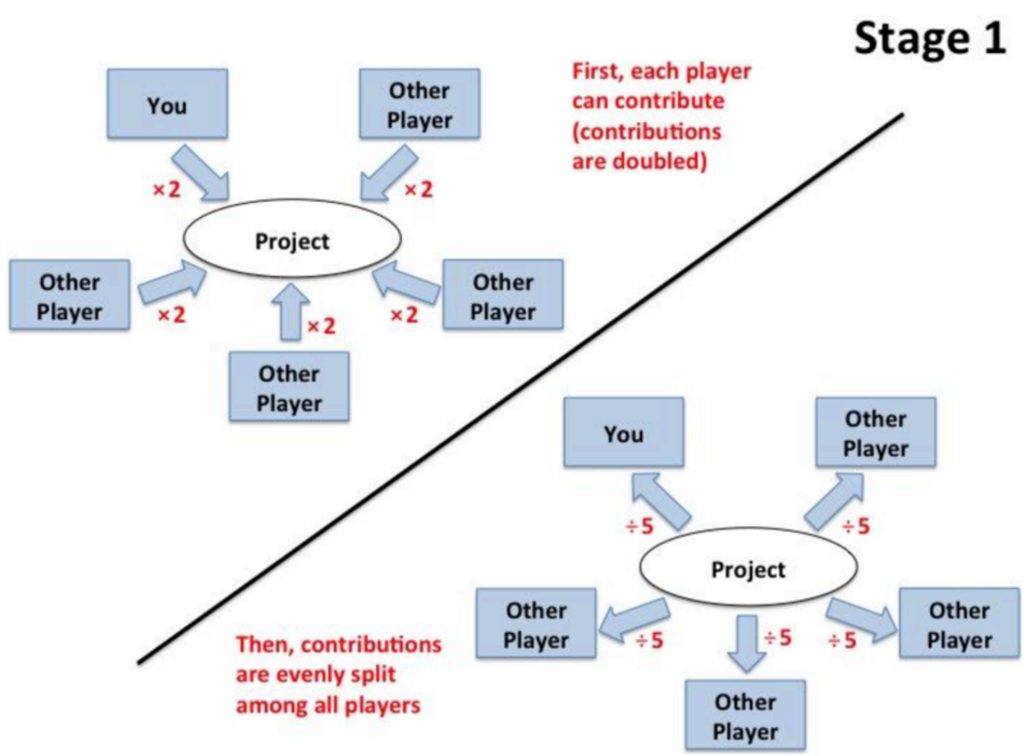
Furthermore, at the beginning of every round, you and all other players receive a new player name. Player names are made up of two alphabetical letters (e.g., AE, UQ, VG, ...) and are randomly assigned at the beginning of every round. **However, despite the change of player names each round, every person keeps their same income.**

Stage 1:

You have to decide how many of the units you receive you want to contribute to the common project and how many of them to keep for yourself.

All units that are contributed to the common project are **doubled**. The doubled contributions are then **split evenly among all participants**, regardless of whether they contributed.

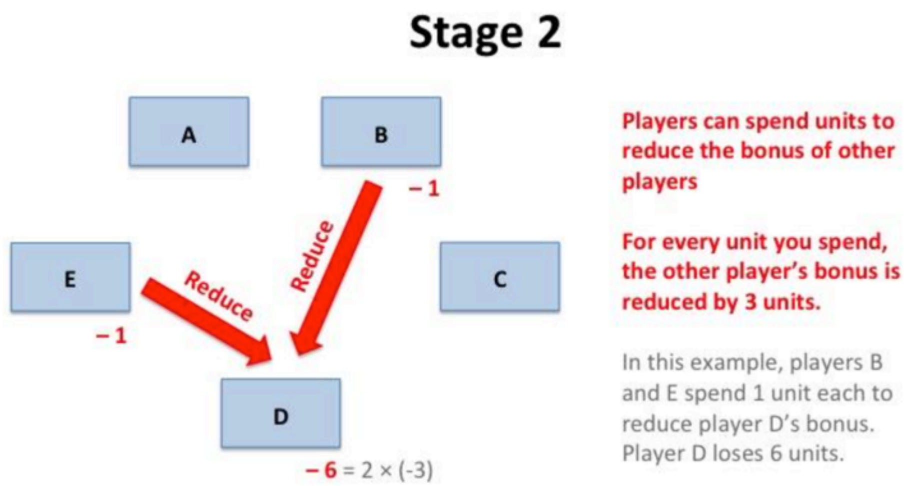
This picture summarises Stage 1:



Stage 2:

You will be given the option to pay to reduce someone else's payoff. For **every unit** that you spend, the **other player loses 3 units**. You cannot spend more than the number of units that you've earned in Stage 1.

This picture summarises Stage 2:



In summary:

Your bonus in each round equals the units you keep in Stage 1 plus the units you receive from the common project minus the units that you spend on reducing others' payoff and that others have taken from you in Stage 2.

The exchange rate is 100 units = $0.40.

There will be **many rounds** like the one described above.

Please make sure that you understand the instructions completely, as you will be asked to answer a few comprehension questions. You **cannot return** to this page later.

Public goods game (repeated 10 times):

In this round, your player name is "MQ".

The other players have the following names in this round:

Player **"EM"**

Player **"QQ"**

Player **"ZX"**

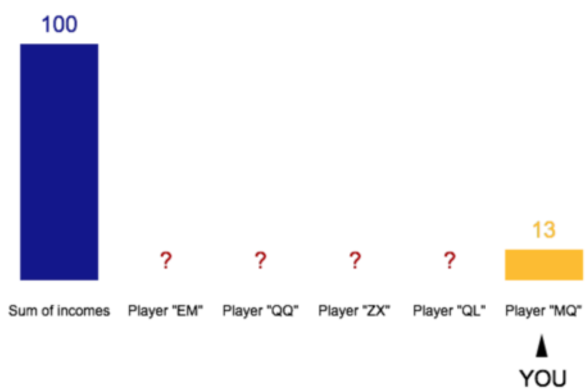
Player **"QL"**

Continue ...

There are always **100 units in total** in your group and you always receive the **same income every round (13 units)**.

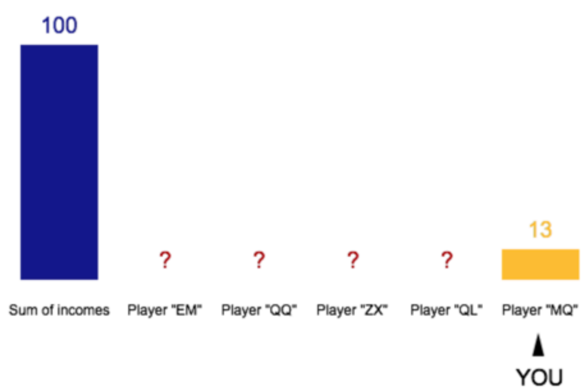
However, your **player name** and the player name of other participants is **randomly generated and assigned each round**.

In this round you are player "MQ":



THIS IS STAGE 1.

You are the **only one** who knows your own income:



You have **13 units available** in this stage.

How many units do you want to contribute to the common project?

Continue

You can type any integer between 0 and 13 units.

In Stage 1, a total of **63 units were contributed**.

Everyone received an **equal share of 26 units**.

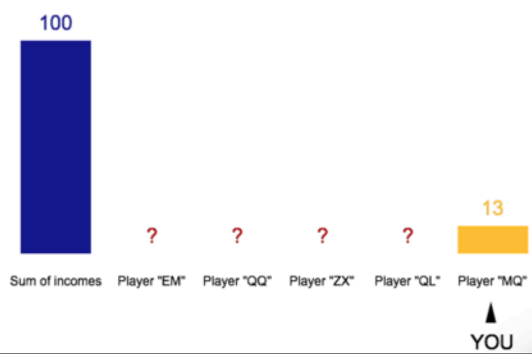
You kept 1 units for yourself.

Thus **you earned 27 units** in Stage 1.

THIS IS STAGE 2.

In Stage 2, you have the option to **pay to reduce someone else's payoff.**

You are the **only one** who knows your own income:



Player	**Contribution in Stage 1**	**Options in Stage 2**
Player "EM"	7	A B C
Player "QQ"	15	A B C
Player "ZX"	4	A B C
Player "QL"	25	A B C
Player "MQ" **(This is YOU!)**	12	

Please choose from the following options:

	You	**Other**
Option **A**	0	0
Option **B**	-1	-3
Option **C**	-2	-6

Continue

STAGE 2 SUMMARY

Units you spent to reduce others' payoffs:	**-3**

Other people's actions towards you:

Player	**Action**	**You got**
Player "EM"	A	0
Player "QQ"	B	-3
Player "ZX"	A	0
Player "QL"	B	-3

Your payoff in Stage 2:	**-9**

YOUR PAYOFF IN THIS ROUND:

Stage 1	27
Stage 2	-9
Total	**18**

4.2.2 Revealed condition

Detailed Instructions:

At the beginning of every round:

Each person in your group receives an income (in units) per round. In total, there are 100 units across all players per round. On the next pages you will be told your income in each round. You receive the same income every round.

Note that you and every participant in your group will be told the income of everybody else.

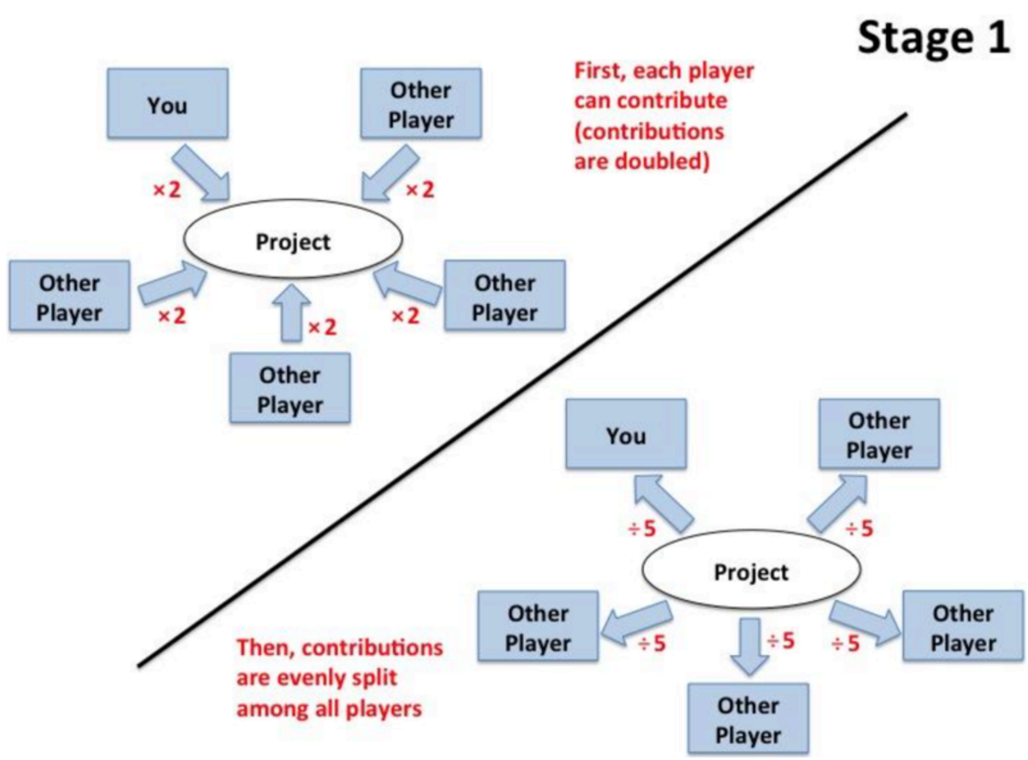
Furthermore, at the beginning of every round, you and all other players receive a new player name. Player names are made up of two alphabetical letters (e.g., AE, UQ, VG, ...) and are randomly assigned at the beginning of every round. **However, despite the change of player names each round, every person keeps their same income.**

Stage 1:

You have to decide how many of the units you receive you want to contribute to the common project and how many of them to keep for yourself.

All units that are contributed to the common project are **doubled**. The doubled contributions are then **split evenly among all participants**, regardless of whether they contributed.

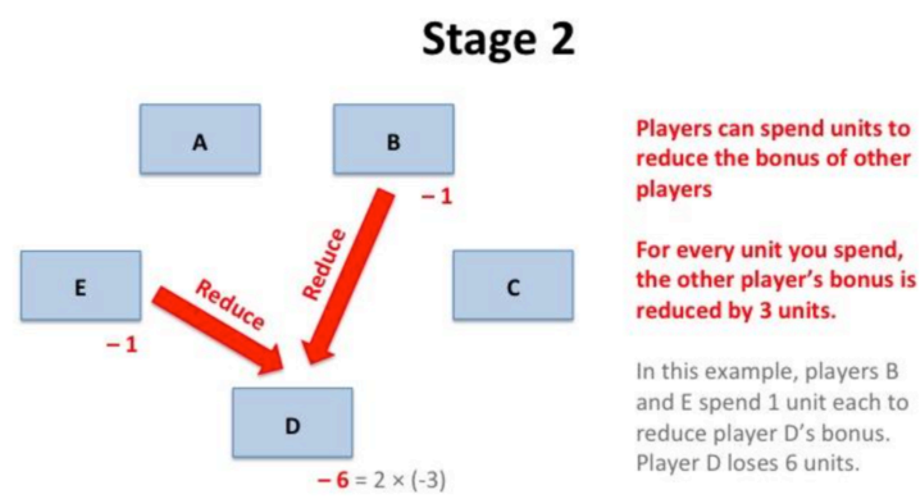
This picture summarises Stage 1:



Stage 2:

You will be given the option to pay to reduce someone else's payoff. For **every unit** that you spend, the **other player loses 3 units**. You cannot spend more than the number of units that you've earned in Stage 1.

This picture summarises Stage 2:



In summary:

Your bonus in each round equals the units you keep in Stage 1 plus the units you receive from the common project minus the units that you spend on reducing others' payoff and that others have taken from you in Stage 2.

The exchange rate is 100 units = $0.40.

There will be **many rounds** like the one described above.

Please make sure that you understand the instructions completely, as you will be asked to answer a few comprehension questions. You **cannot return** to this page later.

Public goods game (repeated 10 times):

In this round, your player name is "XN".

The other players have the following names in this round:

Player **"KU"**

Player **"IP"**

Player **"SI"**

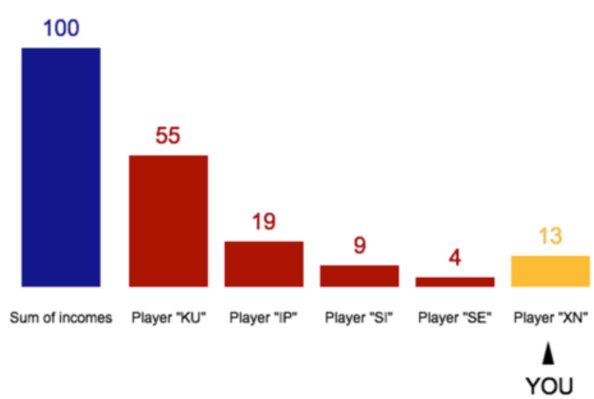
Player **"SE"**

Continue ...

There are always **100 units in total** in your group and you always receive the **same income every round (13 units)**.

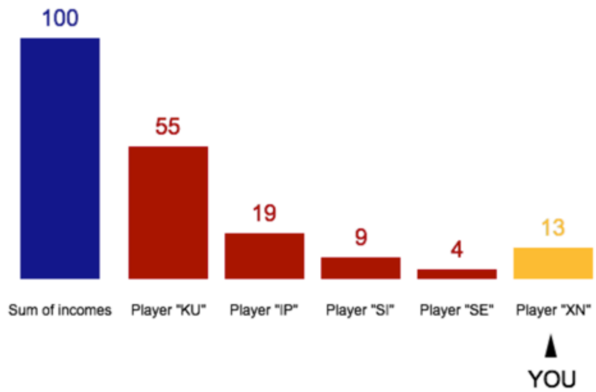
However, your **player name** and the player name of other participants is **randomly generated and assigned each round**.

In this round you are player "XN":



THIS IS STAGE 1.

Here is the **income distribution** in your group:



You have **13 units available** in this stage.

How many units do you want to contribute to the common project?

Continue

You can type any integer between 0 and 13 units.

In Stage 1, a total of **65 units were contributed**.

Everyone received an **equal share of 26 units**.

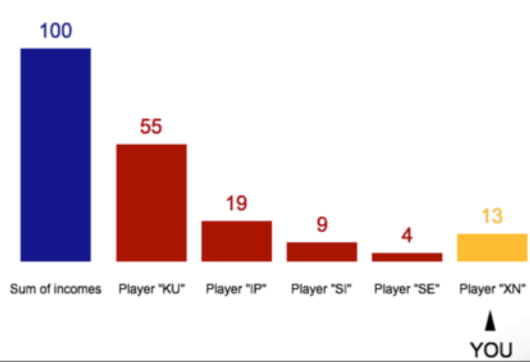
You kept 1 units for yourself.

Thus **you earned 27 units** in Stage 1.

THIS IS STAGE 2.

In Stage 2, you have the option to **pay to reduce someone else's payoff.**

Here is the **income distribution** in your group:



Player	Contribution in Stage 1	Options in Stage 2
Player "KU"	25	○ A ○ B ○ C
Player "IP"	17	○ A ○ B ○ C
Player "SI"	7	○ A ○ B ○ C
Player "SE"	4	○ A ○ B ○ C
Player "XN" **(This is YOU!)**	12	

Please choose from the following options:

	You	Other
Option **A**	0	0
Option **B**	-1	-3
Option **C**	-2	-6

Continue

4.3 Experiment 3

4.3.1 Hidden condition

INDIVIDUAL TASK

Task: Count the number of 0s and submit the number below. Solve <u>as many tables</u> of 0s and 1s as you can in 2 minutes.

The **number of correctly solved tables** (your performance) in the individual task **affects your income** in the upcoming group task. In the group task, you will be randomly matched with 4 other workers and each of you will receive a different income. (You will interact with the same 4 workers for all rounds of the group task.)

Your performance in this individual task relative to your group members' performance will **determine your income level** in the group task. The highest income level will be assigned to the worker in your group who performs the best in the individual task (correctly solving the most tables of 0s and 1s). The second-highest income level will be assigned to the second-best performing worker, and so on. A coin flip will decide if there are performance ties.

When you are ready to get started with the task, <u>click here</u>. The first table of 0s and 1s will appear and the 2-minutes timer will start immediately.

INDIVIDUAL TASK

Task: Count the number of 0s and submit the number below. Solve as many tables of 0s and 1s as you can in 2 minutes.

1	0	1	1	1	1	1	1
1	0	1	1	0	1	1	1
0	1	0	1	0	1	1	1
0	1	0	0	1	0	1	1
1	1	1	0	1	1	1	0
1	1	0	1	1	1	1	1
1	1	0	1	0	1	1	1
1	0	1	1	1	1	1	1

Submit solution & next puzzle

1 min 54 sec

Detailed Instructions (1/3): GROUP TASK

In this group task, you are matched with 4 other workers who are taking this HIT. You will interact with the same 4 participants for the entire group task. All interactions take place in real time.

Please be considerate of others' time!

In this group task, there are multiple rounds. At the beginning of every round, each participant receives a certain number of units (income). The total number of units distributed in each round will be 100 units, though different people will receive different incomes. Once assigned their income, participants will receive that same income every round.

Also at the beginning of every round, each participant will receive a new player name. Player names are made up of 2 random alphabetical letters, for example: "AE", "UQ", "VG", etc. In sum, in every round, a participant's **income will remain the same, *but* they will have a different player name**.

There will be at least 8 rounds of the group task. After the 8th round, there will be a 50% chance that another round will follow. After that round, there will again be a 50% chance that another round will follow, and so on until the group task ends. When the group task is over, you will see a brief questionnaire before you can submit the HIT.

Detailed Instructions (2/3): GROUP TASK

Players' incomes will reflect the distribution of **incomes in the United States**. We sorted the U.S. population by income and divided them into five equally large groups (quintiles). The 1st quintile reflects the top 20% of earners, while the 5th quintile reflects the bottom 20% of earners.

In this group task, we will **distribute 100 units among participants proportional** to each quintile of U.S. earners. In other words, one participant will receive the number of units (out of 100) equivalent to the proportion of income of the top 20% of earners out of all earners, and so on for each quintile.

Which income level you are assigned to is affected by your performance in the individual task you just completed. Each group member's income in the group task is determined by their **performance rank** in the previous individual task.

You will only be told your own income and not the income of others. The other group members will only be told their own income. This means that you are the only one in your group who knows your own income.

Detailed Instructions (3/3): GROUP TASK

There are 2 stages in every round of the group task.

Stage 1: You have to decide how much of your income (in units) you want to contribute to the common project and how much to keep for yourself. **All units that are contributed** to the common project are **doubled**. Once doubled, the total amount is **split evenly** among all players, regardless of whether they contributed. Thus for every 2 units you contribute, you get 1 unit back: so no matter what the other group members contribute, you personally lose units on contributing, but contributing benefits the group as a whole. **For example:**

- Imagine everyone had an income of 20 units (**note**: this is a hypothetical example; each participant in the real game will have a different amount).
- If everyone contributed all of their 20 units, everyone's units would double: each of you would earn 40 units.
- If everyone else contributed their 20 units, while you kept your 20 units, you would earn 52 units, while the others would earn only 32 units.

Stage 2: Everyone will be given the option to pay some of their income to increase or decrease someone else's payoff. **For every 2 units that you spend, the other player either gains or loses 6 units.** You cannot spend more than the number of units that you've earned in Stage 1.

Your bonus in each round is calculated as follows:

- the units you keep for yourself in Stage 1 **PLUS**
- the units you receive from the common project (the sum of everyone's contributions doubled and split five ways) in Stage 1 **PLUS**
- the units that others have given you in Stage 2 **MINUS**
- the units that others have taken from you in Stage 2 **MINUS**
- the units that you spend on increasing or decreasing others' payoff in Stage 2.

The exchange rate is 200 units = $1.00. If your earnings are below 0 units at the end of the game, your bonus will be truncated at $0. You cannot earn a negative bonus.

Public goods game (repeated 10 times):

	Players	Incomes
	Player "VX"	? units
	Player "RG"	? units
	Player "XJ"	? units
YOU >>>	Player "OD"	23 units
	Player "LV"	? units
		Total income: 100 units

Your income is 23 units. Your player name in this round is "OD".

Please decide now how many units you want to contribute to the common pool.

(Type a value between 0 and 23.)

Stage 1 is completed. The payoff you earned in Stage 1 is 26.2 units.

STAGE 2: Please choose an action towards the other players:

- A : You pay 2 units to **decrease** the other player's payoff by 6 units.
- B : You pay 0 units. The other player's payoff remains **unchanged**.
- C : You pay 2 units to **increase** the other player's payoff by 6 units.

	Players	Incomes	Contributions	Choose action:
	Player "VX"	? units	6 units	A B C
	Player "RG"	? units	3 units	A B C
	Player "XJ"	? units	15 units	A B C
YOU >>>	Player "OD"	23 units	20 units	
	Player "LV"	? units	14 units	A B C
		Total income: 100 units	Total contributed: 58 units	

Your player name in this round is "OD".

The four other players have chosen to following actions towards you (***presented in random order***):

Option A	**-6** units
Option B	**0** units
Option C	**+6** units
Option C	**+6** units
From other players:	**6 units**

You have spent a total of **-4 units** in Stage 2 towards the other players, which is subtracted from your Stage 2 payoff.

YOUR PAYOFF IN THIS ROUND:

Stage 1: **26.2 units**

Stage 2: **2 units**

TOTAL EARNED: 28.2 UNITS

4..2 Revealed condition

The *hidden* and *revealed* conditions were identical except for the following three pages:

Detailed Instructions (2/3): GROUP TASK

Players' incomes will reflect the distribution of **incomes in the United States**. We sorted the U.S. population by income and divided them into five equally large groups (quintiles). The 1st quintile reflects the top 20% of earners, while the 5th quintile reflects the bottom 20% of earners.

In this group task, we will **distribute 100 units among participants proportional** to each quintile of U.S. earners. In other words, one participant will receive the number of units (out of 100) equivalent to the proportion of income of the top 20% of earners out of all earners, and so on for each quintile.

Which income level you are assigned to is affected by your performance in the individual task you just completed. Each group member's income in the group task is determined by their **performance rank** in the previous individual task.

You and everyone in your group will be told the income of everybody else.

	Players	Incomes
	Player "KO"	8 units
YOU >>>	Player "JS"	15 units
	Player "WS"	23 units
	Player "CO"	51 units
	Player "KD"	3 units
		Total income: 100 units

Your income is 15 units. Your player name in this round is "JS".

Please decide now how many units you want to contribute to the common pool.

(Type a value between 0 and 15.)

Stage 1 is completed. The payoff you earned in Stage 1 is 21.4 units.

STAGE 2: Please choose an action towards the other players:

- A: You pay 2 units to **decrease** the other player's payoff by 6 units.
- B: You pay 0 units. The other player's payoff remains **unchanged**.
- C: You pay 2 units to **increase** the other player's payoff by 6 units.

	Players	Incomes	Contributions	Choose action:
	Player "KO"	8 units	7 units	A B C
YOU >>>	Player "JS"	15 units	14 units	
	Player "WS"	23 units	20 units	A B C
	Player "CO"	51 units	7 units	A B C
	Player "KD"	3 units	3 units	A B C
		Total income: 100 units	Total contributed: 51 units	

Your player name in this round is "JS".

4.4 Experiment 4

4.4.1 Hidden condition

Every year, the Parent-Teacher Association (PTA) in schools across the U.S. organizes fundraisers. The PTA fundraisers ask parents to donate money to help schools afford expenditures that are not covered by federal or state funding, such as a science lab, a choral program, teacher aides, a telescope, dance lessons, math and literacy coaching, and so on.

PTA donations are pooled and distributed across schools and neighborhoods in a school district.

The city government wants to raise more money for school funding, in addition to the independently organized PTA fundraisers. They are planning on adding $2,000 funding per school.

Raising the necessary money for all five schools can be achieved through a small increase in taxes in **from only some parents**.

Which parents do you think should pay the additional tax to cover the $2,000 per school in all five schools?

School	Average PTA donation per school last year	Your decision:
Parents of School PV	$1,227	Taxes should be raised in this neighborhood
Parents of School RW	$4,249	Taxes should be raised in this neighborhood
Parents of School QM	$353	Taxes should be raised in this neighborhood
Parents of School NG	$9,759	Taxes should be raised in this neighborhood
Parents of School HT	$1,486	Taxes should be raised in this neighborhood

4.4.2 Revealed condition

Every year, the Parent-Teacher Association (PTA) in schools across the U.S. organizes fundraisers. The PTA fundraisers ask parents to donate money to help schools afford expenditures that are not covered by federal or state funding, such as a science lab, a choral program, teacher aides, a telescope, dance lessons, math and literacy coaching, and so on.

PTA donations are pooled and distributed across schools and neighborhoods in a school district.

The city government wants to raise more money for school funding, in addition to the independently organized PTA fundraisers. They are planning on adding $2,000 funding per school.

Raising the necessary money for all five schools can be achieved through a small increase in taxes in **from only some parents**.

Which parents do you think should pay the additional tax to cover the $2,000 per school in all five schools?

School	Average PTA donation per school last year	Average household income last year last year	Your decision:
Parents of School PV	$1,486	$75,575	Taxes should be raised in this neighborhood
Parents of School NG	$1,227	$51,141	Taxes should be raised in this neighborhood
Parents of School RW	$353	$35,176	Taxes should be raised in this neighborhood
Parents of School HT	$4,249	$60,422	Taxes should be raised in this neighborhood
Parents of School QM	$9,759	$71,622	Taxes should be raised in this neighborhood

References

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference Points and Effort Provision. *American Economic Review*, *101*(2), 470–492.

Allison, P. D. (1978). Measures of inequality. *American Sociological Review*, *43*(6), 865. http://doi.org/10.2307/2094626

Ambrus, A., & Ben Greiner. (2012). Imperfect Public Monitoring with Costly Punishment: An Experimental Study. *American Economic Review*, *102*(7), 3317–3332. http://doi.org/10.1257/aer.102.7.3317

Amir, O., Rand, D. G., Gal, Y. K., & Gal, Y. K. (2012). Economic Games on the Internet: The Effect of $1 Stakes. *PLoS ONE*, *7*(2), e31461. http://doi.org/10.1371/journal.pone.0031461

Anderson, L. R., Mellor, J. M., & Milyo, J. (2008). Inequality and public good provision: An experimental analysis. *The Journal of Socio-Economics*, *37*(3), 1010–1028. http://doi.org/10.1016/j.socec.2006.12.073

Antinyan, A., Corazzini, L., & Neururer, D. (2015). Public good provision, punishment, and the endowment origin: Experimental evidence. *Working Paper*.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk. *Political Analysis*, *20*, 351–368. http://doi.org/10.1093/polana/mpr057/-/DC1

Bornstein, G., & Weisel, O. (2010). Punishment, Cooperation, and Cheater Detection in “Noisy” Social Exchange. *Games*, *1*, 18–33.

Brick, K., & Visser, M. (2012). Heterogeneity and Voting: A Framed Public Good Experiment. *Working Paper*, 1–18.

Buckley, E., & Croson, R. (2006). Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics*, *90*(4-5), 935–955.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, *6*(1), 3–5. http://doi.org/10.1177/1745691610393980

Cardenas, J. C. (2007). Wealth Inequality and Overexploitation of the Commons: Field Experiments in Colombia1. *Inequality*.

Cardenas, J. C., Stranlund, J., & Willis, C. (2002). Economic inequality and burden-sharing in the provision of local environmental quality. *Ecological Economics*, *40*(3), 379–395. http://doi.org/10.1016/S0921-8009(01)00285-3

Cardenas, J.-C. (2003). Real wealth and experimental cooperation: experiments in the field lab. *Journal of Development Economics*, *70*(2), 263–289. http://doi.org/10.1016/S0304-3878(02)00098-6

Carpenter, J., & Matthews, P. H. (2009). What norms trigger punishment? *Experimental Economics*, *12*(3), 272–288. http://doi.org/10.1007/s10683-009-9214-z

Chan, K. S., Mestelman, S., Moir, R., & Muller, R. A. (1996). The voluntary provision of public goods under varying income distributions. *Canadian Journal of Economics*, *29*(1), 54. http://doi.org/10.2307/136151

Chan, K. S., Mestelman, S., Moir, R., & Muller, R. A. (1999). Heterogeneity and the Voluntary Provision of Public Goods. *Experimental Economics*, *2*(1), 5–30. http://doi.org/10.1023/A:1009984414401

Cherry, T. L., Kroll, S., & Shogren, J. F. (2005). The impact of endowment heterogeneity

and origin on public good contributions: evidence from the lab. *Journal of Economic Behavior & Organization*, *57*(3), 357–365. http://doi.org/10.1016/j.jebo.2003.11.010
Congressional Budget Office. (2007). Average Federal Tax Rates. Retrieved from http://www.cbo.gov/publication/42870
Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. *PLoS ONE*, *8*(3), e57410. http://doi.org/10.1371/journal.pone.0057410
Dal Bo, P. (2005). Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *American Economic Review*, *95*(5), 1591–1604. http://doi.org/10.2307/4132766
Fehr, E., & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, *90*(4), 980–994.
Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140.
Fisher, J., Isaac, R. M., Schatzberg, J. W., & Walker, J. M. (1995). Heterogenous demand for public goods: Behavior in the voluntary contributions mechanism. *Public Choice*, *85*(3-4), 249–266.
Gächter, S., Renner, E., & Sefton, M. (2008). The Long-Run Benefits of Punishment. *Science*, *322*(5907), 1510–1510. http://doi.org/10.1126/science.1164744
Grechenig, K., & Nicklisch, A. (2010). Punishment despite reasonable doubt—a public goods experiment with sanctions under uncertainty. *Journal of Empirical Legal …*.
Hackett, S., Schlager, E., & Walker, J. (1994). The role of communication in resolving commons dilemmas: experimental evidence with heterogeneous appropriators. *Journal of Environmental Economics and …*.
Hauser, O. P., Hendriks, A., Rand, D. G., & Nowak, M. A. (2016). Think Global, Act Local: Preserving the Global Commons. *Scientific Reports*, *6*.
Hendriks, A. (2012). SoPHIE - Software Platform for Human Interaction Experiments. *Working Paper*.
Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial Punishment Across Societies. *Science*, *319*(5868), 1362–1367. http://doi.org/10.1126/science.1153808
Hofmeyr, A., Burns, J., & Visser, M. (2007). Income inequality, reciprocity and public good provision: an experimental analysis. *South African Journal of …*.
Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: conducting experiments in a real labor market. *Experimental Economics*, *14*(3), 399–425. http://doi.org/10.1007/s10683-011-9273-9
Isaac, R. M., & Walker, J. M. (1988). Communication and free-riding behavior: The voluntary contribution mechanism. *Economic Inquiry*, *26*(4).
Keser, C., Markstädter, A., & Schmidt, M. (2014a). Mandatory Minimum Contributions, Heterogeneous Endowments and Voluntary Public-Good Provision. *Working Paper*. http://doi.org/10.2139/ssrn.2539601
Keser, C., Markstädter, A., Schmidt, M., & Schnitzler, C. (2014b). Social cost of inequality - Heterogeneous endowments in public-good experiments. *Working Paper*, 1–28.
Kiatpongsan, S., & Norton, M. I. (2014). How Much (More) Should CEOs Make? A Universal Desire for More Equal Pay. *Perspectives on Psychological Science*, *9*(6), 587–593.

Maryland CPA. (2016, December). Charitable deduction averages 2014. Retrieved April 2017, from http://www.maryland-cpa.com/news/charitable-deduction-averages-2014/
Mason, W., & Suri, S. (2011). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*(1), 1–23. http://doi.org/10.3758/s13428-011-0124-6
McDonald, J. H. (2014). Handbook of Biological Statistics (pp. 1–305). Sparky House Publishing.
Norton, M. I., & Ariely, D. (2011). Building a Better America—One Wealth Quintile at a Time. *Perspectives on Psychological Science*, *6*(1), 9–12.
Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science*, *23*(3), 184–188. http://doi.org/10.1177/0963721414531598
Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D., & Nowak, M. A. (2009). Positive interactions promote public cooperation. *Science*, *325*, 1272–1275.
Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, *489*(7416), 427–430. http://doi.org/10.1038/nature11467
Rand, D. G., Nowak, M. A., Fowler, J. H., & Christakis, N. A. (2014). Static network structure can stabilize human cooperation. *Proceedings of the National Academy of Sciences*, *111*(48), 17093–17098. http://doi.org/10.1073/pnas.1400406111
Reuben, E., & Riedl, A. (2008). Public Goods Provision and Sanctioning in Privileged Groups. *Journal of Conflict Resolution*, *53*(1), 72–93. http://doi.org/10.1177/0022002708322361
Reuben, E., & Riedl, A. (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior*, *77*(1), 122–137.
Sefton, M., Shupp, R., & Walker, J. M. (2007). The effect of rewards and sanctions in provision of public goods. *Economic Inquiry*, *45*(4), 671–690.
Sullivan, M., & Felton, E. (2014, August 6). PTA fundraising data shows massive gap between haves and have-nots. Retrieved April 29, 2017, from http://www.thenewyorkworld.com/2014/08/06/pta-2/
Sutter, M., Haigner, S., & Kocher, M. G. (2010). Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. *Review of Economic Studies*, *77*(4), 1540–1566.
U.S. Census Bureau. (2013). *Income and Poverty in the United States*. Retrieved from http://www.census.gov/hhes/www/income/data/incpovhlth/2013/table4.pdf
Van Dijk, E., & Wilke, H. (1994). Asymmetry of wealth and public good provision. *Social Psychology Quarterly*.
Weissman Center for International Business. (2016). New York City Median Household Income. Retrieved April 29, 2017, from http://www.baruch.cuny.edu/nycdata/income-taxes/med_hhold_income.htm
World Bank. (n.d.). Gini Index. Retrieved from http://wdi.worldbank.org/table/2.9