

Dwenger, Nadja; Sjursen, Ingrid Hoem; Vietz, Jasmin

Working Paper

What Is Fair? Experimental Evidence on Fair Equality vs Fair Inequality

CESifo Working Paper, No. 11289

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Dwenger, Nadja; Sjursen, Ingrid Hoem; Vietz, Jasmin (2024) : What Is Fair? Experimental Evidence on Fair Equality vs Fair Inequality, CESifo Working Paper, No. 11289, CESifo GmbH, Munich

This Version is available at:

<https://hdl.handle.net/10419/305531>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

What Is Fair? Experimental Evidence on Fair Equality vs Fair Inequality

Nadja Dwenger, Ingrid Hoem Sjursen, Jasmin Vietz

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

What Is Fair? Experimental Evidence on Fair Equality vs Fair Inequality

Abstract

Many societies aim to design policies based on meritocratic fairness, which involves two principles: (i) paying individuals with equal performance equally (fair equality) and (ii) paying individuals with higher performance more (fair inequality). Yet, often it is impossible to respect both simultaneously. This paper provides novel evidence on the importance individuals attach to each principle from a large-scale experiment in the United States. We document large heterogeneity in preferences. Individuals incur substantial personal costs to implement their preferred principle. Republican supporters are more likely to prefer fair inequality. The findings offer insights into the political economy of redistribution and public policy design.

Nadja Dwenger
University of Hohenheim / Germany
Nadja.Dwenger@uni-hohenheim.de

Ingrid Hoem Sjursen
Chr. Michelsen Institute, Bergen / Norway
Ingrid.Sjursen@cmi.no

Jasmin Vietz
University of Hohenheim / Germany
Jasmin.Vietz@uni-hohenheim.de

August 9, 2024

The authors would like to thank Ingvild Ålmas, Steffen Altmann, Alexander W. Cappelen, Adrien Dautheville, Dietmar Fehr, Odd-Helge Fjeldstad, Jon Einar Flåtnes, Dorothea Kübler, Andrea Martinangeli, Ricardo Perez-Truglia, Klaus Schmidt, Sebastian Schweighofer-Kodritsch, Vincent Somville, Eric Ø. Sørensen, Robert Stüber, Bertil Tungodden, Georg Weizsäcker, and numerous seminar and conference participants for valuable comments and suggestions. Adriana Condarco-Quesada and Sebastian Fest provided input on survey implementation. Lukas Hofbauer, Sunniva Nygård, and Philip Mielcarek provided excellent research assistance. The project received financial support from the Research Council of Norway through research grant 314479. The experiment obtained Institutional Review Board (IRB) approval and is registered at the Registry for Randomized Controlled Trials operated by the American Economic Association: RCT ID AEARCTR-0010425.

1 Introduction

In Western societies, there is broad support for the concept of meritocracy. Meritocrats consider income inequalities to be fair if they reflect differences in performance, but not otherwise. The meritocratic fairness view thus encompasses two principles: (i) equal income for people with equal performance (fair equality) and (ii) higher income for people with higher performance (fair inequality). Many tax-and-transfer systems worldwide are designed to respect both principles of meritocracy. However, there are situations where it is impossible to perfectly respect both principles simultaneously.

To illustrate this, consider a tax system that preserves fair inequality but accepts violations of fair equality in terms of net income. Violations of fair equality can occur when income from certain sources remains untaxed due to the tax administration's imperfect information and differences in the observability of various income types, leading to different treatment of individuals with identical pre-tax incomes. Alternatively, consider a transfer system that pays the same amount of benefits to two unemployed individuals, irrespective of their performance in retraining or job search. This preserves fair equality but violates fair inequality, as performance differences do not translate into income differences. How would individuals value these tax-and-transfer systems? Or, more generally, how much weight do individuals place on each of the two principles—fair equality and fair inequality? While the necessity of trading off fair equality and fair inequality is a common feature not only in the design of tax-and-transfer systems, but also in many other distributive situations—such as allocation of bonus payments, university admission, and immigration procedures¹—the literature on social preferences has mostly focused on fair inequality considerations alone (see, e.g., research collection in Cappelen and Tungodden, 2019). How individuals navigate distributive conflicts between fair equality and fair inequality has not been explored yet. Given that the theoretical literature indicates that fulfilling both principles is often impossible (Roemer and Trannoy, 2015), this represents a

¹When designing bonus schemes, firms may allocate a bonus to all workers who contributed to the firm's success (prioritizing fair equality), or they can allocate a larger bonus to the best performing workers (prioritizing fair inequality). Often, the trade-off between fair equality and fair inequality leads governments to use several instruments or procedures in parallel. For instance, countries apply fair inequality to immigration, offering special tracks for the most qualified immigrants. Concurrently, they implement the equality principle in immigration programs such as the Diversity Immigrant Visa Program in the US, providing equal chances of a visa to all applicants from countries with historically low rates of immigration. In the realm of university admissions, countries create fair inequality by offering special tracks for applicants with the best grades. Simultaneously, university admission adheres to the fair equality principle by reserving university seats for weaker applicants with an equally long waiting time.

significant blind spot in the empirical literature on distributional preferences.

This paper uses a new experimental approach to provide evidence from the general population in the United States on how individuals trade off fair equality and fair inequality. Our results show that individuals differ in their preference for fair equality versus fair inequality. In the experiment, participants are presented with two pairs of real-world workers and are asked to determine their earnings. Workers differ in the number of standardized tasks that they completed as well as in the initial payments allocated to them. In one pair, both workers completed the same number of tasks. In the other pair, one worker completed 0 tasks and the other worker completed 6 tasks. We tentatively assign initial earnings to each worker such that initial earnings deviate from what meritocrats may consider fair. In the pair with the same number of tasks completed, workers are allocated different earnings (USD 0 and USD 6), violating fair equality. In the pair with different number of tasks completed, workers are allocated identical earnings (USD 3 each), violating fair inequality. Participants in our experiment take the role of spectators and are tasked with redistributing initial earnings within one of the two pairs only. That is, spectators can either implement fair equality (at the cost of accepting a violation of fair inequality) or they can implement fair inequality (accepting a violation of fair equality).

The spectators are randomly assigned to one of three treatments (the No Cost treatment, Costly Equality treatment, and Costly Inequality treatment), where the treatments only differ with respect to the personal cost for spectators. In the No Cost treatment (baseline), both decision alternatives are costless. In the Costly Equality treatment, implementing fair equality has a personal cost for spectators. That is, they receive lower remuneration for participating in the experiment if they choose to establish equality by redistributing within the pair of workers with the same number of tasks completed. In the Costly Inequality treatment, implementing fair inequality has a personal cost for spectators. In both the Costly Equality and the Costly Inequality treatments, we vary the size of the personal cost for spectators in a within-subject design (presented in a randomized order). Each spectator makes two redistributive decisions, which only vary in the size of the personal cost. We randomly select one of the decisions for implementation.

Our analysis provides four main findings. First, we show that the choices of spectators express a clear preference for either fair equality or fair inequality. This is surprising as one may ex-

pect meritocrats to attach equal weights to the fair equality and the fair inequality principles of meritocracy. The proportion of meritocrats who prioritize fair equality is roughly equivalent to the proportion of individuals who prioritize fair inequality. While the literature (surveyed in Cappelen et al., 2020) has mainly focused on the reward principle of meritocracy (fair inequality), we observe that the equality principle of meritocracy (fair equality) is just as important. Second, in our experiment, almost 40% of individuals redistribute in a way to *increase* inequality to reward higher output in one pair, accepting that the two equally performing workers in the other pair are paid unequally (thus violating fair equality). This strong expression of a preference for fair inequality is noteworthy because incentives should not have played a role in our experiment, as the workers were only working before the spectators made their decisions. The earnings assigned to the workers by the spectators thus solely serve as a form of ex-post reward. Third, spectators are ready to bear a substantial personal cost to implement their fairness preferences, being willing to forgo 30% of their own compensation for participation in the experiment. Fourth, the results offer compelling evidence indicating an important link between individuals' political attitudes and their prioritization of the principles of fair equality versus fair inequality. We provide the first set of evidence demonstrating that individuals who self-identify as supporters of the Republican Party tend to place less emphasis on fair equality relative to fair inequality, which is consistent with a preference for lower levels of redistributive taxation.

Our paper ties into various strands of the literature. First, it advances a literature measuring preferences for distribution (Alesina and Guiliano, 2011; Alesina et al., 2018; Almås et al., 2020; Fehr et al., 2022a; Fong, 2001; Krawczyk, 2010; Mollerstrom et al., 2015; Hvidberg et al., 2023). Prior literature documents the prevalence of meritocratic fairness views both in the laboratory and in the field (e.g., Almås et al., 2020; Andre, 2024; Bhattacharya and Mollerstrom, 2023; Cappelen et al., 2013; Fong, 2001). Although the two principles inherent in meritocracy (reward higher performance through fair inequality and treat equal performance equally through fair equality) are well-described in the theoretical literature (e.g., Bossert, 1995; Fleurbaey, 2008; Fleurbaey and Maniquet, 2011), we are not aware of prior empirical research that studies the importance individuals attach to each of these principles. This is surprising, especially considering that theoretical literature concludes that respecting both principles is generally impossible in an environment which is sufficiently rich (see discussion in Roemer and Trannoy, 2015, p. 243). We thus add a new and important aspect to the empirical description

of people’s preferences for distribution. In previous studies, researchers have presented individuals with unequal income distributions to assess their willingness to redistribute income, thus gauging their tolerance for inequality. Starting with a pair of workers who have different performance levels but identical initial earnings, we discover that many participants are willing to *create* income inequality to reward higher performance. This choice comes at the expense of their ability to equalize earnings for another pair of workers with identical performance levels. We therefore interpret our results as a stronger indication of a fairness preference for rewarding performance.² Our empirical findings are also relevant for the traditional public finance literature regarding the fundamental requirements of a tax system. Earlier literature described two requirements: “horizontal equity”, which stipulates that individuals in equal position should be treated equally, emphasizing fair equality, and “vertical equity”, which posits that individuals in different positions should be treated differently, emphasizing fair inequality. While earlier literature on this topic was primarily theoretical (Atkinson and Stiglitz, 1980; Kaplow, 1989; Musgrave, 1990, 1993; Berg, 2021), our paper introduces empirical evidence to enhance the understanding of the relevance of both concepts.

Second, this study makes methodological contributions to the experimental literature on social preferences and experimental research in general (Almås et al., 2020; Bolton and Ockenfels, 2000; Cappelen et al., 2007, 2013; Charness and Rabin, 2002; Cherry et al., 2002; Engelmann and Strobel, 2004; Fehr and Schmidt, 1999; Konow, 2000, 2009). Previous studies identify the determinants of distributional decisions by comparing distributional outcomes across various experimental treatments (e.g., Almås et al., 2020; Cappelen et al., 2007, 2013; Engelmann and Strobel, 2004). Typically, participants have been presented with *one* distribution problem at a time. In contrast, our approach is different. Using a spectator design, we explicitly model the trade-offs involved by presenting study participants with *two* redistribution problems simultaneously, with each problem representing one aspect of the trade-off. Spectators can only choose to solve one of the redistribution problems, which reveals the aspect of the trade-off to which they give greater importance. The fact that spectators choose fair equality or fair inequality in the treatments, even though they have to bear personal costs, serves as additional evidence of reliability. Our approach, which directly represents the trade-off, may also prove beneficial in other contexts where previous research has assessed trade-offs by comparing outcomes (such

²Also Abeler et al. (2010) find evidence in favor of individuals perceiving inequality as fair if workers perform differently, even though their focus is mainly on effort choices in a lab experiment.

as individuals' marginal willingness to pay) across treatments.

Third, our results contribute to the literature on the foundations of political attitudes (Alesina and Angeletos, 2005; Alesina and Guiliano, 2011; Bénabou and Ok, 2001; Bénabou and Tirole, 2006a; Fehr et al., 2022a,b; Fisman et al., 2017; Roth and Wohlfahrt, 2018). We advance this literature by showing that supporters of the Republican party are more likely to prioritize fair inequality over fair equality. Our findings suggest that political disputes over distributional situations may arise not only from divergent views on the importance of luck and effort for people's entitlements (Almås et al., 2020) or the management of false positives and false negatives (Cappelen et al., 2023) but also from fundamental distinctions in how individuals balance fair equality against fair inequality in their policy considerations.

The paper is organized as follows. Section 2 provides a simple theoretical framework to guide our empirical analysis. The experimental design for our main experiment is presented in Section 3, while the pre-test design is explained in Section 4. Section 5 describes our empirical strategy. The results on the choices of spectators are in Section 6. Section 7 concludes.

2 Theoretical Framework

To guide our empirical analysis, we provide a simple theoretical framework, extending the framework developed in Cappelen et al. (2013) and Almås et al. (2020). This framework assumes that a spectator dislikes if an individual's payment deviates from what she considers to be the fair payment for this individual.

We consider a situation in which a fixed payment is divided among two workers who completed a certain number of tasks. Let F^{equal} be the fair payment share of a worker in a pair where both workers completed the same number (or share) of tasks. We refer to this situation as fair equality. If workers completed a different number of tasks, let $F^{unequal}$ be the fair payment share of the worker who completed the higher share of tasks in this pair. We refer to this situation as fair inequality. I denotes a worker's actual payment share in the respective pair. In our experiment (see Section 3), spectators are faced with a pair of equal workers and a pair of unequal workers simultaneously. They need to trade off deviations from fair equality against deviations from fair inequality. The following utility function $V(I)$ for spectators captures this trade-off,

embodying both deviations from fair equality in the pair of equal workers (first summand) and deviations from fair inequality in the pair of unequal workers (second summand):

$$V(I) = -(1 - \beta)[I - F^{equal}]^2 - \beta[I - F^{unequal}]^2. \quad (1)$$

We allow for heterogeneity in the importance spectators place on treating unequals unequally (fair inequality) and on treating equals equally (fair equality), captured by $0 \leq \beta \leq 1$ (weight attached to fair inequality) and $1 - \beta$ (weight attached to fair equality). A spectator places more weight on fair inequality if $\beta > 0.5$ and vice versa.

In the model, spectators may differ along two dimensions: First, in what they consider to be the fair payment shares F^{equal} and $F^{unequal}$ (the “fairness norm”). Second, in how they evaluate deviations from these fair payment shares, captured by β .

The vast majority of individuals in the US can be identified as **meritocrats** (Almås et al., 2020; Cappelen et al., 2022), i.e., individuals who believe that inequalities in earnings are fair if they reflect differences in performance, but not otherwise. For them, the fair payment share depends on whether workers completed an equal or unequal number of tasks. Specifically, for proportionalist meritocrats, the fair share of total payments for an individual corresponds to the individual’s share in total tasks completed.

If the deviations from the fairness norms of fair equality ($[I - F^{equal}]^2$) and of fair inequality ($[I - F^{unequal}]^2$) are identical, the model allows classifying spectators into three different categories based on β :

1. **Rewarding** individuals with $\beta > 0.5$
2. **Indifferent** individuals with $\beta = 0.5$
3. **Equalizing** individuals with $\beta < 0.5$.

Ex ante, we expect meritocrats to place equal importance on treating unequals unequally (fair inequality) and on treating equals equally (fair equality). If this holds true, many meritocrats should be indifferent when faced with the choice of implementing either fair inequality or fair equality, exhibiting $\beta = 0.5$. Of course, in principle, individuals could also attach greater

importance to one of the two principles. Ultimately, the distribution of β in the population is an empirical question for which we currently lack evidence.

3 Experimental Design

Based on the theoretical framework, we investigate the choices made by spectators when they can either eliminate deviations from fair equality *or* eliminate deviations from fair inequality. That is, when spectators can either equally pay workers with the same number of tasks completed (create fair equality) *or* reward workers who completed more tasks with higher pay (create fair inequality).

Following Almås et al. (2020) and Cappelen et al. (2013), we use a spectator design with two types of participants: *workers* and *spectators*. The workers can perform standardized real effort tasks. The spectators determine the pay of workers, by redistributing payments within pairs of workers. Below, we outline the details of our pre-registered and IRB-approved experimental design.³

3.1 Spectators

The spectators participate in an online experiment. They are recruited via the survey company Norstat and are only paid if they complete the survey fully.⁴ Our sample consists of 1,525 spectators, recruited from the general population in the US. Compared to the full population, our respondents are slightly more likely to be female, less likely African-American, more educated, somewhat older, and more likely to dispose of below median income (median equivalized income is between USD 25,000 and USD 49,999 in our sample compared to median equivalized income of between USD 50,000 and USD 74,000 in the full US population) (Table 1). To address these imbalances, we reweight the sample in robustness checks so that the sample is representative along the gender, age, ethnicity, education, and income dimensions as well. Results are robust to reweighting (Table A.2).

The main focus of the experiment is the redistribution decision, which is outlined below (with incentivized treatments). In the instructions for spectators, we emphasize that the spectator's

³We provide the experimental protocols for workers in the online appendix. Spectators' redistribution decision is also displayed in the online appendix. The complete spectator survey can be accessed [via this link](#).

⁴The baseline compensation for participation for the spectators is equivalent to USD 2.

Table 1: Descriptives sample and US population

	Sample	US Population
Female (share in %)	61	51
Ethnicity		
European American (share in %)	82	62
African American (share in %)	6	12
Higher Education (share in %)	53	33
Republican (share in %)	30	31
Age, median (in years)	49	39
Equivalized Income, median (US census category)	3	4
Observations	1525	.

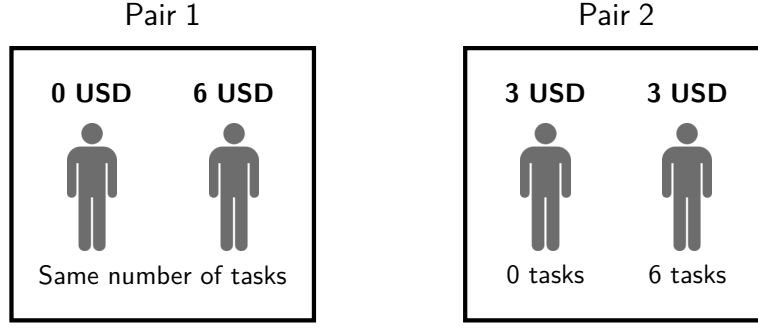
Notes: This table contains summary statistics for our sample of spectators alongside nationally representative statistics from the Census Bureau and the Current Population Survey. Table A.1 in the appendix contains further details. Education: An individual is defined as having higher education if his highest level of education is a completed 4-year college degree or more. Income: Equivalized income calculated with the same method used in the US census. Income category 3 corresponds to equivalence-adjusted incomes between USD 25,000 and USD 49,999, income category 4 corresponds to equivalence-adjusted incomes between USD 50,000 and USD 74,999. Republican: An individual is defined as Republican if his political affiliation is Republican.

decision has real consequences for others and determines the distribution of payments between real individuals. Furthermore, spectators are aware that workers were paid a participation fee and had the opportunity (but no obligation) to complete standardized tasks with the prospect of receiving additional payment.

In the redistribution decision, spectators are presented with two pairs of workers, as shown in Figure 1. In one pair, both workers have completed the same number of tasks. For these pairs, spectators are informed only that the number of tasks completed is equal, not the exact number. We deliberately do not inform spectators on the exact number of tasks that equal workers completed to prevent comparisons across pairs.⁵ In the other pair, one worker completed 0 tasks and the other worker completed 6 tasks. The pairs are displayed in a randomized order. We assign initial payments to each worker such that initial payments deviate from what spectators may consider fair. In the pair with the same number of tasks completed, workers are allocated different payments (USD 0 and USD 6). In the pair with different number of tasks completed, workers are allocated identical payments (USD 3 each).

⁵The pairs may consist of two workers who both completed 1 task, both completed 2 tasks etc. As shown in Figure A.2, the most common pair consists of two workers who both completed 5 tasks. Details on the pairing are in Section 3.3.

Figure 1: Spectator decision



We ask spectators to redistribute payments in *one* of the two pairs. Thus, the trade off for spectators is between treating equals equally vs. treating unequals unequally. More precisely, spectators can either (i) redistribute USD 3 within the pair of workers with the same number of tasks completed and establish fair equality *or* (ii) redistribute USD 3 in the pair of workers with different number of tasks completed and establish fair inequality. Redistributing within the pair of workers with same number of tasks (in Figure 1: pair 1) implies that both workers receive USD 3. That is, equal payment for equal workers. Redistributing within the pair of workers with different number of tasks completed (in Figure 1: pair 2) implies that the worker who completed 6 tasks will receive USD 6 and the worker who completed 0 tasks will receive USD 0. That is, unequal payment to reward a higher number of tasks completed.

Remember from the theoretical framework in Section 2 that spectators may differ along two dimensions: (i) in what they consider to be the fair payments F^{equal} and $F^{unequal}$; and (ii) in the weight they place on fair equality ($1-\beta$) versus fair inequality (β). In this study, we are interested in β . As we show in Section 4, the majority of individuals in a pre-test (recruited in the same way as individuals in the main experiment) are proportionalists. We therefore designed our experiment in such a way, that proportionalists face identical deviations from the fair payment in both pairs of workers $[I - F^{equal}]^2 = [I - F^{unequal}]^2$. In the pair of workers who completed the same number of tasks, proportionalists consider an equal payment to be fair ($F^{equal} = \text{USD } 3$). Yet, initial payment assigned is USD 6, resulting in a deviation between initial and fair payment of 9 ($[I - F^{equal}]^2 = [6 - 3]^2 = 9$).⁶ In the pair with one worker completing all the 6 tasks (and the other worker completing 0 tasks), the fair payment to this worker

⁶Without loss of generality, we focus on the worker who initially was assigned a payment of USD 6. The deviation from the fair payment for the other worker is the same ($[I - F^{equal}]^2 = [0 - 3]^2 = 9$).

is USD 6. With an initial payment of USD 3, the deviation between fair and initial payment in this pair is again 9 ($[I - F^{unequal}]^2 = [3 - 6]^2 = 9$).⁷ To sum up, the initial payments and tasks completed are such that proportionalist meritocrats face identical deviations from what they consider a fair payment in both pair of workers. As a result, whether these proportionalist meritocrats implement fair equality or fair inequality in our experiment only depends on β , the weight that they attach to deviations from fair inequality.

Before and after the redistribution decision, participants answer non-incentivized questions on standard background characteristics (gender, age, geographic location, education, equivalized income, political orientation) as well as questions on their opinions on and experiences with fair and unfair (in)equality.

3.2 Experimental Treatments

We implement three different treatments to study the importance that individuals attach to fair equality versus fair inequality in their distributive preferences. In all treatments, spectators are presented with two pairs of workers as described above (one pair with equal number of tasks completed and one pair with different number of tasks completed). The initial distribution of payments is identical across treatments and spectators are asked to redistribute within one of the two pairs in all treatments. The treatments only differ in the *personal* cost that spectators need to incur for choosing one of the decision alternatives: fair equality or fair inequality. We randomize one third of spectators into each of the following treatments (see Table 2): (i) No Cost treatment; (ii) Costly Equality treatment; (iii) Costly Inequality treatment. The between-subject implementation enables us to rule out that spectators are indifferent between fair equality and fair inequality and it allows to investigate the strength of individual preferences for either fair equality or fair inequality. Background characteristics are balanced across all three treatments (see Table OA.1 in the online appendix), i.e., randomization into treatments was successful.

We now provide details on each treatment. In the No Cost treatment (baseline), none of the two

⁷In our calculations, we focus on the worker with more tasks completed. As the deviations from the fair payment are symmetric for the workers completing 6 and 0 tasks, respectively, we do so without loss of generality.

Table 2: Treatment variation

	Treatments (between-subjects)				
	No Cost	Costly Equality		Costly Inequality	
Decision 1	No Cost	High Cost	Low Cost	High Cost	Low Cost
Decision 2	Endogenous Cost	Low Cost	High Cost	Low Cost	High Cost
# of Spectators	509	244	264	264	244

Notes: This table provides an overview of the between-subject variation from treatments and the within-subject variation from asking participants to take two decisions in the experiment. In the No Cost treatment, our focus is on the first decision. “No Cost” indicates that none of the decision alternatives is associated with a personal cost for spectators. In “High Cost” the personal cost for spectators is 30% of what they receive for participating in the survey, and in “Low Cost” the personal cost equals 10%. The personal cost for spectators for the decision alternative which they chose in their decision 1 in “Endogenous Cost” equals 1% of their remuneration.

decision alternatives is associated with a cost.⁸ In the Costly Equality treatment, implementing fair equality has a personal cost for spectators. That is, they receive lower remuneration for participating in the experiment if they choose to establish equality by redistributing within the pair of workers with the same number of tasks completed. In the Costly Inequality treatment, it is the implementation of fair inequality which has a personal cost for spectators. That is, they receive lower remuneration if they choose to create fair inequality by rewarding more tasks completed with a higher pay.

In both the Costly Equality and the Costly Inequality treatments, we vary the size of the personal cost for spectators in a within-subject design. To do so, we sequentially present spectators with two redistributive decisions, which only vary in the size of the personal cost. In one decision (“low cost decision”), the personal cost for spectators is 10% of what they receive for participating in the survey. In the other decision (“high cost decision”), the personal cost is 30% of their participation remuneration. The order of high and low cost decisions is random-

⁸This refers to the first decision in the No Cost treatment, which forms part of our main analysis. In the second decision of the No Cost treatment, we introduce a very small cost (of 1%), which is determined endogenously: The decision alternative that the spectators choose in their first redistribution decision becomes slightly costly in their second redistribution decision. This small cost in the second decision problem is designed to examine within-subject whether individuals are indifferent between fair equality and fair inequality. We anticipate that indifferent spectators will change their choice to maximize their payoff from the experiment. Only those spectators with a preference for fair equality or fair inequality will incur the cost and maintain their choice. A potential concern with the results from the within-subject variation is that spectators may maintain their choice for other reasons. For instance, individuals might choose the same decision alternative again as they try to behave consistently (Falk and Zimmermann, 2013). When describing our results, we therefore primarily focus on those identified through between-subject variation, while also noting that the results identified through within-subject variation are qualitatively the same.

ized across spectators. Hence, half of spectators in each the Costly Equality and the Costly Inequality treatments take the high cost decision first, while the other half of spectators take the low cost decision first (see Table 2).

3.3 Workers

To conclude the exposition of our experimental design, we next describe the workers whose payment is determined by spectators. We recruit 781 workers⁹ via the online crowd-sourcing platform [Connect](#). As also communicated to spectators, workers are paid a participation fee of USD 2 which may be topped up by an additional bonus pay. Workers are aware that a third person decides on the bonus pay, i.e., whether and how much they receive as a bonus. They also know that working on standardized real effort assignments may increase their bonus.

Workers can complete up to four standardized real effort assignments. For each assignment, workers decide whether they want to work on the assignment or not. If they decide to work, the assignment starts and workers are subsequently presented with six tasks. Each task is standardized and lasts for one minute. During this minute, workers are presented with a matrix of 48 three-digit numbers. They are asked to tick off every number that lies within a given interval of numbers.¹⁰ The matrix is automatically submitted after 60 seconds. If workers correctly tick off every code, the task is labeled as *completed* (and as not completed otherwise). Workers can complete between 0 to 6 tasks in each assignment. Workers who decide not to work on the assignment, have 0 completed tasks for this assignment and are forwarded to the next assignment.

Figure A.1 in the appendix shows the distribution of tasks per assignment that the 781 workers completed (all assignments pooled). There are two modes in the distribution: working 0 or working 6 tasks. While there are some workers with 1 task (2.6%) completed, the share of workers is monotonically increasing in the number of tasks completed (reaching 20% for 5 tasks and 24% for 6 tasks completed).

We form pairs of workers by matching each worker with another worker for each assignment.

⁹To ensure an adequate number of equality and inequality pairings, we had to recruit an additional 31 workers in addition to the originally planned 750 workers. The bonus payments of the surplus pairs are determined by randomly drawn spectators from the pre-test (see Section 4).

¹⁰See Section E in the online appendix for an example of this code recognition task.

We first combine the worker-assignments with 0 or 6 tasks completed to 0-6 pairs. Then we combine the remaining worker-assignments such that the two workers in a pair completed the same number of tasks (forming 1-1, 2-2, ..., and 5-5 pairs). Finally, we tentatively assign payments (see Section 3.1) and match two pairs of workers (one 0-6 pair and one pair with the same number of tasks completed) to each spectator. We pay out the bonus determined by the spectators within a few weeks after the spectators had taken their redistribution decision.¹¹

4 Pre-test

An important aspect of our experimental design is that the initial payment allocations (I) equally deviate from the fairness norm (captured in F^{equal} and $F^{unequal}$, see Section 2) in both pairs. This ensures that spectators choose to implement fair inequality or fair equality because they have a stronger preference for one of the two principle embodied in meritocracy, not because the degree of unfairness is greater in one pair than the other.

In our main experiment, piece-rate pay is an obvious way to implement fair inequality in the pair with different number of tasks completed. That is, to pay the worker who completed zero tasks USD 0 and the worker who completed 6 tasks USD 6.

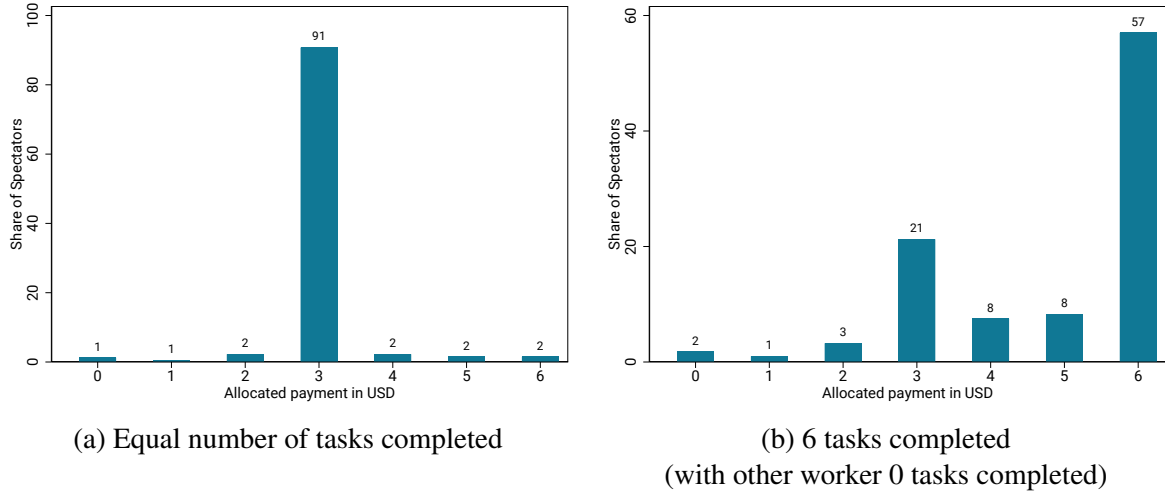
We subject the presumption of piece-rate pay to an empirical pre-test by recruiting 400 different spectators, following the same recruitment approach as in the main experiment. In the non-incentivized pre-test, we asked them to distribute payments freely in a hypothetical payment situation.

The experimental set-up closely follows that of our main experiment. Participants are presented with two pairs of workers identical to those in our main experiment. Unlike in the main experiment, workers are not allocated any initial payment. Instead, we instruct spectators to distribute USD 6 freely within each pair.¹²

¹¹To determine workers' bonus payments, we randomly select one decision and one pair per spectator, independent of which pair the spectator chose for redistribution.

¹²In addition, we ask participants to make a redistributive choice as a second decision. As the first (distributive decision) and the second (redistributive decision) are structurally very similar but obviously differed in the presentation of the decision problem, we are afraid that an experimenter demand effect may affect the second decision. In what follows, we therefore focus on participants' first decisions. A test version of the pre-test survey can be accessed [via this link](#).

Figure 2: Payment for worker 2 allocated in pre-test



The figure illustrates the distribution of payments allocated by spectators in both pairs. Panel (a) pertains to the workers who completed an equal number of tasks. Panel (b) focuses on the worker who completed 6 tasks, with the other worker completing 0 tasks.

Figure 2 displays the payments allocated by spectators in both pairs. In the pair where workers completed the same number of tasks, the majority of participants (91%) allocates USD 3 to each worker (Panel (a)). That is, workers with an equal number of completed tasks are generally allocated equal payments. Panel (b) illustrates the payments allocated to the worker who completed 6 tasks in the unequal pair, where the other worker completed 0 tasks. Two observations from Panel (b) are important for our main analysis.

Observation 1: 21% of participants allocate USD 3 to both workers in the pair, regardless of their differing task completion.¹³ We label these participants *egalitarians*, individuals who view equal payment as fair regardless of output.

Observation 2: The majority of participants (73%) allocate a larger payment to the worker who completed 6 tasks. We label these participants *meritocrats*, individuals who reward higher output with higher payment.¹⁴

¹³This number is in the same ballpark as the share of egalitarians identified in Almås et al. (2020).

¹⁴The proportion of meritocrats closely matches the 76% of individuals, as reported by (Cappelen et al., 2022, p. 2107 et seq.), who allocate more than half of the total earnings to the better-performing individual when performance is certain.

5 Empirical Strategy

We specified our empirical strategy in a pre-analysis plan before the start of the data collection.¹⁵ The main outcome of interest in the study is the indicator variable $D_{i,d}$, which takes value one if a spectator i implements fair inequality (and zero if the respondent implements fair equality) in decision d .

We analyze all redistributive decisions made by spectators in both the Costly Equality and Costly Inequality treatments. In the No Cost treatment, we concentrate on spectators' first decisions, as second decisions involve an endogenous cost (see Section 3.2).¹⁶ In all regressions, we cluster standard errors at the spectator level.

5.1 Main Analysis

As our main empirical specification, we estimate the following equation, using both a Linear Probability Model and a Probit Model:

$$D_{i,d} = \alpha + \delta^E \text{CostEqual}_{i,d} + \delta^I \text{CostInequal}_{i,d} + \beta \mathbf{X}_i + \varepsilon_{i,d} \quad (2)$$

where $\text{CostEqual}_{i,d}$ and $\text{CostInequal}_{i,d}$ indicate spectators i in the Costly Equality and in the Costly Inequality treatments, respectively; the No Cost treatment is the reference category. Note that each spectator makes two redistributive decisions; $d = 1, 2$ indicates spectators' first and second decisions, respectively. We estimate equation (2) with and without control variables. The specification with controls \mathbf{X}_i for individual background characteristics includes indicator variables for gender (*Female*), age above the median (*Older*), equivalized income above the median (*Higher Income*), completion of 4 or more years of college (*Higher Education*), and support of the Republican party (*Republican*).

To test the impact of cost size on spectators' choices, we estimate equation (3), using separate indicators for low cost (L) and high cost (H). Specifically, the indicator variable $\text{CostEqual}_{i,d}^L$ ($\text{CostEqual}_{i,d}^H$) equals one if spectator i incurs a *low* cost L (*high* cost H) for restoring equality

¹⁵See doi.org/10.1257/rct.10425-2.1

¹⁶As specified in the pre-analysis plan, we test for order effects in spectators' choices in the Costly Equality and the Costly Inequality treatments. We examine whether there is a systematic difference between decisions taken first and decisions taken second. We find no significant effect on our outcome variable. Hence, as pre-specified, we do not control for the order of decisions in the analysis.

in decision d and 0 otherwise.

$$D_{i,d} = \alpha + \rho^{EL} CostEqual_{i,d}^L + \rho^{IL} CostInqual_{i,d}^L + \rho^{EH} CostEqual_{i,d}^H + \rho^{IH} CostInqual_{i,d}^H + \beta \mathbf{X}_i + \varepsilon_{i,d}. \quad (3)$$

We estimate equation (3) with and without the vector of controls \mathbf{X}_i (see above).

5.2 Heterogeneity Analysis

To explore heterogeneity in preferences for fair equality and fair inequality with respect to background characteristic Z_i , we introduce an indicator Z_i and its interactions with the treatment indicators into Equation (2), resulting in Equation (2').

$$D_{i,d} = \alpha + \gamma^Z Z_i + \delta^E CostEqual_{i,d} + \delta^I CostInequal_{i,d} + \delta^{EZ} CostEqual_{i,d} \times Z_i + \delta^{IZ} CostInequal_{i,d} \times Z_i + \beta \mathbf{X}_i + \varepsilon_{i,d} \quad (2')$$

In this regression, \mathbf{X}_i includes all background characteristics from above, except the variable captured by Z_i .

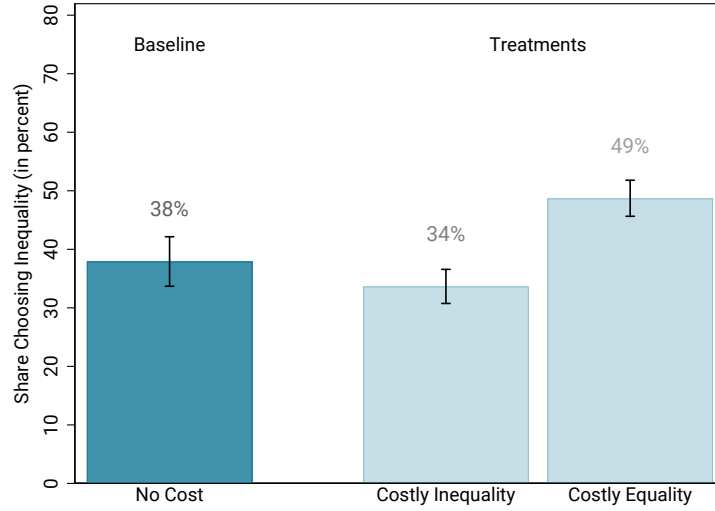
We consider heterogeneity with respect to age, gender, income, education, and political affiliation.¹⁷ In Sections 6.2 and 6.3, we discuss two dimensions of heterogeneity. First, we examine variations in the distributive behavior among different subgroups. Second, we explore whether treatment effects are consistent across subgroups, investigating the interactions between background variables and treatments.

6 Results

We start by presenting the main results before moving on to the analysis of subgroup differences, heterogeneous treatment effects, and the correlation between stated and revealed preferences.

¹⁷As outlined in the pre-analysis plan, we also investigated how individuals' stated preferences and experiences related to fair equality and fair inequality impact behavior. We do not observe any significant effects, though we acknowledge that this could be attributed to limited variation in our measure of personal experience, see discussion in footnote 21 in Section 6.2.

Figure 3: Percentage of spectators establishing fair inequality by treatment



Notes: The figure depicts the percentage of spectators who choose to establish fair inequality (bars) together with 95% confidence bands (error bars). The first bar (dark blue) pertains to the baseline. The second and third bars (light blue) pertain to the Costly Inequality and Costly Equality treatments, respectively. Sample: All decisions from Costly Inequality and Costly Equality treatments, first decisions from No Cost treatment.

6.1 Main Analysis

Figure 3 presents our core findings concerning the percentage of spectators who choose to establish fair inequality. In our experiment, spectators are given the choice to either implement fair equality (workers with equal number of tasks completed receive identical payment) *or* establish fair inequality (higher payment for the worker with more tasks completed). Remember that in the baseline, neither of these alternatives involves any cost. The first bar demonstrates that, in the baseline, 38% of individuals choose to establish fair inequality. We refer to these individuals as *rewarding meritocrats*: For them, it is more crucial to reward a higher number of completed tasks with greater payment than to ensure equal payment for individuals with the same number of tasks completed, revealing $\beta > 0.5$.

The remaining 62% of individuals choose to implement fair equality. This percentage conceals two distinct types of spectators. First, there are the egalitarians (21% of the population, as described in Observation 1 in Section 4). These individuals believe that equal payment is inherently fair, regardless of worker performance. Second, there are the meritocrats who place greater emphasis on the equality aspect of meritocracy. For them, it is more important to ensure equal payment for individuals with equal number of tasks completed than to reward higher

output. Their behavior reveals $\beta < 0.5$. We refer to these individuals as *equalizing meritocrats*. They represent 41% ($= 62\% - 21\%$) of the population, with a standard error of 2.25 percentage points.¹⁸

The similar percentages of rewarding meritocrats (38%) and equalizing meritocrats (41%) raise concerns about whether meritocratic spectators in the baseline might be simply indifferent between the two decision alternatives. In other words, spectators might attach equal weight to deviations from fair equality and deviations from fair inequality ($\beta = 0.5$). If $\beta = 0.5$ (or β close to 0.5), we would anticipate significant differences between the baseline and our treatments, which involve personal costs for spectators for either fair equality or fair inequality. Therefore, we proceed by comparing the behavior observed in the treatments to that observed in the baseline.

First, we assess whether there is a change in the proportion of individuals creating fair inequality in the Costly Inequality treatment. A comparison between the bars for Costly Inequality and No Cost in Figure 3 shows a small difference in the share of spectators choosing fair inequality ($p = 0.06$, one-sided test): Making the choice of fair inequality personally costly, reduces the share of spectators choosing fair inequality by 4 percentage points. Table 3 reports the corresponding regression results and marginal effects for fair inequality based on equation (2), estimated using both a Linear Probability model (LPM) and a Probit model. In all regressions, we observe a small and statistically barely significant effect of the Costly Inequality treatment, even when including control variables in columns (3) and (4).¹⁹ When considering only spectators' first decisions (as reported in Table A.4 in the appendix), the coefficient for the Costly Inequality treatment is statistically insignificant at all common significance levels. Remarkably, even with the introduction of a substantial monetary cost for spectators themselves, most rewarding meritocrats continue to create fair inequality. This is noteworthy, as it indicates their willingness to incur a substantial *personal cost* to reward higher output among others. This underscores the importance of distinguishing equality from fairness. While a vast literature on

¹⁸Standard error for the difference in the two proportions calculated as $\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = \sqrt{\frac{0.62(1-0.62)}{2541} + \frac{0.21(1-0.21)}{400}}$.

¹⁹For eight observations, the vector of controls is incomplete. These observations are excluded from the analysis in columns (3) and (4). The estimates in columns (1) and (2) remain unchanged when we estimate the regressions from columns (1) and (2) using the restricted sample. The coefficient for the Costly Inequality treatment is more precisely estimated and becomes significant due to the inclusion of control variables, not because of the exclusion of observations with incomplete vector of controls.

Table 3: Regression results and marginal effects on created fair inequality

	Dependent Variable: Creating Fair Inequality			
	(1) LPM	(2) Probit	(3) LPM	(4) Probit
Costly Inequality	-0.043 (0.028)	-0.043 (0.028)	-0.047* (0.028)	-0.047* (0.028)
Costly Equality	0.108*** (0.028)	0.108*** (0.028)	0.104*** (0.029)	0.104*** (0.029)
Control Group Mean	0.379*** (0.022)		0.424*** (0.033)	
Controls	No	No	Yes	Yes
Observations	2541	2541	2385	2385
R^2	0.020		0.041	

Sample: All decisions, except the 2nd decisions in the No Cost Treatment.

Notes: LPM stands for Linear Probability model. The table displays marginal effects for the Probit model. The Control Group Mean refers to the No Cost treatment as the reference category. Control variables include indicators for gender, age above median, equivalized income above median, higher education (4-year college degree or more), and Republican affiliation. Clustered standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

preferences for redistribution primarily examines individuals' willingness to promote equality in outcomes, our findings highlight the need for more research on individuals' willingness to facilitate inequality.

Next, we investigate whether a cost for fair equality affects the share of individuals choosing fair inequality. The third bar in Figure 3 represents results from the Costly Equality treatment. Compared to the baseline, it shows an increase in the percentage of spectators choosing fair inequality from 38% to 49%. This observation is confirmed by a positive and statistically significant coefficient for the Costly Equality variable in Table 3 ($p < 0.01$, both with and without controls). In other words, an additional 11% of all spectators (without controls) choose fair inequality if they must incur a personal cost to pay workers with equal output equally. However, even for the Costly Equality condition, this difference across treatments remains relatively small when compared to the surprisingly stable share of spectators creating fair inequality under

all conditions.²⁰

As noted in Table 1, we observe slight imbalances in our sample compared to the US population. Reweighting the sample to match the composition of the full US population across gender, age, ethnicity, education, and income dimensions leaves our results unaffected (see Table A.2 in the appendix).

The analyses above lead us to our first set of results:

Result 1a: Nearly 40% of the US population are *rewarding meritocrats*. They choose to *create inequality* to reward higher output, preferring fair inequality over equalizing payments among those with the same output.

Result 1b: Approximately 40% of the US population are *equalizing meritocrats*. They choose to *equalize* payments and treat individuals with the same output equally, preferring fair equality over rewarding higher output.

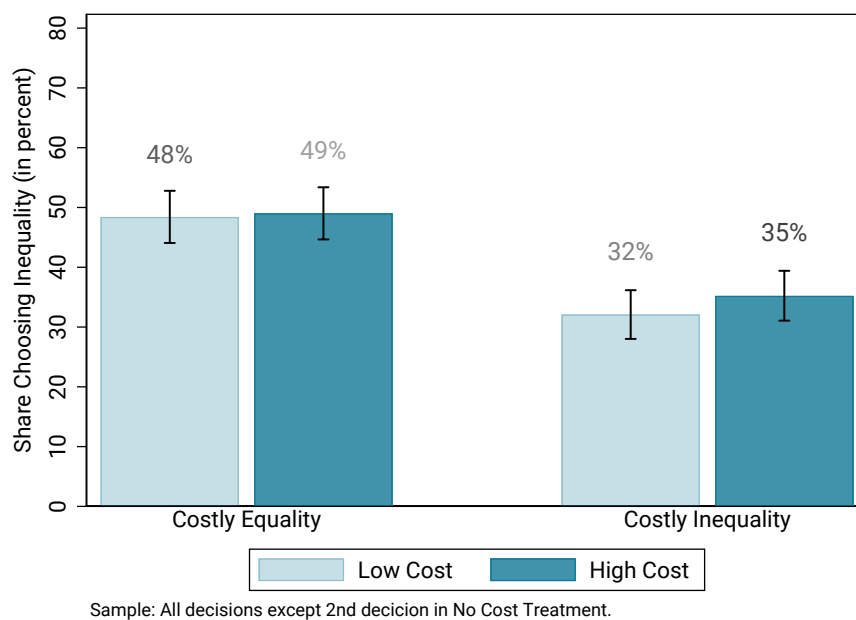
Result 2a: Choices in the Costly Equality and Costly Inequality treatments are very similar to those in the baseline. That is, the choices of spectators are driven by fairness preferences, not indifference.

Result 2b: Overall, we find that only 15% of all spectators reveal a β close to or equal to 0.5. This means that only a small share of the US population is indifferent between fair equality and fair inequality, underlining the importance of carefully distinguishing between both principles.

We conclude our main analyses by examining whether the size of the personal cost for spectators makes a difference. In Figure 4, we dissect the overall impact into separate effects for low and high costs in both the Costly Equality and Costly Inequality treatments (estimated using equation (3)). This analysis reveals an additional insight: the effects are quite similar for low and high personal cost, leading us to summarize this finding in the following way:

²⁰In a within-subject design, we further varied the cost of choosing fair equality/ fair inequality in the No Cost treatment. Specifically, one of the decision alternatives in the second decision problem incurred a small personal cost for spectators, depending on what spectators chose in their first decision problem. If spectators chose to implement fair equality (and fair inequality, respectively) in their first decision problem, choosing fair equality (and fair inequality, respectively) became costly in their second decision problem. Remarkably, spectators tended to stick to their first choice in their second decision problem, even when faced with a personal cost. We acknowledge that this behavior may partly result from spectators trying to maintain consistency. Therefore, in the main text, we primarily focus on the between-subject variation.

Figure 4: Little impact of size of personal cost on percentage of spectators choosing fair inequality



Notes: The figure depicts the percentage of spectators who choose to establish fair inequality (bars) together with 95% confidence bands (error bars). The first two bars pertain to the Costly Equality treatment and the last two bars to the Costly Inequality treatment. The bars in light blue refer to the low cost treatment and the bars in dark blue to the high cost treatment.

The figure depicts the regression results from Table A.5, column (1) in the appendix.

Result 3: In both treatments, the coefficients for high (30% of participation remuneration) and low (10% of participation remuneration) costs do not significantly differ from each other. Individuals are ready to bear substantial personal costs for implementing their preference.

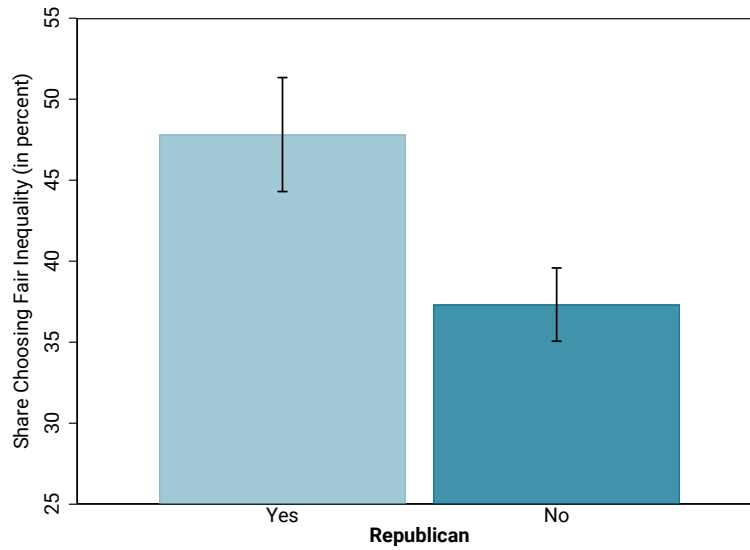
6.2 Differences in Distributional Preferences across Subgroups

We now investigate whether distributive behavior varies among subgroups of spectators. As outlined in the pre-analysis plan, our analysis focuses on the following individual characteristics: political orientation, age, gender, equivalized income, and education. Political orientation is of primary interest in the context of fairness preferences, as these preferences may influence voting behavior and political affiliation. Our first set of results therefore refers to political orientation. Thereafter, we present results for the other subgroups. In the figures, we present raw differences between subgroups. Unless otherwise noted, these differences are consistent with the coefficients estimated in Table A.3 in the appendix.

Figure 5 displays the percentage of spectators who prefer fair inequality over fair equality, categorized by whether they report to support the Republican party. It is evident that individuals who identify as Republican party supporters are significantly more inclined to favor fair inequality compared to non-supporters (48% vs. 37%, $p < 0.01$). This finding suggests a noteworthy link between political orientation and the prioritization of fair inequality over fair equality in public policy. It could potentially contribute to our understanding of differing party views on issues such as the design of transfer systems or university admission procedures.

Next, we investigate several further background characteristics. Age helps us to understand how fairness preferences may evolve in an aging society. Gender is relevant due to a growing body of literature indicating the importance of gender identity in economic decisions (e.g., Akerlof and Kranton (2000); Bénabou and Tirole (2006b); Bursztyn et al. (2017)) and the potential impact of gender inequalities on preferences (Falk and Hermle, 2018). Equivalized income is essential for assessing self-serving behavior—specifically, whether high-income individuals are more inclined to attribute their higher income to the reward for higher output (Almås et al., 2017). Since better-educated individuals also have higher earning prospects, their distributional preferences might differ (Alesina and La Ferrara, 2005). Furthermore, considering that the propensity for a political career increases with education, examining the differences in dis-

Figure 5: Higher prevalence of fair inequality among Republican party supporters



Notes: The figure depicts the percentage of spectators who choose to establish fair inequality (bars) together with 95% confidence bands (error bars) categorized by background variables.

Sample: All decisions, except 2nd decision in No Cost Treatment.

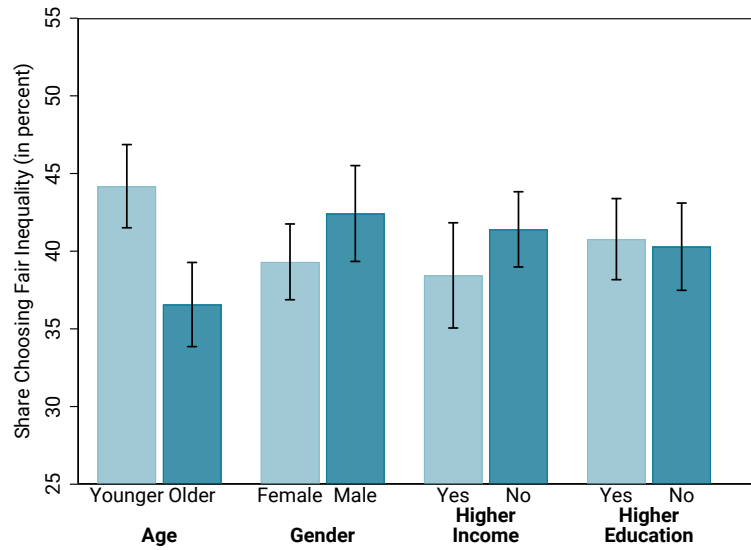
tributional preferences by level of education can shed further light on the disparities between implemented policies and the preferences of the general population (Fisman et al., 2015).

Figure 6 presents the percentage of spectators who prefer fair inequality, categorized by sub-groups. The first two bars represent individuals aged below and above the median of 49 years. Notably, 44% of younger spectators choose fair inequality, while only 37% of older spectators do so ($p < 0.01$).

In terms of gender, we observe that males are slightly more inclined to create fair inequality than females (42% vs. 39%), although this raw difference is not statistically significant ($p = 0.17$). Conditional on other covariates, the coefficient for female is negative (-0.04) and statistically significant at the 10% level, suggesting that women are slightly less likely to create fair inequality than men (Table A.3 in the appendix). However, we do not want to overemphasize the effect given its small size and weak statistical significance.

We also explore potential differences between higher- and lower-income individuals. Surprisingly, the share of lower-income spectators preferring fair inequality is slightly larger compared to that of higher-income spectators (41% vs. 38%), but confidence intervals overlap ($p = 0.222$). Finally, we examine differences with respect to education and find no statistically

Figure 6: Differences in percentage of spectators choosing fair inequality across subgroups



Notes: The figure depicts the percentage of spectators who choose to establish fair inequality (bars) together with 95% confidence bands (error bars). Indicators for age above the median of 49 years (older), gender (female), income, and completion of 4 or more years of college (higher education).

Sample: All decisions, except 2nd decision in No Cost Treatment.

significant differences for spectators with higher education (those who completed a 4-year college degree or more, $p = 0.83$).

Overall, we document significant **subgroup differences** in the propensity to create fair inequality:

Result 4a: Supporters of the Republican party and younger individuals are significantly more likely to create fair inequality, accepting unequal pay for workers with equal output.

Result 4b: There is no or little association between the likelihood of creating fair inequality and gender, income, and education, respectively.

6.3 Consistent Treatment Effects across Subgroups

Next, we turn to an analysis of how the background variables interact with the treatments (results are reported in Table OA.2 in the online appendix). We observe remarkably consistent treatment effects across subgroups: a cost for fair equality slightly increases the proportion of individuals creating fair inequality (coefficient of 0.1 with $p < 0.1$; except for females with $p = 0.173$), while a cost for fair inequality causes no statistically significant effect (in all cases $p > 0.1$, except for older ($p = 0.097$) and higher educated individuals ($p = 0.011$)).

Figure A.3 in the appendix illustrates the proportion of individuals creating fair inequality, categorized by background and treatment. While we observe a significantly weaker gradient in the Costly Inequality treatment for individuals with higher education compared to those without higher education, we regard these observations primarily as exploratory findings that need to be further investigated in future research.²¹ Furthermore, we generally find little evidence for differences between the low- and high-cost treatments across our heterogeneity dimensions (results are reported in Table OA.3 in the online appendix). We summarize the above in

Result 5: The estimated effects of the Costly Equality and Costly Inequality treatments are robust across subgroups.

6.4 Correlation between Stated and Revealed Preferences

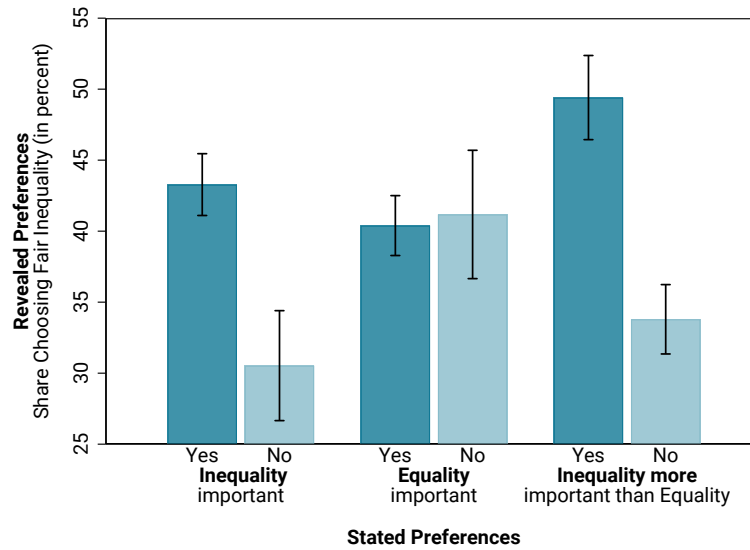
We complete the analysis by investigating whether individuals' stated preferences regarding fair equality and fair inequality, as expressed in the post-questionnaire, align with their revealed preferences in the decision problem. After the experiment, we posed three questions on distributional preferences.

The first question aimed to gauge individuals' preference for fair inequality, inquiring whether participants agree with the statement that "A society should ensure that people who contribute more earn more." About 79% of spectators agreed or strongly agreed with this statement, indicating a preference for fair inequality. Encouragingly, these spectators were also significantly more inclined to implement fair inequality in the decision problem (Figure 7, leftmost bars), with a Spearman's correlation coefficient of 0.21. This pattern persisted across all treatments (see Figure A.4).

The second question explored the importance of fair equality by asking whether spectators agree with the statement: "A society should ensure that people who contribute equally have

²¹As outlined in the pre-analysis plan, we also investigate how personal experiences of fair equality and fair inequality violations impact behavior. We do not observe any significant effects, though we acknowledge that this could be attributed to limited variation in our measure of personal experience. Specifically, the coefficient for the personal experience of a fair inequality violation is 0.005 (p -value: 0.813) and the coefficient for the personal experience of a fair equality violation is 0.029 (p -value: 0.198). It is worth noting that the inclusion of the personal experience variables does not affect the coefficients of our treatment variables. We obtain similar results when employing a stricter definition of personal experience, wherein examples provided by survey respondents were classified by a non-involved research assistant to determine whether they represented fair equality or fair inequality violations. In this case, the coefficient for personal experience of a fair equality violation is -0.032 (p -value: 0.467), and it is -0.004 for personal experience of a fair inequality violation (p -value: 0.930).

Figure 7: Correlation between stated and revealed preferences



Notes: The figure documents the correlation between stated preferences (horizontal axis) and preferences revealed in the experiment (vertical axis).

Sample: All decisions, except 2nd decision in No Cost Treatment.

equal earnings.” This statement also garners substantial support, with 82% of spectators agreeing or strongly agreeing. However, spectators who emphasize the importance of equality are not more inclined to implement fair equality in the decision problem. This holds true for the pooled sample, as shown in Figure 7, and for each of the treatments in Figure A.4 in the appendix. This result might not be surprising, considering that a majority of spectators express concerns for both fair equality and fair inequality—69% in our sample agree or strongly agree with both statements.

To assess whether individuals prioritize fair equality or fair inequality, we posed a third question in the post-questionnaire, asking spectators whether society should place greater emphasis on (i) ensuring that people who contribute more are rewarded more *or* (ii) ensuring that people who contribute equally are equally rewarded. The first statement reflects a preference for fair inequality, while the second reflects a preference for fair equality. The distribution of stated preferences closely mirrors the revealed preferences observed in our baseline. In the questionnaire, 43% of spectators state that they prioritize fair inequality (stated $\beta > 0.5$), which is numerically close to the nearly 40% of spectators implementing fair inequality in the experimental part. Notably, individuals who prioritize fair inequality in the post-questionnaire are also significantly more inclined to implement fair inequality in the decision problem (see Fig-

ure 7). As Figure A.4 illustrates, this effect is particularly pronounced in the No Cost treatment, but it is also evident in the Costly Equality and Costly Inequality treatments.²²

Although a stated preference for fair inequality over fair equality correlates positively with choosing fair inequality in the decision problem, a noteworthy portion of these spectators still choose fair equality in the decision problem. This discrepancy likely arises from conceptual differences between eliciting stated and revealed preferences, highlighting the relevance of experimental evidence. Primarily, the question regarding stated preferences encompasses society as a whole, including spectators themselves, whereas the spectator assumes the role of a third party in the decision problem. Consequently, it is unsurprising that fairness preferences differ when self-interest is at play compared to situations where it is not. This distinction in decision-making concerning oneself and others is a fundamental motivation for assessing distributive and fairness preferences through spectator designs (Cappelen et al., 2013). Secondly, individual survey responses may also depend on the perceived prevalence of violations of fair equality and of fair inequality. For example, an individual might place greater importance on fair equality but observe violations of fair inequality more frequently in daily life. In this case, the individual might be inclined to state that counteracting violations of fair inequality is more important. In our experimental design, we explicitly control for these differences in prevalence by presenting both types of violations simultaneously.

Lastly, stated preferences are ascertained through hypothetical questions, whereas preferences revealed in the decision problem entail real consequences for other individuals. Finally, directly asking for preferences regarding the equal treatment of equals may lead to socially desirable response bias.

Result 6: Stated preferences show a positive correlation with distributive behavior in the experiment, although the correlation is far from perfect. This underscores the importance of eliciting preferences in experiments where distributive choices have real consequences for others and where there is no self-serving bias.

²²We do not detect statistically significant differences for the stated preferences between the low- and high-cost treatments. For additional details, please feel free to request these results directly from the authors.

7 Conclusion

Meritocrats believe that income differences should only be driven by differences in performance, embodying two principles: rewarding equal performance equally (fair equality) and rewarding higher performance more (fair inequality). Yet, in many real-world situations, it is difficult to fully respect both principles simultaneously. Best et al. (2023) show that the perceived unfairness of public institutions can have adverse effects on behavior. To optimally design policies and institutions, it is therefore crucial to understand the importance that individuals attach to fair equality versus fair inequality. Surprisingly, this question has not been previously explored in the empirical literature on fairness preferences.

This study experimentally investigates which principle is more important to individuals from the general population in the US: fair equality or fair inequality. To do so, we presented spectators with two pairs of workers: one pair violating fair equality and the other violating fair inequality. Spectators were allowed to redistribute within one of the pairs only, enabling spectators to either restore fair equality *or* fair inequality. At the outset, we expected individuals to value both principles equally. In other words, we expected them to be indifferent when choosing between implementing fair inequality or fair equality. Yet, our research unveils that most individuals are willing to incur significant personal costs (up to 30% of their remuneration for participating in the experiment) to implement a certain principle. We interpret this as evidence of a clear preference for one principle over the other.

Our research also reveals a connection between political affiliation and the emphasis placed on fair equality versus fair inequality. Supporters of the Republican party tend to prioritize fair inequality over fair equality. This insight contributes to our understanding of the relationship between political attitudes and distributive preferences.

By explicitly modeling the trade-off, our research contributes methodologically to the experimental literature on social preferences. It enhances the ability to investigate trade-offs by offering participants a concrete choice between two aspects of fairness. This approach may have broader applications in other contexts where previous research has assessed trade-offs by comparing individuals' willingness to pay across different treatments.

References

- Abeler, J., Altmann, S., Kube, S., and Wibrals, M. (2010). Gift exchange and workers' fairness concerns: when equality is unfair. *Journal of the European Economic Association*, 8(6):1299–1324.
- Akerlof, G. and Kranton, R. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115:715–753.
- Alesina, A. and Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, 95(4):960–980.
- Alesina, A. and Guiliano, P. (2011). Preferences for redistribution. In Benhabib, J., Bisin, A., and Jackson, M. O., editors, *Handbook of Social Economics*, volume 1, pages 99–131. Elsevier, Amsterdam, the Netherlands.
- Alesina, A. and La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics*, 89(5):897–931.
- Alesina, A., Stantcheva, S., and Teso, E. (2018). Intergenerational mobility and preferences for redistribution. *American Economic Review*, 108(2):521–54.
- Almås, I., Cappelen, A. W., Kjell, S. G., Sørensen, E. Ø., and Tungodden, B. (2017). Fairness and family background. *Politics, Philosophy & Economics*, 16(2):117–131.
- Almås, I., Cappelen, A. W., and Tungodden, B. (2020). Cutthroat capitalism versus cuddly socialism: are Americans more meritocratic and efficiency-seeking than Scandinavians? *Journal of Political Economy*, 128(5):1753–1788.
- Andre, P. (2024). Shallow meritocracy: an experiment on fairness views. Technical report.
- Atkinson, A. B. and Stiglitz, J. E. (1980). *Lectures on public economics*. McGraw Hill Book Company, New York.
- Berg, K. (2021). Revealing inequality aversion from tax policy. Technical report, Oxford Working Paper 2021-18.
- Best, M. C., Caloi, L., Gerard, F., Krech, E., Naritomi, J., and Zoratto, L. (2023). Greener on the other side? Inequity and tax compliance. Technical report, mimeo.

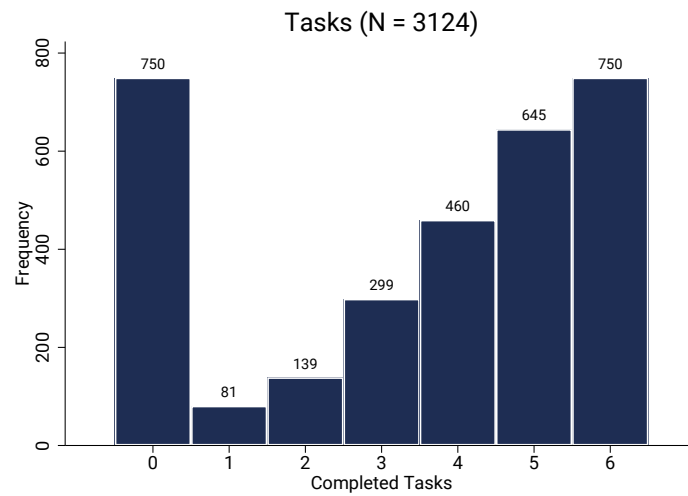
- Bhattacharya, P. and Mollerstrom, J. (2023). Luck to work. Technical report, mimeo.
- Bolton, G. E. and Ockenfels, A. (2000). ERC: a theory of equity, reciprocity, and competition. *American Economic Review*, 90(1):166–193.
- Bossert, W. (1995). Redistribution mechanisms based on individual characteristics. *Mathematical Social Sciences*, 29(1):1–17.
- Bursztyn, L., Fujiwara, T., and Pallais, A. (2017). ‘Acting wife’: marriage market incentives and labor market investments. *American Economic Review*, 107(11):3288–3319.
- Bénabou, R. and Ok, E. A. (2001). Social mobility and the demand for redistribution: the POUM hypothesis. *Quarterly Journal of Economics*, 116(2):447–487.
- Bénabou, R. and Tirole, J. (2006a). Belief in a just world and redistributive politics. *Quarterly Journal of Economics*, 121(2):699–746.
- Bénabou, R. and Tirole, J. (2006b). Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678.
- Cappelen, A. W., Cappelen, C. W., and Tungodden, B. (2023). Second-best fairness: the trade-off between false positives and false negatives. *Journal of Political Economy*.
- Cappelen, A. W., Falch, R., and Tungodden, B. (2020). Fair and unfair income inequality. *Handbook of Labor, Human Resources and Population Economics*, pages 1–25.
- Cappelen, A. W., Hole, A. D., Sørensen, E. Ø., and Tungodden, B. (2007). The pluralism of fairness ideals: an experimental approach. *American Economic Review*, 97(3):818–827.
- Cappelen, A. W., Konow, J., Sørensen, E. Ø., and Tungodden, B. (2013). Just luck: an experimental study of risk-taking and fairness. *American Economic Review*, 103(4):1398–1413.
- Cappelen, A. W., Mollerstrom, J., Reme, B.-A., and Tungodden, B. (2022). A meritocratic origin of egalitarian behaviour. *The Economic Journal*, 132(646):2101–2117.
- Cappelen, A. W. and Tungodden, B. (2019). *The economics of fairness*. Edward Elgar Publishing.
- Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3):817–869.

- Cherry, T. L., Frykblom, P., and Shogren, J. F. (2002). Hardnose the dictator. *American Economic Review*, 92(4):1218–1221.
- Engelmann, D. and Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review*, 94(4):857–869.
- Falk, A. and Hermle, J. (2018). Relationship of gender differences in preferences to economic development and gender equality. *Science*, 362(6412):eaas9899.
- Falk, A. and Zimmermann, F. (2013). A taste for consistency and survey response behavior. *CESifo Economic Studies*, 59(1):181–193.
- Fehr, D., Mollerstrom, J., and Perez-Truglia, R. (2022a). Your place in the world: relative income and global inequality. *American Economic Journal: Economic Policy*, 14(4):232–268.
- Fehr, E., Epper, T., and Senn, J. (2022b). Other-regarding preferences and redistributive politics. Working Paper 9545, CESifo.
- Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868.
- Fisman, R., Jakiela, P., and Kariv, S. (2017). Distributional preferences and political behavior. *Journal of Public Economics*, 155:1–10.
- Fisman, R., Jakiela, P., Kariv, S., and Markovits, D. (2015). The distributional preferences of an elite. *Science*, 349(6254):aab0096.
- Fleurbaey, M. (2008). *Fairness, responsibility, and welfare*. Oxford University Press, Oxford.
- Fleurbaey, M. and Maniquet, F. (2011). *A theory of fairness and social welfare*. Cambridge University Press, Cambridge, UK.
- Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics*, 82(2):225–246.
- Hvidberg, K. B., Kreiner, C. T., and Stantcheva, S. (2023). Social positions and fairness views on inequality. *Review of Economic Studies*, 90(6):3083–3118.

- Kaplow, L. (1989). Horizontal equity: measures in search of a principle. *National Tax Journal*, 42(2):139–154.
- Konow, J. (2000). Fair shares: accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 90(4):1072–1091.
- Konow, J. (2009). Is fairness in the eye of the beholder? An impartial spectator analysis of justice. *Social Choice and Welfare*, 33(1):101–127.
- Krawczyk, M. (2010). A glimpse through the veil of ignorance: equality of opportunity and support for redistribution. *Journal of Public Economics*, 94(1-2):131–141.
- Mollerstrom, J., Reme, B.-A., and Sørensen, E. Ø. (2015). Luck, choice and responsibility—an experimental study of fairness views. *Journal of Public Economics*, 131:33–40.
- Musgrave, R. A. (1990). Horizontal equity, once more. *National Tax Journal*, 43(2):113–122.
- Musgrave, R. A. (1993). Horizontal equity: a further note. *Florida Tax Review*, 1(6):354–359.
- Roemer, J. and Trannoy, A. (2015). Equality of opportunity. In Atkinson, A. B. and Bourguignon, F., editors, *Handbook of Income Distribution*, pages 217–300. Elsevier, Amsterdam, the Netherlands.
- Roth, C. and Wohlfahrt, J. (2018). Experienced inequality and preferences for redistribution. *Journal of Public Economics*, 167:251–262.

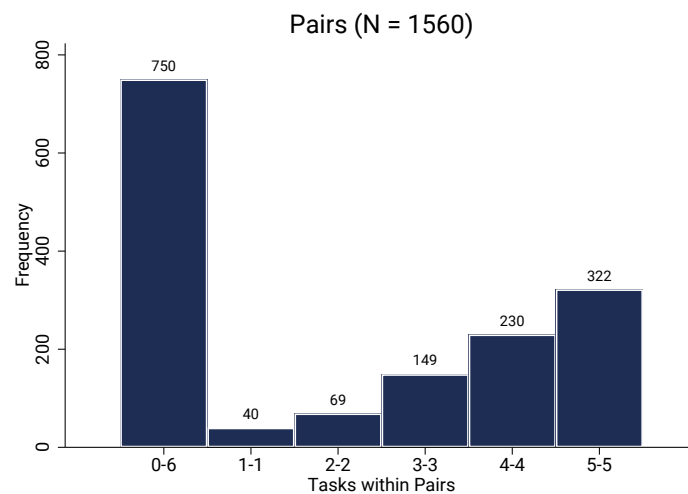
Appendix

Figure A.1: Distribution of number of tasks completed (all assignments pooled)



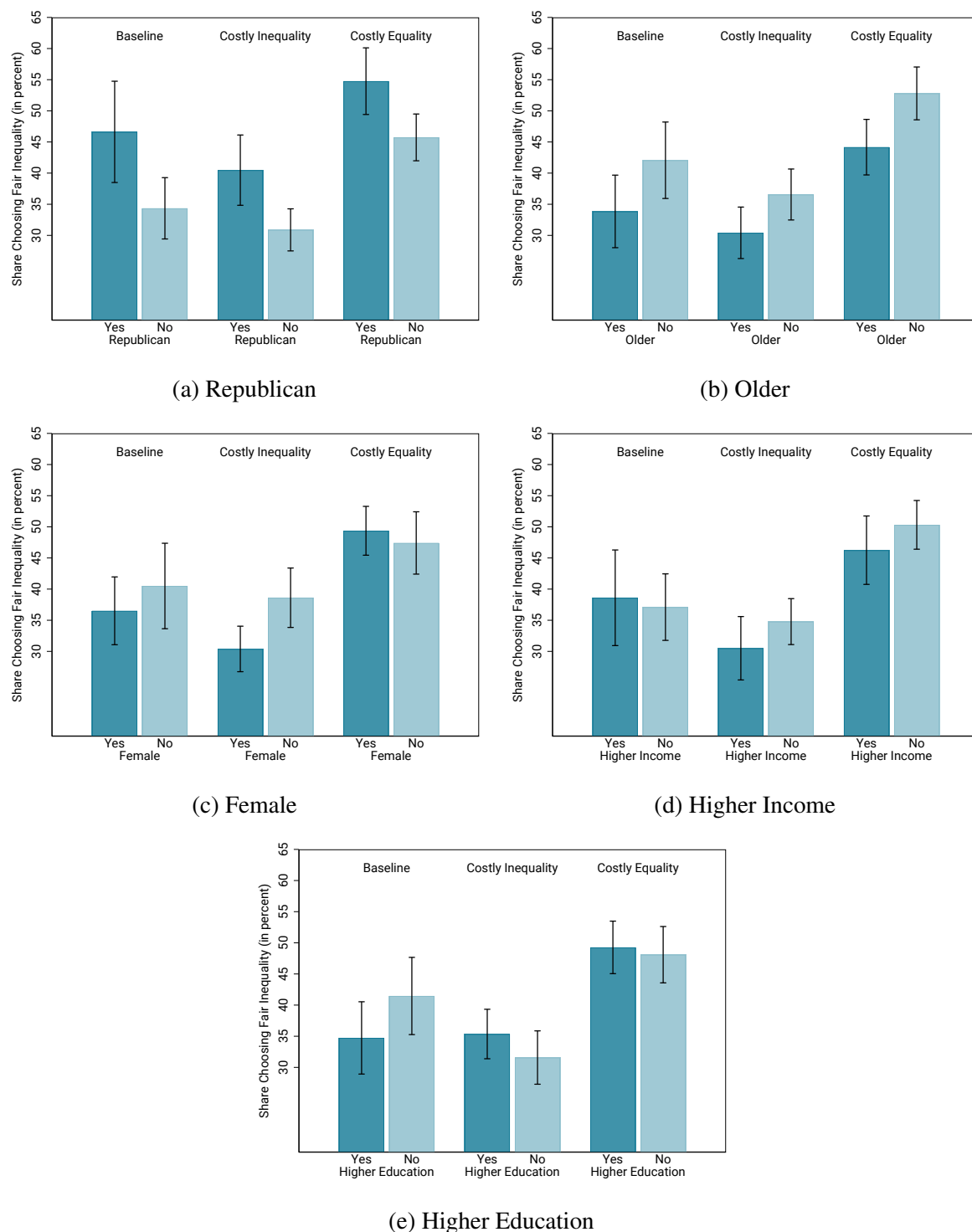
Notes: The figure shows the distribution of the number of tasks that workers completed (all assignments pooled). Ten assignments with 5 tasks completed were upgraded to having completed 6 tasks.

Figure A.2: Distribution of worker pairs (all assignments pooled)



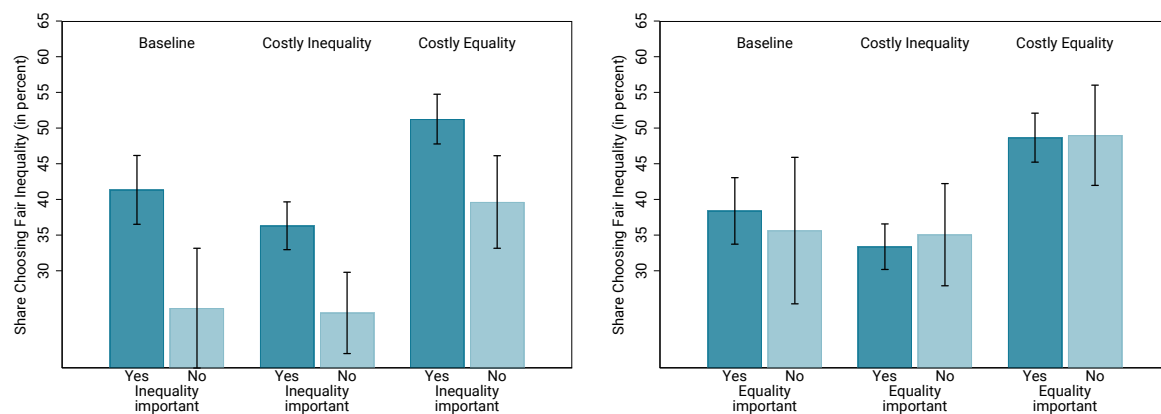
Notes: Four assignments remained unused in the formation of worker pairs (one with 1 task, 2 tasks, 3 tasks and 5 tasks completed, respectively). No bonus was issued for these four assignments.

Figure A.3: Heterogeneity in the propensity to create fair inequality persists across treatments



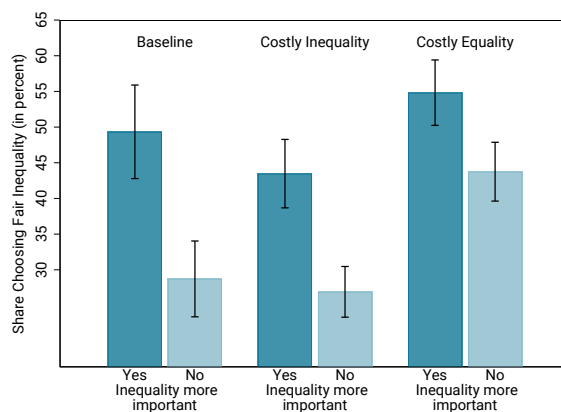
Notes: The figure depicts the percentage of spectators who choose to establish fair inequality (bars), categorized by treatment, together with 95% confidence bands (error bars). It documents that the large differences in distributional preferences across subgroups observed within the pooled sample (Figures 5 and 6) persist across treatments. *Sample:* All decisions from Costly Inequality and Costly Equality treatments, first decisions from No Cost Treatment.

Figure A.4: Correlation pattern btw. stated and revealed preferences persists across treatments



(a) Stated preference: Inequality important

(b) Stated preference: Equality important



(c) Stated preference: Inequality more important than Equality

Notes: The figure depicts the percentage of spectators who choose to establish fair inequality (bars), categorized by treatment, together with 95% confidence bands (error bars). It documents the positive correlation between stated preferences (horizontal axis) and preferences revealed in the experiment (vertical axis).

Sample: All decisions from Costly Inequality and Costly Equality treatments, first decisions from No Cost Treatment.

Table A.1: Our sample compared to the US population: descriptives

	Sample	US Population
Female	0.608	0.513
Age		
18 to 39	0.334	0.383
40 to 59	0.339	0.323
60 to 69	0.171	0.150
above 70	0.155	0.144
Marital status		
Never-married	0.260	0.339
Married	0.586	0.476
Legally separated or divorced	0.117	0.128
Widowed	0.037	0.057
Ethnicity		
European American/White	0.833	0.616
African American/Black	0.062	0.124
Hispanic/Latino	0.035	0.187
Asian/Asian American	0.041	0.060
American Indian / Alaska Native / Mixed race	0.024	0.013
Education		
Less than high school	0.014	0.114
High school graduate	0.161	0.269
Some college	0.190	0.200
College degree	0.422	0.289
Professional or graduate	0.209	0.128
Income		
Less than \$15,000	0.179	0.098
\$15,000 to \$24,999	0.186	0.083
\$25,000 to \$49,999	0.303	0.203
\$50,000 to \$99,999	0.244	0.302
\$100,000 to \$199,999	0.086	0.229
More than \$200,000	0.003	0.085

Notes: This table contains summary statistics for our sample of spectators alongside nationally representative statistics from the Census Bureau and the Current Population Survey. Education: highest level of education. Income: Equivalized income calculated with the same method used in the US census.

Table A.2: Regression results robust to weighted sample matching population values

	Unweighted	Weighted by				
	(1)	(2) Gender	(3) Age	(4) Ethnicity	(5) Education	(6) Income
Costly Inequality	-0.047* (0.028)	-0.045 (0.028)	-0.056* (0.029)	-0.052 (0.037)	-0.065 (0.040)	-0.089 (0.054)
Costly Equality	0.104*** (0.029)	0.101*** (0.029)	0.099*** (0.030)	0.103*** (0.039)	0.099** (0.039)	0.150** (0.059)
Control Group Mean	0.424*** (0.033)	0.423*** (0.033)	0.422*** (0.035)	0.403*** (0.045)	0.429*** (0.046)	0.352*** (0.064)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2385	2385	2385	2314	2374	2385
R^2	0.041	0.040	0.039	0.040	0.045	0.076

Sample: All decisions, except 2nd decision in No Cost Treatment.

Notes: The No Cost treatment is the reference category; the constant reflects the baseline share of individuals choosing fair inequality. Indicators for being female, above median age, above median income, higher education and being a republican. Observations are reweighted to represent population values in terms of gender, age, ethnicity, education, and income dimensions. Clustered standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Regression results and marginal effects on created fair inequality

	Dependent Variable: Creating Fair Inequality			
	(1) LPM	(2) Probit	(3) LPM	(4) Probit
Costly Inequality	-0.043 (0.028)	-0.043 (0.028)	-0.047* (0.028)	-0.047* (0.028)
Costly Equality	0.108*** (0.028)	0.108*** (0.028)	0.104*** (0.029)	0.104*** (0.029)
Republican			0.113*** (0.024)	0.112*** (0.023)
Older			-0.094*** (0.023)	-0.095*** (0.023)
Female			-0.050** (0.023)	-0.050** (0.023)
Higher Income			-0.013 (0.026)	-0.013 (0.026)
Higher Education			0.004 (0.024)	0.003 (0.023)
Control Group Mean	0.379*** (0.022)		0.424*** (0.033)	
Observations	2541	2541	2385	2385
R^2	0.020		0.041	

Sample: All decisions, except the 2nd decisions in the No Cost Treatment.

Notes: LPM stands for Linear Probability model. The table displays marginal effects for the Probit model. The No Cost treatment is the reference category; the constant reflects the baseline share of individuals choosing fair inequality. Indicators include gender, age above median, equivalized income above median, higher education (4-year college degree or more), and Republican affiliation. Clustered standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Regression results and marginal effects on created fair inequality, 1st decisions only

	Dependent Variable: Creating Fair Inequality			
	(1) LPM	(2) Probit	(3) LPM	(4) Probit
Costly Inequality	-0.033 (0.030)	-0.033 (0.030)	-0.034 (0.031)	-0.034 (0.031)
Costly Equality	0.127*** (0.031)	0.127*** (0.031)	0.129*** (0.032)	0.129*** (0.032)
Republican			0.110*** (0.028)	0.109*** (0.027)
Older			-0.119*** (0.027)	-0.119*** (0.026)
Female			-0.068** (0.027)	-0.068*** (0.026)
Higher Income			0.009 (0.030)	0.009 (0.030)
Higher Education			-0.000 (0.027)	-0.000 (0.027)
Control Group Mean	0.379*** (0.022)		0.443*** (0.036)	
Observations	1525	1525	1429	1429
R^2	0.020		0.047	

Sample: First decision only.

Notes: LPM refers to a Linear Probability model. For the Probit model, the table shows marginal effects. Control Group Mean refers to the No Cost treatment as the reference category. Indicators for being female, above median age, above median equivalized income, higher education (4-year college degree or more) and being a republican. Clustered standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Regression results on created fair inequality, by size of personal cost for spectators

	Dependent Variable: Creating Fair Inequality			
	(1) LPM	(2) Probit	(3) LPM	(4) Probit
High Cost Inequality	-0.027 (0.030)	-0.027 (0.030)	-0.031 (0.031)	-0.031 (0.031)
Low Cost Inequality	-0.058* (0.030)	-0.058* (0.030)	-0.063** (0.030)	-0.063** (0.030)
High Cost Equality	0.111*** (0.031)	0.111*** (0.031)	0.106*** (0.032)	0.106*** (0.032)
Low Cost Equality	0.105*** (0.031)	0.105*** (0.031)	0.102*** (0.032)	0.102*** (0.032)
Republican			0.113*** (0.024)	0.112*** (0.023)
Older			-0.094*** (0.023)	-0.095*** (0.023)
Female			-0.050** (0.023)	-0.050** (0.023)
Higher Income			-0.013 (0.026)	-0.013 (0.026)
Higher Education			0.004 (0.024)	0.003 (0.023)
Control Group Mean	0.379*** (0.022)		0.424*** (0.033)	
Observations	2541	2541	2385	2385
R^2	0.020		0.041	

Sample: All decisions, except 2nd decision in No Cost Treatment.

Notes: The No Cost treatment is the reference category; the constant reflects the baseline share of individuals choosing fair inequality. Indicators for being female, above median age, above median income, higher education and being a republican. Clustered standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

ONLINE APPENDIX

Not for publication.

A Treatments Balanced

Table OA.1: Balancing table

	Mean				P-value of t-test		
	Total	NC	CE	CI	NC vs. CE	NC vs. CI	CE vs. CI
Republican	0.30 (0.01)	0.29 (0.02)	0.33 (0.02)	0.29 (0.02)	0.184	0.977	0.175
Older	0.48 (0.01)	0.50 (0.02)	0.47 (0.02)	0.47 (0.02)	0.286	0.286	1.000
Female	0.61 (0.01)	0.60 (0.02)	0.62 (0.02)	0.60 (0.02)	0.571	0.959	0.536
Higher Income	0.33 (0.01)	0.33 (0.02)	0.34 (0.02)	0.33 (0.02)	0.873	0.982	0.855
Higher Education	0.53 (0.01)	0.52 (0.02)	0.54 (0.02)	0.55 (0.02)	0.509	0.244	0.614

Mean (standard error) and p-value of t-test. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Treatment Heterogeneity

Table OA.2: Heterogeneity results: personal characteristics

	Var				
	Republican	Older	Female	Higher Inc.	Higher Educ.
Costly Inequality	-0.041 (0.033)	-0.068* (0.041)	-0.019 (0.045)	-0.032 (0.034)	-0.103** (0.041)
Costly Inequality X Var	-0.020 (0.063)	0.043 (0.056)	-0.048 (0.057)	-0.046 (0.060)	0.105* (0.056)
Costly Equality	0.113*** (0.035)	0.101** (0.042)	0.063 (0.046)	0.119*** (0.036)	0.072* (0.042)
Costly Equality X Var	-0.030 (0.063)	0.005 (0.058)	0.068 (0.059)	-0.045 (0.061)	0.062 (0.058)
Var	0.134*** (0.050)	-0.114** (0.045)	-0.057 (0.045)	0.023 (0.048)	-0.063 (0.045)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	2385	2385	2385	2385	2385
R^2	0.041	0.041	0.044	0.041	0.042

Sample: All decisions, except the 2nd decisions in the No Cost Treatment.

Notes: The No Cost treatment is the reference category. Indicators include gender, age above median, equivalized income above median, higher education (4-year college degree or more), and Republican affiliation. Clustered standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C No Impact of Size of Cost

To investigate whether the size of cost impacts behavior, we estimate Equation (OA.1). Results are presented in Tables OA.3.

$$\begin{aligned}
D_{i,d} = & \alpha + \gamma^Z Z_i + \delta^{EL} CostEqual_{i,d}^L + \delta^{IL} CostInequal_{i,d}^L \\
& + \delta^{ELZ} CostEqual_{i,d}^L \times Z_i + \delta^{ILZ} CostInequal_{i,d}^L \times Z_i \\
& + \delta^{EH} CostEqual_{i,d}^H + \delta^{IH} CostInequal_{i,d}^H \\
& + \delta^{EHZ} CostEqual_{i,d}^H \times Z_i + \delta^{IHZ} CostInequal_{i,d}^H \times Z_i \\
& + \beta \mathbf{X}_i + \varepsilon_{i,d}
\end{aligned} \tag{OA.1}$$

Table OA.3: Heterogeneity results with respect to size of cost: personal characteristics

	Var				
	Republican	Older	Female	Higher Inc.	Higher Educ.
High Cost Equality	0.113*** (0.038)	0.097** (0.046)	0.044 (0.051)	0.110*** (0.039)	0.061 (0.046)
High Cost Equality X Var	-0.024 (0.070)	0.018 (0.064)	0.102 (0.065)	-0.010 (0.067)	0.087 (0.064)
Low Cost Equality	0.113*** (0.038)	0.105** (0.046)	0.082 (0.051)	0.129*** (0.039)	0.083* (0.047)
Low Cost Equality X Var	-0.036 (0.069)	-0.008 (0.064)	0.034 (0.065)	-0.079 (0.067)	0.037 (0.064)
High Cost Inequality	-0.013 (0.036)	-0.068 (0.045)	0.027 (0.050)	-0.019 (0.038)	-0.080* (0.045)
High Cost Inequality X Var	-0.062 (0.069)	0.076 (0.062)	-0.099 (0.063)	-0.036 (0.066)	0.092 (0.062)
Low Cost Inequality	-0.069* (0.035)	-0.068 (0.045)	-0.064 (0.048)	-0.044 (0.037)	-0.126*** (0.044)
Low Cost Inequality X Var	0.022 (0.069)	0.010 (0.061)	0.003 (0.062)	-0.056 (0.064)	0.119* (0.061)
Var	0.134*** (0.050)	-0.114** (0.045)	-0.057 (0.045)	0.023 (0.048)	-0.063 (0.045)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	2385	2385	2385	2385	2385
R ²	0.042	0.042	0.045	0.042	0.043

Sample: All decisions, except the 2nd decisions in the No Cost Treatment.

Notes: The No Cost treatment is the reference category. Indicators include gender, age above median, equalized income above median, higher education (4-year college degree or more), and Republican affiliation. Clustered standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D Instructions for Spectators

Redistribution decision displayed below. Complete instructions can be accessed [via this link](#).

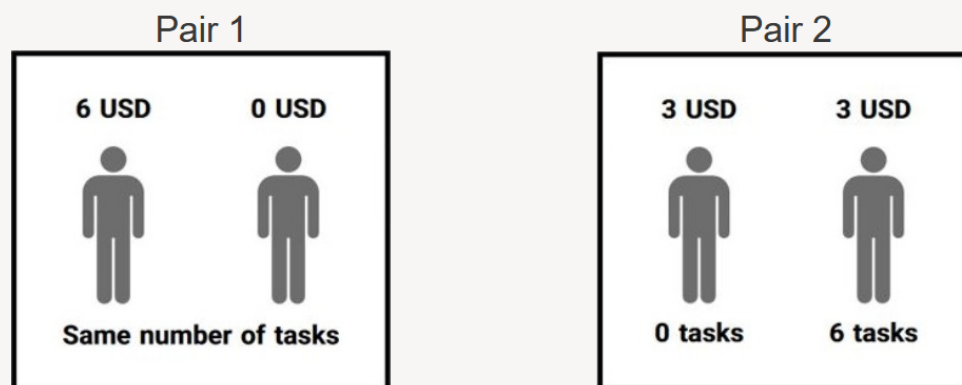
Below, we present the first of the two situations.

Individuals were recruited via an international online labor market and were paid a participation fee of 2 USD. They had the opportunity to complete standardized tasks with the prospect of receiving additional earnings.

Each individual was paired with another one, and total earnings of 6 USD were divided between them. The individuals were not informed how the total earnings were distributed. Below, you see two of the pairs:

In Pair 1, both individuals completed the **same number of tasks** but were given **different earnings**.

In Pair 2, both individuals completed a **different number of tasks** but were given the **same earnings**.



Individuals were told that a third person would be given the opportunity to redistribute their total earnings and thus determine their final payments. You are the third person and can redistribute within **one of the two pairs**.

You can choose between the following two alternatives:

Please indicate what you want to choose by ticking one of the two alternatives below. Your decision is completely anonymous. Individuals will not receive any further information. One of the two pairs will be drawn at random and the individuals in this pair will receive their final payments within a few days.

- ☐ Redistribute within Pair 1, so that both are paid 3 USD. The earnings in Pair 2 remain unchanged. If you choose this alternative, you will receive the equivalent of 1.4 USD for completing this survey.
- ☐ Redistribute within Pair 2, so that the individual with 6 tasks is paid 6 USD (and the individual with 0 tasks is paid 0 USD). The earnings in Pair 1 remain unchanged. If you choose this alternative, you will receive the equivalent of 2 USD for completing this survey.

E Instructions for Workers

Instructions Workers

Please read the instructions below carefully.

General instructions:

The results from this study will be used in a research project. It is therefore important that you carefully read and follow all instructions. Note that you will remain anonymous throughout the study. We will only use your Participant ID to assign payments and check that you have not participated in this study before.

You will be paid a fixed participation fee of **2 USD** and you may, depending on the actions you and others take, earn additional money.

You will be given detailed instructions on your screen before each part of the study. Please read the instructions to each part carefully.

If you have any questions regarding this study, you may contact **survey@cmi.no**

I have read and understood the the above and want to participate in this study:

☐ Yes

☐ No

>>

Production phase

The first part of the study is a production phase in which you get the opportunity to work on 4 assignments. In each assignment, you are asked to solve 6 tasks that each take 1 minute.

You can **choose between working and not working** at the start of **each assignment**.

Your performance will be measured by the number of tasks you complete in each assignment.

Go on to the next page to receive instructions for the tasks.



Assignment 1

In the first assignment you are asked to work on six code recognition tasks.

In each task, you will get a range of numbers and you should find and check off all numbers within that range from a matrix. There will be multiple numbers to be checked off in the same matrix.

You have successfully completed the task if you reach the maximum number of points. This is the case if you **correctly tick off all numbers in the range without ticking off any wrong numbers**. After each task you will be informed whether or not you reached the maximum number of points (e.g. 10/10).

Your matrix will be submitted automatically after 60 seconds. Each assignment will last for 6 minutes, and thereafter you will be taken to the next assignment.

Below you are shown a simplified example to make sure you understand the task. When you have read and understood the instructions press >> to advance to the next screen.

This is an example:

You must check the codes in the interval between 100 and 149 (including interval borders).

<input type="checkbox"/> 123	<input type="checkbox"/> 283
<input type="checkbox"/> 231	<input type="checkbox"/> 149
<input type="checkbox"/> 952	<input type="checkbox"/> 641
<input type="checkbox"/> 864	<input type="checkbox"/> 820
<input type="checkbox"/> 333	<input type="checkbox"/> 462
<input type="checkbox"/> 791	<input type="checkbox"/> 940

Do you want to work on this assignment for the next six minutes?

☐ Yes

☐ No

>>

0051

You must check the codes in the interval between **157** and **246** (including interval limits).

<input type="checkbox"/> 272	<input type="checkbox"/> 021	<input type="checkbox"/> 481	<input type="checkbox"/> 391
<input type="checkbox"/> 209	<input type="checkbox"/> 817	<input type="checkbox"/> 871	<input type="checkbox"/> 759
<input type="checkbox"/> 879	<input type="checkbox"/> 489	<input type="checkbox"/> 672	<input type="checkbox"/> 027
<input type="checkbox"/> 719	<input type="checkbox"/> 428	<input type="checkbox"/> 402	<input type="checkbox"/> 514
<input type="checkbox"/> 911	<input type="checkbox"/> 307	<input type="checkbox"/> 430	<input type="checkbox"/> 485
<input type="checkbox"/> 014	<input type="checkbox"/> 295	<input type="checkbox"/> 576	<input type="checkbox"/> 043
<input type="checkbox"/> 581	<input type="checkbox"/> 506	<input type="checkbox"/> 591	<input type="checkbox"/> 884
<input type="checkbox"/> 917	<input type="checkbox"/> 799	<input type="checkbox"/> 008	<input type="checkbox"/> 851
<input type="checkbox"/> 693	<input type="checkbox"/> 586	<input type="checkbox"/> 387	<input type="checkbox"/> 177
<input type="checkbox"/> 581	<input type="checkbox"/> 418	<input type="checkbox"/> 215	<input type="checkbox"/> 462
<input type="checkbox"/> 101	<input type="checkbox"/> 784	<input type="checkbox"/> 165	<input type="checkbox"/> 789
<input type="checkbox"/> 060	<input type="checkbox"/> 478	<input type="checkbox"/> 824	<input type="checkbox"/> 721

Determination of payments

You have completed your work. Your payment for this work is determined by a two-stage process.

First stage:

After you have completed this study, we will, for each assignment, pair you with another participant that has

- (i) completed the same number of tasks or
- (ii) completed a different number of tasks.

Then, earnings of 6 USD will be tentatively divided between you and the other participant.

Second stage:

For each assignment, a third person will be given the opportunity to redistribute the earnings between you and the other participant. This person will not know your or the other participant's identity but will be informed about the number of tasks you and the other participant have completed in the assignment.

Your payment for each assignment will be either 0 USD, 3 USD or 6 USD.

You will receive your payments for the four assignments within a few weeks and it will be paid separately from your fixed participation fee of 2 USD.

Please click >> to continue.

>>