# Essays in Applied Microeconomics

Inauguraldissertation

zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften

durch

die Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

**Timo Freyer**

aus Waldbröl

2023

| | |
|---|---|
| Dekan: | Prof. Dr. Jürgen von Hagen |
| Erstreferent: | Prof. Dr. Sebastian Kube |
| Zweitreferent: | Prof. Dr. Thomas Dohmen |
| Tag der mündlichen Prüfung: | 25. September 2023 |

# Acknowledgements

In this world nothing can be said to be certain, except that I was incredibly lucky to have the support of many capable and cordial people, and that this support was essential for the development of this thesis.

First and foremost, I would like to thank my supervisors Sebastian Kube and Thomas Dohmen, who not only provided excellent input and feedback on paper drafts, projects, experimental designs, and stupid research ideas, but also were extraordinarily accessible and role models in many regards. I believe it is not common that a supervisor takes the time for an emergency meeting on data collection between Christmas and New Year's Eve. Besides my supervisors, special thanks go to Florian Zimmermann and Matthias Kräkel for many feedback sessions, a lot of advice, and being particularly supportive. I am thankful to and was incredibly lucky with my coauthors, with whom working together was mostly pleasant, sometimes exhausting, but always the right decision. Further, I am indebted to Holger Gerhardt for providing various public goods, a sharp eye on experimental instructions and typography, being a welcoming office mate, and the occasional dime.

I am grateful to the whole Uni Bonn econ faculty and everyone at IAME, especially Simone Jost and Ilona Krupp, who have the rare talent to magically make things work. I am grateful for the opportunity to study alongside a fantastic group of BGSE students, and for being able to call many of them close friends today. In that regard, I am particularly grateful to Luca, Martin, Matthias, Chui Yee, Ximeng, Christian, Laurenz, and Marton for memorable trips and vacations, (sometimes overly) competitive Doppelkopf and table tennis sessions, inspiring (and sometimes heated) debates, delicious dinners, and generally for making the Ph.D. a mostly enjoyable episode of my life. Finally, I am grateful to all my friends and family who accompanied me on this journey. Thank you all!

Timo Freyer
Bonn, September 2023

# Contents

# List of Figures

# List of Tables

# Introduction

The concept of preferences is widely used in economics and related disciplines to put structure on individual decision making and behaviour. By formalizing how individuals make decisions based on their preferences, researchers are able to explain behavioural regularities within and heterogeneity across individuals and environments. Provided that it yields a better assessment of how individuals will react to potential policy interventions or behave under various institutions, a thorough understanding of individuals' preferences also helps policy-makers to implement those policies that are likely to have the desired effects and safeguards against those ones that have unintended consequences. Moreover, policy debates often revolve around normative questions such as whether the distributional consequences of a given intervention are desirable or not. In that regard, knowledge of individuals' preferences facilitates the design of policies that reflect commonly held values.

This thesis consists of three essays that advance our knowledge of the distribution of individual preferences, how they shape social interactions and translate into collective outcomes, and how they interact with the decision environment to determine individual behaviour. Thereby, the thesis contributes to our understanding of human behaviour and may help to develop effective and widely accepted policies.

One topic that has sparked heated normative debates in the past and remains a root of contemporary political disagreement is distributive justice, i.e., the perceived fairness of resource allocations. Given that in many contemporary societies the majority of individuals subscribes to the same meritocratic fairness ideal, which holds that resource distributions should reflect factors within but not outside individuals' control, it is striking that this disagreement persists. In Chapter 1, titled "Inherited Inequality and the Dilemma of Meritocracy" (joint work with Laurenz Günther), we argue that one potential source of disagreement originates from a fundamental tension in the meritocratic logic. As human beings routinely invest efforts and resources to the benefit of others—but tend to be more altruistic toward their family members, friends, and compatriots than toward non-relatives, strangers and foreigners—individuals are often not responsible for their outcomes themselves but, to a differential extent, benefit from the efforts of others. By meritocratic standards, the resulting *inherited inequality* is just and unjust at the same time because it reflects factors within the benefactors' but outside the beneficiaries' control. Hence, inherited inequality confronts meritocrats with a moral dilemma because redistribution decisions require them to necessarily violate their fairness preferences.

We run a survey experiment with a representative sample of US citizens to investigate how people deal with this dilemma. In the experiment, impartial spectators who have no stakes in the situation redistribute payments between pairs of individuals. We vary a) whether the initial payment distribution is based on a random draw or on relative effort and b) whether spectators redistribute between individuals who have worked themselves or who merely benefit from the work of real-life friends. Redistribution levels are substantially higher if inequality is based on luck instead of effort. However, whether individuals worked themselves or inherited their initial payoffs does not matter much for spectators' redistribution decisions. These results suggest that many US citizens accept inherited inequality as long as it is merited at some stage, which may explain why many people oppose redistributive policies.

Individual preferences are also relevant for collective action problems such as the fight against climate change or the containment of the recent COVID-19 pandemic. Because reducing one's carbon footprint or engaging in preventive health behaviors is costly for individuals but generates benefits that also accrue to society at large, collective action problems constitute social dilemmas. In Chapter 2, titled "Prosociality predicts individual behavior and collective outcomes in the COVID-19 pandemic" (joint work with Ximeng Fang, Chui-Yee Ho, Zihua Chen, and Lorenz Götte), we examine the relationship between prosociality and individual health behavior as well as collective health outcomes and provide empirical support for this conjecture.

We conduct a nationally representative online survey in Germany to investigate the role of prosociality in reducing the spread of COVID-19 during the second coronavirus wave. At the individual level, higher prosociality is strongly positively related to compliance with public health behaviors such as mask wearing and social distancing. At the regional level, a higher average prosociality is associated with significantly lower weekly incidence and case growth rates, controlling for a host of demographic and socio-economic factors. These associations are driven by higher compliance with public health behaviors in regions with higher prosociality. Our correlational results thus support the common notion that voluntary behavioral change plays a vital role in fighting the pandemic and, more generally, that social preferences may determine collective action outcomes of a society.

Finally, Chapter 3, titled "The Effect of Task (Mis)Matching and Self-Selection on Intrinsic Motivation and Performance" (joint work with Jonas Radbruch and Sebastian Schaube), is concerned with work environments and how they interact with individual preferences over features of that environment. Because different work environments cater to different workers' strengths or preferences, placing workers in a suitable environment is important to help them succeed on their job. We focus on a core aspect of work environments—namely, which task individuals work on—and study whether the (mis)match between tasks and workers' task preferences affects their performance, and whether there is a direct motivational effect of self-selection.

To answer these questions, we conduct an online experiment where subjects work on one of two real effort tasks. We exogenously vary whether they are assigned their preferred or non-preferred task, or whether they can actively self-select a task. The results show that subjects who either self-select a task or are assigned their preferred task produce about 50% more output than subjects who are assigned their non-preferred task. This effect can be attributed to both higher productivity and an increase in the time subjects spend working on their task. Evidence from the post-experimental questionnaire indicates that the increase in performance is not exclusively due to ability-sorting but in part driven by increased effort due to higher intrinsic motivation. In summary, the results suggest that workers' performance depends crucially on whether they work on their preferred task, but not so much on whether that task is self-selected or assigned.

# Chapter 1

# Inherited Inequality and the Dilemma of Meritocracy*

*Joint with Laurenz Günther*

## 1.1 Introduction

In a meritocratic society, inequality is considered to be just if it reflects factors within but not outside individuals' control. However, individuals are often not responsible for their outcomes themselves but benefit differentially from the efforts of others. For example, a child may be lucky to inherit abundant resources acquired by its parents, while another child is born into less favourable circumstances. Such *inherited inequality*[1] exposes a fundamental tension in the meritocratic logic. On the one hand, individuals are entitled to decide how to spend their earned resources, which includes the right to transfer them to others. On the other hand, if two individuals are not involved in the process that generates inequality between them, such inequality does not reflect their individual achievements. In the parent-child example, if one pair of parents works particularly hard such that

* We thank Johannes Goldbeck for excellent research assistance. We thank Thomas Dohmen, Sebastian Kube, and Florian Zimmermann for their guidance and support. We thank Johannes Abeler, Peter Andre, Christian Apenbrink, Roland Bénabou, Alexander Cappelen, Armin Falk, Simon Gächter, Luca Henkel, Chui-Yee Ho, Thomas Kohler, Michael Kosfeld, George Loewenstein, Anna Schulze-Tilling, Louis Strang, as well as audiences at ESA Bologna, NSEF PhD & Postdoc Workshop Naples, Fairness and the Moral Mind Virtual Workshop, and IAME Applied Micro Coffee Bonn for very helpful comments. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC2126/1-390838866.

1. With *inherited inequality* we refer to inequality between individuals that originates from the actions of others. Hence, we interpret the term "inherited" broadly. Our definition encompasses inequality between children who profit differentially from the actions of their parents, but also inequality between people who benefit to a differential extent from their friends, coworkers, or compatriots.

their children “can have a better life”, they have merited to see their child reap the benefits of their efforts. However one child has not merited more favourable circumstances than the other. By meritocratic standards, inherited inequality is just and unjust at the same time and confronts meritocrats with a dilemma—*the dilemma of meritocracy*.

The dilemma of meritocracy is central to various policy debates. Consider as an example the debate on the estate tax. Here, people who seemingly endorse the same fairness ideal—the meritocratic one—can end up taking diametrically opposed positions. Some contend that bequests are a result of the testator’s hard work and usually conclude that it is unfair to redistribute. In this vein, it has been argued that “[s]ince the accumulation of a substantial estate is one of the motivations that drive people to work hard, a death tax on saving is indirectly a tax on work” (Posner, 1972). Other people stress that it was certainly not the heir’s efforts that generated the bequest and label inheritances as unmerited income, concluding that it should be heavily taxed. For instance, US investor Warren Buffet is quoted in Obama (2006) saying that “[w]hen you get rid of the estate tax, you’re basically handing over command of the country’s resources to people who didn’t earn it”. Hence, the meritocratic fairness ideal is being used as a justification for policies at opposite ends of the political spectrum.

A potential explanation for this disagreement is that people differ in whether they prioritize meritocratic fairness toward the benefactors or the beneficiaries. Different priorities may, in turn, translate into different views on policies and demand for redistribution in the context of inherited inequality. To develop policies that are politically implementable and meet the fairness preferences of citizens, it is therefore necessary to better understand people’s attitudes toward economic inequality and, in particular, how they deal with the dilemma of meritocracy.

In pursuit of this aim, this study introduces a stylized theoretical framework that formalizes how individuals evaluate (inherited) unequal distributions and reports results from a survey experiment that puts its predictions to the test. The framework covers situations in which money is distributed between two individuals who each benefit from the effort of an associated worker. An impartial spectator observes this situation and makes a fairness judgment based on his or her fairness ideal. This setup nests the case of noninherited inequality, where a beneficiary and the associated worker are identical and, therefore, being fair toward workers is the same as being fair toward beneficiaries. If beneficiaries and their associated workers are not identical, however, meritocrats need to balance two potentially conflicting fairness views: if the two workers exert different levels of effort, the distribution that is considered fair toward the two workers may be different from the distribution that is considered fair toward the two beneficiaries, who both exert no effort. Given that fairness toward the workers calls for no redistribution whereas fairness toward the beneficiaries demands full equalization, individuals face a dilemma because they infringe meritocratic fairness no matter how they redistribute. Because beneficiaries merit similar but inherit different outcomes, meritocrats may be less willing to accept inherited inequality as compared to noninherited inequality.

The corresponding experiment builds on the impartial spectator paradigm (Konow, 2000; Cappelen et al., 2013) and consists of two stages. In the earnings stage, an initial distribution of $10 between two stakeholders is determined. In the first of two treatment dimensions, we vary whether the two stakeholders themselves work on a real-effort task to generate earnings (Noninherited Inequality), or whether they each profit from the work of a real-life friend (Inherited Inequality). In the second treatment dimension, we vary whether workers complete the same fixed number of tasks and the initial distribution is determined by a random draw (Luck), or whether workers choose how many tasks to complete and the initial distribution is proportional to the relative number of completed tasks (Effort). In the redistribution stage, we sample 543 impartial spectators representative of the general US population who can redistribute the $10 between pairs of workers (Noninherited Inequality conditions) or workers' friends (Inherited Inequality conditions). Based on the treatment variation in the earnings stage, we implement a 2$x$2 within-subjects design in the redistribution stage: spectators make redistribution decisions for each of the four types of situations. For each situation, they observe the initial distribution and workers' relative effort before they determine the final allocation. Spectators are impartial in the sense that they have no stakes in the distribution themselves. Because redistribution is costless, we interpret the final allocation as the allocation they consider fair.

Besides the absence of spectator self-interest, this experimental setting has a number of additional advantages. First, it allows to abstract from other factors that affect distributional preferences and support for redistributive policies, such as efficiency considerations or trust in the government (Almås, Cappelen, and Tungodden, 2020; Stantcheva, 2021). Second, the comparability of redistribution decisions across experimental conditions enables us to isolate how variations in our two dimensions of interest—whether the initial distribution is tied to workers' relative efforts or based on a random draw, and whether beneficiaries are responsible for their outcomes themselves or not—affect which distribution spectators find fair. Finally, while the intergenerational transmission of wealth will be our leading example, the phenomenon that individuals derive advantages from the achievements of others is more widespread. Besides inheriting from family members, people might also profit differentially from friendship ties, coworkers, or their countries' institutional environments. Hence, studying fairness preferences in an abstract setting may yield insights into behavior, policy preferences, and fairness views in a variety of settings that have inherited inequality in general and the dilemma of meritocracy in particular at their core.

Our empirical results are in line with our theoretical framework and yet surprising. Consistent with the existing literature, we find that in the Noninherited Inequality & Luck condition redistribution levels are substantially higher than in Noninherited Inequality & Effort (Cappelen et al., 2020). Spectators equalize about 80% of the initial inequality on average in the Luck case but only about 5% in the Effort case. Comparing redistribution levels between the two Luck conditions reveals that spectators redistribute in a similar way when beneficiaries profit from the random draw of their friends compared to a random draw of themselves. In the Effort domain, however, spectators indeed redistribute significantly more if inequality is inherited. While in the Noninherited Inequality & Effort condition spectators equalize 5% of the inequality in the initial distribution, this share increases to 8% in Inherited Inequality & Effort.

The key takeaway though is that spectators redistribute a small fraction of the initial inequality in Inherited Inequality & Effort, close to the Noninherited Inequality & Effort benchmark but far away from the Luck benchmark of 80%. In other words, most spectators handle the dilemma of meritocracy by prioritizing fairness toward the benefactors over fairness toward the beneficiaries. This result seems to be a general feature of the US population, as it does not vary much by demographic variables like age, gender, or political ideology. Hence, there appears to be a broad consensus among US citizens that inherited inequality is acceptable as long as it is merited by those who bequest.

We examine potential reasons why spectators tend to handle the dilemma of meritocracy in favor of the benefactors by analyzing open-ended responses in which spectators explain their redistribution decisions. Consistent with their decisions, most spectators state to redistribute based on the workers' (and not their non-working friends') relative efforts in the Inherited Inequality & Effort condition. Zooming in on spectators who acknowledge the dilemma, i.e. that they infringe meritocratic fairness irrespective of how they redistribute, reveals a more instructive consideration behind redistribution decisions: many of these spectators argue that passive friends are not entitled to payoffs whatsoever, such that fairness toward the workers receives a much larger weight in their decision process. Under the assumption that workers prefer their own friends to receive the earnings they have merited through their efforts, this relative weighting of conflicting fairness judgments calls for the low level of redistribution that we observe in the experiment.

These considerations suggest that spectators observe workers' relative efforts, derive their relative entitlements, and then implement redistribution decisions trying to take into account (in particular the more industrious worker's) preferences over the distribution of payoffs between passive friends. To substantiate that this is a common rationale behind spectator's decisions, we explore how decisions are associated with spectators' (incentivized) beliefs about workers' preferred distributions of the $10 between their own and the other worker's friend. Indeed, spectators who believe that workers prefer distributions that more strongly favor their own friends redistribute less. Despite being neither causal nor conclusive, these observations suggest that spectators prioritize meritocratic fairness toward workers and try to respect workers' distributional preferences.

Due to the within-subjects design employed in the spectator stage, we can relate a given spectator's decisions across the four treatment conditions. Both within the Noninherited Inequality and the Inherited Inequality domain, we use this feature to classify spectators into one of three fairness types that have received the most attention in the literature, and a residual type: egalitarians who prioritize equality and always redistribute, libertarians who prioritize property rights and personal freedom and never redistribute, and meritocrats who prefer distributions that reflect relative efforts. In the Noninherited Inequality domain, we can classify all but one spectator into one of the three fairness types. By far the most prevalent fairness type is the meritocratic one (76%), followed by libertarians (21%) and only few egalitarians (3%). Most spectators display similar redistribution patterns in situations with Noninherited Inequality and Inherited Inequality. While we observe some switching between meritocrats and libertarians that is not in line with our theoretical framework, more than 85% of the spectators behave in a way that is consistent. We conclude that our theoretical framework can accommodate spectators' redistribution behavior well.

We also relate our experimental measures of fairness preferences to attitudes toward various redistribution-related policies including income and estate taxation, disability and unemployment insurance, and support for equal opportunity programs. Because redistribution decisions across Noninherited Inequality and Inherited Inequality situations are highly correlated both within the Luck and the Effort domain, we apply a factor analysis to reduce the four behavioral measures elicited in the experiment to two factor variables. One of these factor variables captures variation in redistribution behavior in the Luck domain while the other one captures variation in redistribution behavior in the Effort domain. We find that more redistribution in the experiment is related to more support for redistribution regarding all policies. This suggests that the fairness preferences identified in this experiment are a fundamental preference underlying attitudes toward various policies.

Finally, researchers who seek to relate survey responses to individual fairness preferences may often not have the resources to accommodate a thorough experimental elicitation of these preferences. We validate that unincentivized survey questions included in the post-experimental questionnaire correlate strongly with the experimentally elicited preferences in Noninherited Inequality situations. Hence, these survey items may constitute an economical alternative in the presence of organizational constraints.

This paper contributes to a growing literature that explores how contextual and personal factors determine individuals' fairness views and redistributional preferences (Cappelen et al., 2020). With regard to personal factors, it has been studied how redistributional preferences are associated with risk preferences (Gärtner, Mollerstrom, and Seim, 2017), depend on experienced inequality (Roth and Wohlfart, 2018), and respond to information on intergenerational mobility (Alesina, Stantcheva, and Teso, 2018) or inequality and the tax system (Kuziemko et al., 2015). In terms of contextual factors, it is well documented that many people reject inequality that is based on luck but accept inequality if stakeholders are responsible for their outcomes, for example due to investment decisions (Cappelen et al., 2007), effort provision (Cappelen, Sørensen, and Tungodden, 2010; Cappelen and Tungodden, 2017; Andre, 2022; Cappelen et al., 2022; Schaube and Strang, 2022), or risk-taking (Cappelen et al., 2013; Mollerstrom, Reme, and Sørensen, 2015). Relative to this literature, our study differs in two key aspects: first, we are primarily interested in situations where individuals are not responsible for their outcomes themselves but profit—potentially to a differential extent—from the actions of others. Second, the situations studied in existing papers usually yield interesting decision problems because individuals face uncertainty regarding decision-relevant aspects of the situation, such as to what extent the initial distribution is based on factors within versus outside individuals' control. In contrast, in our case individuals who endorse a meritocratic fairness ideal face a non-trivial decision problem even if they are perfectly informed about all relevant aspects of the situation; the dilemma originates from the fact that they will infringe meritocratic fairness no matter how they redistribute.

Our results may also help to explain why many people oppose redistributive policies. Several studies show that people's preferences regarding redistributive policies are strongly related to whether they find inequality fair or unfair (Alesina and Angeletos, 2005; Alesina and Giuliano, 2011; Stantcheva, 2021). At the same time, economic inequality is often inherited either directly through bequests or indirectly through differential education, social environments, and parenting (Bowles and Gintis, 2002; Björklund, Roine, and Waldenström, 2012; Chetty, Hendren, and Katz, 2016; Kosse et al., 2020). Hence, our finding that individuals tend to consider inequality as fair if it is based on effort at some stage suggests that people may reject redistributive policies based on fundamental fairness preferences. Faced with two similarly unattractive options, many people might perceive inherited inequality or unequal opportunity as the lesser evil and prioritize rewarding the efforts of those who pass on resources.

While Bowles and Gintis (2002) and Stantcheva (2021) briefly discuss the dilemma of meritocracy and Bénabou (2000) and Piketty and Saez (2013) study related issues theoretically, Cohen, Maltz, and Ofek-Shanny (2022) is most closely related to our paper. They employ the impartial spectator design to experimentally study fairness preferences in a setting where inequality between two non-working individuals originates from the decision of a worker who has to pass on all earned money to one of these two individuals. Contrary to our results, they find that impartial spectators redistribute between the non-working subjects in a similar way as between two workers who are randomly assigned unequal initial endowments. A key difference to our design, where workers generate payments for real-life friends, is that in Cohen, Maltz, and Ofek-Shanny (2022) the worker can differentiate between the two individuals only based on their favorite hobbies, which they had to list beforehand. Because the non-working subjects are otherwise strangers to the worker, spectators may wonder whether the worker would not actually prefer an egalitarian split. Notably, the design of Cohen, Maltz, and Ofek-Shanny (2022) requires workers to pass on all of the money to one individual, precluding an equal split. If spectators indeed try to respect workers' preferences — as our analysis suggests — one would then expect redistribution toward an egalitarian split, which is common in the luck case. Hence, the results in Cohen, Maltz, and Ofek-Shanny (2022) can be well reconciled with ours.

The remainder of the paper is structured as follows: Section 1.2 introduces the theoretical framework to study fairness preferences under inherited inequality in general and the Dilemma of Meritocracy in particular. Section 1.3 details the experimental design, Section 1.4 outlines the empirical strategy, and Section 1.5 reports the results. Finally, Section 1.6 concludes.

## 1.2 Theoretical Framework

We are primarily interested in situations where individuals are not responsible for their outcomes themselves but profit—potentially to a differential extent—from the efforts of others. In such situations, fairness judgments may not only need to take into account whether inequality reflects differential luck or differential efforts but also balance fairness toward individuals who generated payments and toward individuals who receive these payments. To accommodate these situations, we extend the framework in Cappelen et al. (2013) and Almås, Cappelen, and Tungodden (2020) to allow for cases of inherited inequality, in which the person responsible for an outcome is not identical to the person who receives that outcome. We derive behavioral hypotheses in Section 1.4.3, after introducing the experimental design.

### 1.2.1 Setup

We study distributional preferences in a situation in which a fixed sum of money $P$ is distributed between two individuals ("beneficiaries" $B_X$ and $B_Y$), who each benefit from the effort of an associated worker ($W_X$ and $W_Y$). Workers exert effort for their respective beneficiaries because they are interested in their well-being; for example, one may think of workers as parents caring for their respective child. Let $e_{W_i} \geq 0$ denote the effort of worker $i \in \{X, Y\}$ and $e_{B_X} = e_{B_Y} = 0$ the effort of the two beneficiaries, who are entirely passive. After workers have exerted effort, an initial distribution of $P$ between the two beneficiaries is realized, which may depend on effort levels and a random process. This distribution is described by $(s_0, 1-s_0)$, with $s_0$ being the initial (relative) share of $B_X$. Without loss of generality, we assume that $B_X$ is the initially weakly disadvantaged beneficiary, i.e., $s_0 \leq 0.5$.

Consider an impartial spectator who observes this situation and contemplates whether the distribution is fair or should be altered. The spectator is impartial in the sense that he does not receive a material benefit but incurs disutility if he perceives the distribution between the two beneficiaries to be unfair. We assume that the spectator's utility function is given by

$$V(s|\sigma) = -\frac{\alpha}{2}(\underbrace{s - s_W^f(\sigma)}_{\substack{\text{deviation from}\\ \text{what is fair}\\ \text{toward workers}}})^2 - \frac{1-\alpha}{2}(\underbrace{s - s_B^f(\sigma)}_{\substack{\text{deviation from}\\ \text{what is fair}\\ \text{toward beneficiaries}}})^2. \tag{1.1}$$

In that expression, $\sigma$ encodes information about the situation. The spectator's fairness judgments in situation $\sigma$ are expressed by the relative shares $s_W^f(\sigma)$ and $s_B^f(\sigma)$, which describe the distributions $(s_L^f(\sigma), 1-s_L^f(\sigma))$, $L \in \{W, B\}$, that the spectator considers fair toward the workers and beneficiaries, respectively. Quadratic loss functions capture the disutility from distributions that deviate from what is considered fair, and $\alpha \in [0,1]$ governs how the spectator balances fairness toward workers and beneficiaries. Solving the corresponding maximization problem yields the distribution the spectator finds fair overall, given by

$$s^r(\sigma) = \alpha s_W^f(\sigma) + (1-\alpha) s_B^f(\sigma). \tag{1.2}$$

Under the given functional form assumptions, the spectator's preferred distribution is a linear combination of the distribution considered fair toward the workers and the distribution considered fair toward the beneficiaries, with weights $\alpha$ and $1-\alpha$, respectively.

### 1.2.2 Fairness Types, Fairness Judgments, and the Dilemma of Meritocracy

Let us turn to the question of how spectators make fairness judgments. We follow the literature by assuming that spectators endorse either an egalitarian ($E$), libertarian ($L$), or meritocratic ($M$) fairness type $\tau$.

**Egalitarians ($\tau = E$).** An egalitarian is convinced that total resources should be distributed equally in any case. Hence, the distribution perceived fair toward workers as well as beneficiaries is given by $s_W^f(\sigma) = s_B^f(\sigma) = s^f(\sigma) = \frac{1}{2}$. Because perceived fair shares coincide, egalitarians do not encounter a conflict in the case of inherited inequality, and the preferred distribution is $s^r(\sigma) = \frac{1}{2}$.

**Libertarians ($\tau = L$).** A libertarian does not value equality but advocates the opposing standpoint that one should not intervene in the allocation process and therefore accepts the initial allocation. The perceived fair distributions are given by $s_W^f(\sigma) = s_B^f(\sigma) = s^f(\sigma) = s_0$ and the overall preferred distribution is $s^r(\sigma) = s_0$.

**Meritocrats ($\tau = M$).** In between, meritocrats think that distributions should reflect individual merits: $s_L^f(\sigma) = \frac{e_{L_X}}{e_{L_X}+e_{L_Y}}$ if $e_{L_X} + e_{L_Y} > 0$ and $s_L^f(\sigma) = \frac{1}{2}$ if $e_{L_X} + e_{L_Y} = 0$, with $L \in \{W, B\}$. Hence, in the case of inherited inequality, meritocrats may face a dilemma: because beneficiaries do not exert any effort but their associated workers may exert different levels of effort ($e_{W_X} \neq e_{W_Y}$), it follows that $s_B^f = \frac{1}{2}$ but usually $s_W^f = e_{W_X}/(e_{W_X} + e_{W_Y}) \neq \frac{1}{2}$ — merit judgments conflict! As a consequence, meritocrats need to balance fairness toward workers and beneficiaries, and the overall perceived fair share is given by

$$s^r(\sigma) \;=\; \alpha \frac{e_{W_X}}{e_{W_X} + e_{W_Y}} + (1-\alpha)\frac{1}{2}. \tag{1.3}$$

We denominate this phenomenon the *Dilemma of Meritocracy*. If one worker chose to exert higher effort for the sake of his beneficiary than the other, this pulls the meritocrat toward a distribution between beneficiaries that reflects these differences in effort. Conversely, both beneficiaries are passive and none merited more resources than the other, which pulls the meritocrat toward an egalitarian distribution. The weighting parameter $\alpha$ that governs how this dilemma is handled may be interpreted as the relative importance of the workers' and the beneficiaries' perspectives in the meritocrat's overall fairness judgment.

### 1.2.3 Noninherited Inequality

Our framework nests the case of noninherited inequality studied in existing research, where each worker is identical to his associated beneficiary, $W_i \equiv B_i$. This implies that $e_{W_i} = e_{B_i}$ and fairness judgments toward workers and beneficiaries coincide for all fairness types: $s_W^f = s_B^f = s^f$. The spectator's utility function collapses to $V(s|\sigma) = -(s - s^f(\sigma))^2$, and the solution is simply $s^r(\sigma) = s^f(\sigma)$, such that one reobtains the formulation used in Cappelen et al. (2013) and Almås, Cappelen, and Tungodden (2020).

## 1.3 Experimental Design

Our experiment builds on the impartial spectator paradigm (Konow, 2000; Cappelen et al., 2013) and consists of two stages. In the earnings stage, an initial (pre-redistribution) allocation of \$10 between two stakeholders is determined. In the redistribution stage, impartial spectators may redistribute the \$10 between the two stakeholders to determine the final (post-redistribution) allocation. We are primarily interested in spectators' redistribution decisions; the earnings stage is used to incentivize these decisions.

### 1.3.1 The Earnings Stage

In the earnings stage, we implement four treatment conditions in a between-subjects design. In all conditions, subjects work on a real-effort task in which they have to reposition sliders into the middle position (Gill and Prowse, 2012). Each task has a fixed duration of 30 seconds and requires repositioning 5 sliders, which is easy to achieve. Hence, completing tasks is solely a matter of effort and time, but not ability. After workers have completed their participation, they are divided into pairs of two. Treatments differ in two dimensions. One dimension varies whether the initial distribution of the $10 is determined by a random draw ("Luck") or reflects the relative number of completed tasks ("Effort"). The other dimension varies whether the $10 is distributed between a pair of workers themselves ("Noninherited Inequality") or whether each worker designates a real-life friend and the $10 is distributed between the two friends of a pair of workers ("Inherited Inequality"). Working with real-life friends has organizational advantages over, for example, the stricter requirement that workers designate a beneficiary among their family members. At the same time, friendship ties capture two central aspects of relationships between benefactors and beneficiaries that may be prerequisites for the dilemma of meritocracy: there is a meaningful relationship between workers and their friends, and workers are more altruistic toward their own friend than toward the friend of the other worker (Gächter, Starmer, and Tufano, 2015).

The $2 \times 2$ variation in the earnings stage results in the following four conditions which are summarized in Table 1.1:

- **Noninherited Inequality & Luck:** Workers complete exactly 20 tasks. $10 are distributed between the two workers of a pair. The initial distribution is determined by a random draw. Each distribution is equally likely.
- **Noninherited Inequality & Effort:** Workers choose to complete between 0 and 40 tasks. $10 are distributed between the two workers of a pair. The initial distribution corresponds to the relative number of completed tasks.
- **Inherited Inequality & Luck:** Workers complete exactly 20 tasks. Each worker chooses a real-life friend, and $10 is distributed between the workers' friends. The initial distribution is determined by a random draw. Each distribution is equally likely.
- **Inherited Inequality & Effort:** Workers choose to complete between 0 and 40 tasks. Each worker chooses a real-life friend, and $10 is distributed be-

tween the workers' friends. The initial distribution corresponds to the relative number of completed tasks.

**Table 1.1.** Features of Treatment Arms

| Treatment | \$10 distr. betw. | # Tasks completed | Initial allocation |
|---|---|---|---|
| NONINHERITED INEQ. & LUCK | Workers | $e_x = e_y = 20$ | $s_0 \sim U[0,1]$ |
| NONINHERITED INEQ. & EFFORT | Workers | $e_x, e_y \in [0,40]$ | $s_0 = e_x/(e_x + e_y)$ |
| INHERITED INEQ. & LUCK | Workers' friends | $e_x = e_y = 20$ | $s_0 \sim U[0,1]$ |
| INHERITED INEQ. & EFFORT | Workers' friends | $e_x, e_y \in [0,40]$ | $s_0 = e_x/(e_x + e_y)$ |

*Notes:* $e_x$ and $e_y$ denote the number of tasks by worker X and Y, respectively. $U[\cdot]$ denotes the uniform distribution and $s_0$ denotes the share of the \$10 allocated to stakeholder X according to the initial distribution. The share of the \$10 allocated to stakeholder Y according to the initial distribution always equals $1 - s_0$.

Before they start working, workers know whether they generate earnings for themselves or a real-life friend and how the initial allocation is determined. They also know that another person's decision may affect their (or their friend's) payoff, but not how and why. Workers (and their friends) never observe the initial allocation or spectators' decisions. Friends are entirely passive.

Workers make a final decision at the end of the earnings stage. We ask workers in the NONINHERITED INEQUALITY conditions how they would distribute additional \$10 between themselves and the worker they are matched to if they could freely decide. Likewise, we ask workers in the INHERITED INEQUALITY conditions how they would distribute \$10 between their own friend and the friend of the worker they are matched to. Workers are incentivized to report their preferences truthfully, as we would randomly draw one worker and implement his or her preference. We will later refer to these decisions as dictator decisions.

### 1.3.2 The Redistribution Stage

In the redistribution stage, unrelated subjects ("impartial spectators") can redistribute the \$10 between pairs of workers or workers' friends. Based on the four conditions from the earnings stage, we implement a $2 \times 2$ within-subjects design in the redistribution stage. Before they make a redistribution decision, spectators learn whether \$10 is distributed between workers or passive friends, whether the initial allocation was determined by a random draw or according to the relative number of completed tasks, and the initial allocation. They make their decision by entering the final distribution in the form of relative shares of the two workers (in the NONINHERITED INEQUALITY conditions) or friends (in the INHERITED INEQUALITY conditions) in a table that also contains condensed information about the situation. Figure 1.B.1 shows a screenshot of the decision screen in the INHERITED INEQUALITY & EFFORT condition; the other decision screens had the same structure. To focus on the fairness aspect of the redistribution problem, we abstract from a potential fairness-efficiency tradeoff (Almås, Cappelen, and Tungodden, 2020) by making redistribution costless.

Similar to recent studies that use the impartial spectator design (Schaube and Strang, 2022) we employ a variant of the strategy method (Kube and Traxler, 2011). For each spectator, we construct a set of six initial allocations that consists of one initial allocation from a randomly drawn situation that has occurred in the earnings stage and five hypothetical initial allocations that are constant across all spectators.[2] These initial allocations yield a block of 6 situations within each of the four conditions – 24 situations in total – for which we ask spectators to make redistribution decisions.

Spectators make redistribution decisions for all situations within a block before they proceed to the next one. After each block, they are prompted to briefly describe the reasoning behind their decisions. We randomize the order of blocks as well as the order of situations within each block between subjects. Spectators know that some situations are hypothetical and that we randomly select one spectator for each pair of workers (friends), whose decision for the relevant situation is implemented. Because spectators do not know whether a decision is potentially relevant or not, all decisions are probabilistically incentivized.

After spectators have completed the redistribution part, we elicit their beliefs about workers' dictator decisions. Separately for workers in the Noninherited Inequality and Inherited Inequality conditions, we ask spectators to guess how much workers on average kept for themselves or gave to their own friends, respectively. Spectators receive a bonus of \$0.20 for each guess with less than \$0.20 distance to the actual value, such that guesses are incentivized as well. Finally, spectators complete a brief questionnaire on their general attitudes toward inequality, their assessment of various policies related to inequality and redistribution, and additional demographics.

### 1.3.3 Procedures

#### 1.3.3.1 Workers and Friends

The earnings stage was conducted online in March 2022 and implemented using oTree (Chen, Schonger, and Wickens, 2016). Workers were recruited from the BonnEconLab subject pool via Hroot (Bock, Baetge, and Nicklisch, 2014). The invitation mail informed potential participants that some of them would be able to generate a payment for a real-life friend. In the confirmation email, workers in the Inherited Inequality conditions received a link that they had to pass on to

2. The hypothetical initial allocations were (\$0.00, \$10.00), (\$1.00, \$9.00), (\$2.20, \$7.80), (\$3.00, \$7.00), and (\$3.80, \$6.20). If the initial allocation in the randomly drawn situation was identical to one of the hypothetical initial allocations, the respective hypothetical initial allocation was replaced by a "backup" allocation. This case applied for 52 spectators.

a friend. Via that link, friends had to give us their bank details. On the next day, the corresponding workers received another email with a participation link only if a friend had given us his or her bank details before, such that we could ensure to be able to make all payments that were generated in the study. Workers in the Noninherited Inequality conditions were informed in the confirmation email that they were not among those participants that could generate a payment for a friend and received an email with a participation link on the next day as well. All workers could start immediately when they received the participation link and had time to conclude their participation until the end of the day.

In the earnings stage itself, workers had to enter their own bank details before they received condition-specific instructions and entered the work stage. Workers in the Effort conditions could choose how many tasks to complete, whereas workers in the Luck conditions had to complete exactly 20 tasks.[3] After the work stage, workers had to make their respective dictator decision to conclude their participation.

In total, 43 workers completed their participation in the earnings stage, 21 in the Noninherited Inequality conditions and 22 in the Inherited Inequality conditions. In the Nonhereditary Inequality conditions, each worker received a fixed payment of $3, and $10 was distributed between two workers each. In the Inherited Inequality conditions, each worker received a fixed payment of $5, each friend received a fixed payment of $3, and $10 was distributed between two friends each. In addition, one among all workers' dictator decisions was randomly selected and implemented as announced during the study. Payoffs were presented in the form of experimental currency during the earnings stage but eventually made in euros via bank transfer.

#### 1.3.3.2 Spectators

The redistribution stage was conducted online in late April 2022 and implemented using oTree as well. We recruited a sample of 552 adult US citizens via the survey provider Prolific, which has been shown to provide higher data quality than comparable companies (Palan and Schitter, 2018; Peer et al., 2022). In addition to incentivizing redistribution decisions, we took several measures to further promote quality responses, including two attention checks, control questions for each block of redistribution decisions, and graphical instructions that are arguably more engaging than large blocks of text instructions. Details and data quality checks are presented in Section 1.A, which also provides evidence that spectators recognized and understood the differences between treatments.

3. Workers could at most work on 60 tasks until the work stage was automatically concluded. One worker in the Luck conditions did not manage to complete 20 tasks with 60 attempts and did not generate a payment, as was announced beforehand.

Spectators were recruited in two waves within the same week.[4] The first and second wave contained 75 and 477 spectators, respectively. Because participants from the first wave were not excluded from participating in the second wave, 9 spectators participated twice. We only include the first observation from these participants, such that we end up with a sample of 543 spectators. The median completion time in the first wave was 21 minutes and subjects earned a base rate of £3.03 plus bonus payments. The median completion time in the second wave was slightly longer at 25 minutes and participants earned a base rate of £2.55 plus bonus payments. For the second wave, Prolific recruited a sample representative of the US adult population aged 18 or older regarding the joint distribution of age, sex, and ethnicity. This was impossible for the first wave due to the low number of participants. Yet, as shown in Table 1.C.1, our total spectator sample is representative of the adult US population in terms of age, gender, and ethnicity. In contrast, our sample overrepresents the well-educated and underrepresents the top quartile of the income distribution, which is common for survey samples (Stantcheva, 2022). The study was preregistered at the AER RCT Registry (RCT ID: AEARCTR-0009186). The instructions for the spectator session are presented in Section 1.D, and the pre-analysis plan can be accessed here: https://doi.org/10.1257/rct.9186.

## 1.4 Empirical Analysis

### 1.4.1 Main Variables

**Independent Variables.** Our main independent variables are the indicators $II_\sigma$ ($= 1$ if situation $\sigma$ features inherited inequality) and $E_\sigma$ ($= 1$ if the initial allocation in situation $\sigma$ is based on effort). Both indicators together describe the treatment condition situation $\sigma$ was embedded in. Further, we define the initial extent of inequality $\Delta_\sigma = 0.5 - s_0$, which allows us to investigate whether redistribution decisions depend on how much inequality is present in the initial allocation.

**Dependent Variables.** Observing that a spectator implements (\$4, \$6) as the final allocation indicates very different redistributional preferences if the initial allocation was (\$2, \$8) instead of (\$4, \$6). In the former case, the spectator reduces inequality while in the latter inequality is left constant. To differentiate between such cases, our analysis needs to take into account that the initial allocation varies

4. The two-wave procedure mainly served to test for technical issues. Indeed, during the first wave, we recognized that for some of the spectators one hypothetical initial allocation was always replaced by the backup allocation due to a bug, which we fixed immediately. Because there is nothing inherently special about our preselected hypothetical initial allocations this is not a big issue, though, and the respective decisions/observations are treated like all other decisions and as described in Section 1.4.2.

across situations.[5] Hence, we define as our main outcome variable the extent of redistribution implemented by spectator $i$ in situation $\sigma$,

$$\theta_{i,\sigma} = \frac{s_i^r - s_0}{0.5 - s_0}. \tag{1.4}$$

The extent of redistribution describes the fraction of inequality in the initial situation that is equalized by spectator i's redistribution decision. $\theta_{i,\sigma} = 1$ indicates that spectator i completely equalizes payoffs in situation $\sigma$ while $\theta_{i,\sigma} = 0$ means that spectator i accepts the initial allocation. For some analyses we use the average of spectator i's redistribution decisions within a given condition, which we refer to as the average extent of redistribution, $\bar{\theta}_{i,c}$, $c \in$ {NI-L, NI-E, II-L, II-E}.

### 1.4.2 Exclusion Criteria and Restricted Sample

To ensure high data quality, we remove some observations from our main sample as preregistered. First, we drop spectators who fail both attention checks. Second, if a spectator rushes unreasonably fast through the instructions for a given block of redistribution decisions, we drop the decisions of that spectator for the corresponding condition. Third, we only include observations for situations that all spectators encountered because these are constant across spectators and admit a clean comparison. Hence, the main sample does not include observations based on a true scenario (except if that scenario coincides with a hypothetical one) or the backup scenario.

Based on the main sample, we further construct a restricted sample that disregards observations that cannot be reconciled with the fairness ideals prevalent in the literature, which was preregistered as well. First, we drop observations which imply $\theta_{i,\sigma} < 0$ (the spectator redistributes money from the already disadvantaged beneficiary to the already advantaged beneficiary) or $\theta_{i,\sigma} > 1$ (the spectator redistributes more to the initially disadvantaged beneficiary than what would lead to a 50/50 split). While such decisions should not prematurely be characterized as "noise" or "irrational", we cannot explain these decisions within our framework and our hypotheses do not pertain to such behavior. Second, we completely drop a spectator from the restricted sample if we disregard 3 or more decisions of that spectator within any of the four conditions, either because the spectator rushed or because too many decisions imply $\theta_{i,\sigma} \notin [0, 1]$.

5. This is different from existing studies on fairness preferences in the context of noninherited inequality, where usually one of the two workers receives all of the money in the initial distribution (see e.g. (Cappelen and Tungodden, 2017; Almås et al., 2022; Cappelen et al., 2022; Schaube and Strang, 2022)). In that case, it suffices to normalize that the first worker is the initially disadvantaged one (or vice versa) and consider how much that worker receives after redistribution.

**Table 1.2.** Predicted Extent of Inequality $\theta$ by Condition and Fairness Type

| Condition | Egalitarians | Libertarians | Meritocrats |
|---|---|---|---|
| Noninherited Ineq. & Luck | 1 | 0 | 1 |
| Noninherited Ineq. & Effort | 1 | 0 | 0 |
| Inherited Ineq. & Luck | 1 | 0 | 1 |
| Inherited Ineq. & Effort | 1 | 0 | $1-\alpha$ |

Starting with 543 spectators and 13,032 decision observations, we end up with 543 spectators and 10,236 decision observations in the main sample and 437 spectators and 8,399 observations in the restricted sample. Unless indicated differently, the results presented in the paper are based on the restricted sample. However, results do not differ notably if we consider the main sample or all of the 13,032 observations for which our main outcome measure is defined, that is, where the initial allocation is not 50/50.

### 1.4.3 Behavioral Predictions & Preregistered Hypotheses

The theoretical framework outlined in Section 1.2 makes nuanced individual-level predictions about what kinds of behavioral patterns we should observe across the four treatment conditions, given a subjects' fairness type: egalitarians always prefer equal distributions, libertarians always go with the initial distribution, and meritocrats prefer distributions that reflect relative effort. Given that $e_{W_X}/(e_{W_X} + e_{W_Y})$ equals 1/2 in the Luck conditions and $s_0$ in the Effort conditions, the expression for the perceived fair share (Equation 1.2) collapses to numbers for each of the three fairness types. Plugging these numbers into the definition of the extent of redistribution (Equation 1.4) yields predictions on the extent of redistribution spectators with different fairness types implement in the different conditions. These predictions are summarized in Table 1.2.

Assuming that all types are present in our sample, these predictions imply that the four conditions should be ordered in terms of the average extent of redistribution as follows: $\bar{\theta}_{NI-L} = \bar{\theta}_{II-L} \geq \bar{\theta}_{II-E} \geq \bar{\theta}_{NI-E}$, with at least one of the inequalities being strict. Based on the individual-level predictions and this expected ordering, we derive the following four (preregistered) aggregate-level predictions that we will formally test using ordinary least squares (OLS) regressions and clustering standard errors on the spectator-level:

**Hypothesis 1.1.** *Spectators redistribute less if inequality is based on effort instead of luck.*

Because this hypothesis should hold both in the noninherited inequality domain (H1a) and — weakly — in the inherited inequality domain (H1b), we will test it separately within both domains. Formally, we estimate the following (regression) equation:

$$\theta_{i,\sigma} \;=\; \beta + \beta_E \cdot E_\sigma + \delta \cdot \Delta_\sigma + \varepsilon_{i,\sigma}. \tag{1.5}$$

We preregistered to test $H_0 : \beta_E = 0$ against $H_1 : \beta_E \neq 0$ and interpret $\beta_E < 0$ and the rejection of $H_0$ as evidence in favour of Hypothesis 1.

**Hypothesis 1.2.** *Spectators redistribute more if inequality is inherited.*

Pooling the data from the LUCK and EFFORT conditions, we estimate

$$\theta_{i,\sigma} \;=\; \beta + \beta_{II} \cdot II_\sigma + \delta \cdot \Delta_\sigma + \varepsilon_{i,\sigma}, \tag{1.6}$$

and test $H_0 : \beta_{II} = 0$ against $H_1 : \beta_{II} \neq 0$ as preregistered, interpreting $\beta_{II} > 0$ and the rejection of $H_0$ as evidence in favour of Hypothesis 2.

**Hypothesis 1.3.** *The higher extent of redistribution in the case of inherited inequality is driven by situations in which inequality is based on effort.*

To formally test whether the fact that inequality is inherited indeed only matters if the initial allocation is based on effort, we consider the following difference-in-difference-like regression equation:

$$\theta_{i,\sigma} \;=\; \beta + \beta_E \cdot E_\sigma + \beta_{II} \cdot II_\sigma + \beta_{E,II} \cdot E_\sigma \cdot II_\sigma + \delta \cdot \Delta_\sigma + \varepsilon_{i,\sigma}. \tag{1.7}$$

In accordance with our pre-analysis plan, we test $H_0^a : \beta_{II} = 0$ against $H_1^a : \beta_{II} \neq 0$ and $H_0^b : \beta_{E,II} = 0$ against $H_1^b : \beta_{E,II} \neq 0$. We interpret the results as evidence in favour of Hypothesis 3 if we find $\beta_{E,II} > 0$ and reject $H_0^b$ but not $H_0^a$.

**Hypothesis 1.4.** *The higher extent of redistribution in the case of inherited inequality, driven by situations in which inequality is based on effort, is driven by meritocrats.*

Due to the within-subjects design, we can relate individual redistribution patterns across conditions. We will classify spectators into the three fairness types (and a residual type) based on their decisions in the NONINHERITED INEQUALITY conditions (details follow later) and estimate

$$\begin{aligned}
\theta_{i,\sigma} \;=\;\; & \beta^E && + \beta^L L_i && + \beta^M M_i && + \beta^{NC} NC_i \\
& + \beta_E^E E_\sigma && + \beta_E^L E_\sigma L_i && + \beta_E^M E_\sigma M_i && + \beta_E^{NC} E_\sigma NC_i \\
& + \beta_{II}^E II_\sigma && + \beta_{II}^L II_\sigma L_i && + \beta_{II}^M II_\sigma M_i && + \beta_{II}^{NC} II_\sigma NC_i \\
& + \beta_{E,II}^E E_\sigma II_\sigma && + \beta_{E,II}^L E_\sigma II_\sigma L_i && + \beta_{E,II}^M E_\sigma II_\sigma M_i && + \beta_{E,II}^{NC} E_\sigma II_\sigma NC_i \\
& && && && + \delta \Delta_\sigma + \varepsilon_{i,\sigma}.
\end{aligned} \tag{1.8}$$

Here, egalitarians are the baseline type and $L_i$ (libertarian), $M_i$ (meritocrat), and $NC_i$ (non-classified) are indicators that equal 1 if spectator i is classified into the corresponding fairness type. As preregistered, we test $H_0^a : \beta_{E,II}^M = 0$ against $H_1^a : \beta_{E,II}^M \neq 0$ and $H_0^b : \beta_{E,II}^M = \beta_{E,II}^L$ against $H_1^b : \beta_{E,II}^M \neq \beta_{E,II}^L$ and interpret the results as evidence in favour of the hypothesis if $\beta_{E,II}^M > 0$, $\beta_{E,II}^M > \beta_{E,II}^L$, and we reject both $H_0^a$ and $H_0^b$.

## 1.5 Results

First, we compare the average extent of redistribution between treatment conditions, displayed in Figure 1.1. Averages are taken over all decisions of all subjects in the restricted sample. Comparing redistribution levels between NONINHERITED INEQUALITY & LUCK and NONINHERITED INEQUALITY & EFFORT, we replicate what many studies have documented before: under noninherited inequality, where workers' actions determine their own earnings and spectators do not need to balance potentially conflicting fairness ideals, they redistribute much less if distributions reflect differential effort than if they are based on a random draw. While they, on average, equalize about 80% of the inequality in the initial distribution in the LUCK case, they equalize only about 5% in the EFFORT case. These numbers suggest that many spectators in our sample subscribe to the meritocratic idea that resource distributions should reflect individual effort and achievement.

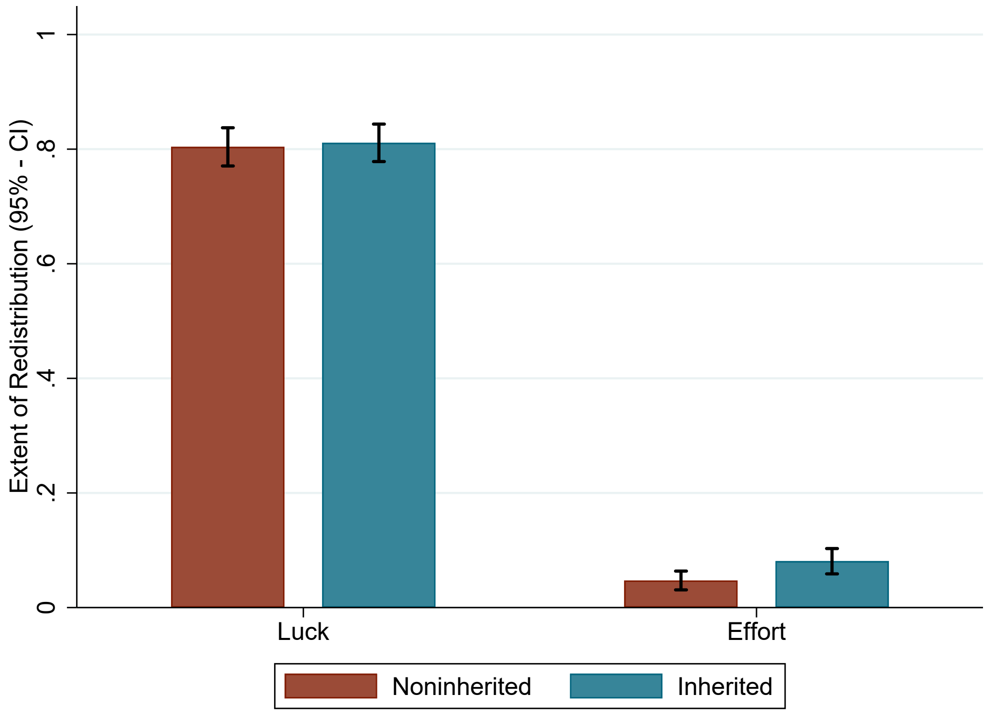



**Figure 1.1.** Average Extent of Redistribution $\bar{\theta}_{i,c}$ by Treatment Condition

*Notes:* This figure displays the average extent of redistribution $\bar{\theta}_{i,c}$ by treatment condition, together with $95-\%$ confidence intervals. Averages are taken over all decisions of all subjects in the restricted sample. Confidence intervals are based on standard errors clustered on the spectator level.

Consistent with our theoretical considerations from Section 1.2, a comparison of redistribution levels between NONINHERITED INEQUALITY & LUCK and INHERITED INEQUALITY & LUCK shows that it makes no difference whether inequality is inherited or not in the LUCK domain: the difference is insignificant and small both in absolute and relative terms.[6] This indicates that in the Luck domain, given that in either case the initial distribution is not tied to relative effort, it does not matter whether the money goes to the workers themselves or is inherited by their passive friends.

To judge how spectators deal with the dilemma of meritocracy, we examine how the average extent of redistribution in Inherited Inequality & Effort compares to the Noninherited Inequality & Luck and Noninherited Inequality & Effort benchmarks. As displayed in Figure 1.1, the fraction of inequality that is equalized in Inherited Inequality & Effort (8%) is significantly higher than the share that is equalized in Noninherited Inequality & Effort (5%).[7] However, the key takeaway is that the average extent of redistribution in Inherited Inequality & Effort is much closer to the Noninherited Inequality & Effort benchmark than to the Noninherited Inequality & Luck benchmark (80%). This is consistent with our theoretical considerations from Section 1.2, but given that any magnitude between the two benchmarks would have been similarly consistent, this result may almost be considered a corner solution. Speaking in model terms, the data suggest that spectators "have a high $\alpha$": they prioritize fairness toward the workers—whose effort is reflected in the initial distribution—and accept that in the Inherited Inequality case the beneficiaries end up with different shares even though one did not "merit" more than the other. Overall, these results suggest that spectators treat the dilemma of meritocracy by prioritizing fairness toward the workers over fairness toward the friends.

### 1.5.1 The Aggregate Level: Testing the Hypotheses

To test the hypotheses from Section 1.4.3, we estimate the corresponding preregistered regression equations using OLS regressions. All reported equations control for the initial extent of inequality in a given situation ($\Delta_\sigma$), and standard errors are always clustered on the spectator level. The results are reported in Table 1.3. The titles below the column numbers indicate which hypothesis is referred to.

The estimates in columns (1) and (2) indicate that, both in the case of Noninherited Inequality and Inherited Inequality, spectators redistribute significantly less if the initial distribution is based on effort rather than luck. The differences in the average extent of redistribution amount to 76%p (Noninherited Inequality) and 73%p (Inherited Inequality), respectively.

We further observe that the initial extent of inequality ($\Delta_\sigma$) has a weakly significant but small effect on the fraction of inequality spectators equalize. The

6. $d = 0.007$ and $p = 0.62$ in an OLS regression of the form $\theta_{i,\sigma} = \beta + \beta_{II} \cdot II_\sigma + \varepsilon_{i,\sigma}$, using only observations from the Luck domain and clustering standard errors on the spectator level.

7. $d = 0.034$ and $p < 0.001$ in an OLS regression of the form $\theta_{i,\sigma} = \beta + \beta_{II} \cdot II_\sigma + \varepsilon_{i,\sigma}$, using only observations from the Effort domain and clustering standard errors on the spectator level.

**Table 1.3.** Treatment Effects on the Extent of Redistribution $\theta_{i,\sigma}$

| | Restricted Sample | | | | Main Sample | Full Sample |
|---|---|---|---|---|---|---|
| | (1) H1a | (2) H1b | (3) H2 | (4) H3 | (5) H3 | (6) H3 |
| Effort ($E_\sigma$) | -0.757*** (0.019) | -0.730*** (0.019) | | -0.757*** (0.019) | -0.747*** (0.020) | -0.741*** (0.020) |
| Inherited ($II_\sigma$) | | | 0.022** (0.009) | 0.007 (0.014) | 0.021 (0.015) | 0.017 (0.016) |
| Effort ($E_\sigma$) × Inherited ($II_\sigma$) | | | | 0.027 (0.016) | 0.022 (0.019) | 0.042** (0.021) |
| Initial Inequality ($\Delta_\sigma$) | 0.031* (0.018) | 0.035* (0.019) | 0.024 (0.015) | 0.033** (0.015) | 0.079*** (0.019) | 0.054 (0.042) |
| Constant | 0.795*** (0.018) | 0.801*** (0.018) | 0.421*** (0.011) | 0.794*** (0.018) | 0.784*** (0.019) | 0.789*** (0.024) |
| Included Treatments | NI-L & NI-E | II-L & II-E | All | All | All | All |
| Clusters | 437 | 437 | 437 | 437 | 543 | 543 |
| Observations | 4203 | 4196 | 8399 | 8399 | 10236 | 12448 |
| $R^2$ | 0.620 | 0.575 | 0.001 | 0.598 | 0.488 | 0.364 |

*Notes:* This table reports results from OLS regressions of the extent of redistribution implemented by spectator $i$ in situation $\sigma$ on treatment indicators, controlling for the initial extent of inequality in situation $\sigma$. Columns (1) and (2) correspond to Equation 1.5 and estimate the difference between redistribution in the Effort versus Luck case, once in the Noninherited Inequality and once in the Inherited Inequality domain. Column (3) corresponds to Equation 1.6 and estimates the difference between redistribution if inequality is inherited versus noninherited, pooling Effort and Luck situations. Columns (4)–(6) correspond to Equation 1.7 and interact both treatment dimensions using observations from all treatment conditions. For information on the composition of the different subsamples, see Section 1.4.2. Standard errors (in parentheses) are clustered on the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

estimates show that the extent of redistribution is $3-4$%p higher on average if the initial extent of inequality is one unit larger. Given that the variable is only defined over the interval from 0 (a 50/50 split) to 0.5 (one stakeholder receives everything), the effect is more tangibly described by saying that, for example, going from a 30/70 split to a 20/80 split increases the average extent of redistribution by $0.3--0.4$%p. Overall, these observations yield strong support for Hypothesis 1.1:

**Result 1.1.** *In both the Noninherited Inequality and the Inherited Inequality domain, spectators redistribute considerably less on average if inequality is based on effort instead of luck.*

Moving to the regression equation in column (3), which makes use of all observations in the restricted sample, we see that spectators redistribute significantly more if inequality is inherited. Consistent with Hypothesis 1.2, the average extent of redistribution is 2.2%p higher if the money is distributed between passive friends instead of the workers themselves. Yet, in contrast to the magnitude of the difference in redistribution levels between Effort and Luck situations, the effect is almost negligible. We summarize these observations in the following result:

**Result 1.2.** *Spectators redistribute significantly more if inequality is inherited. However, the magnitude of the effect is small.*

The remaining columns, (4)-(6), test for an interaction effect: does the fact that payoffs are inherited matter more if the initial distribution is based on workers' relative effort levels instead of a random draw? Whereas the difference in average redistribution levels between Inherited Inequality and Noninherited Inequality situations is less than 1%p if the initial distribution is determined by luck, this difference is about five times as large $(0.007 + 0.027)$ if the initial distribution is proportional to workers' relative effort. The interaction effect is still small, however, and just short of reaching statistical significance. The numbers and qualitative patterns are very similar if the same equation is estimated on the main sample (column (5)), which includes observations that cannot be reconciled with commonly considered fairness ideals, i.e., $\theta_{i,\sigma} \notin [0,1]$. Similarly, results change little if we consider the full sample (column (6)), which includes situations based on true scenarios and from blocks where spectators rushed through the instructions, albeit the interaction effect is statistically significant here. Relative to our main regression equation in column (4) the share of variance explained drops sharply in columns (5) and (6), which indicates that our sample restrictions successfully reduce the amount of noise in the data. Overall, we interpret these observations as (partial) support in favour of Hypothesis 1.3:

**Result 1.3.** *The higher extent of redistribution in the case of inherited inequality is, if anything, driven by situations in which inequality is based on effort.*

### 1.5.2 The Individual Level: Redistribution Patterns & Fairness Types

Our within-subjects setup in the redistribution stage has the advantage that we can relate a given spectator's redistribution decisions across the four different conditions. In this subsection, we use this feature to detect common redistribution patterns. As a first step, we use subjects' decisions in the two Noninherited Inequality conditions to classify them into one of three fairness types discussed in Section 1.2.2: egalitarians (E), libertarians (L), and meritocrats (M). We define a spectator's fairness type in situations of noninherited inequality, $\tau_{i,NI}$, as follows:

$$\tau_{i,NI} = \begin{cases} E & \text{if } \bar{\theta}_{i,NI-L} \geq 0.5 \text{ and } \bar{\theta}_{i,NI-E} \geq 0.5 \\ M & \text{if } \bar{\theta}_{i,NI-L} \geq 0.5 \text{ and } \bar{\theta}_{i,NI-E} < 0.5 \\ L & \text{if } \bar{\theta}_{i,NI-L} < 0.5 \text{ and } \bar{\theta}_{i,NI-E} < 0.5 \\ NC & \text{else,} \end{cases} \tag{1.9}$$

where NC describes a residual type of "Nonclassifieds".

Figure 1.2 plots the distribution of spectators in the $\bar{\theta}_{i,NI-L} \times \bar{\theta}_{i,NI-E}$ space. The horizontal axis indicates the average extent of redistribution in the Noninherited Inequality & Luck condition. Similarly, the vertical axis measures the average extent of redistribution in Noninherited Inequality & Effort. Hence, each circle in Figure 1.2 represents the redistribution behavior of a spectator in the Noninherited Inequality domain, and circle size is proportional to the number of spectators at the corresponding position.

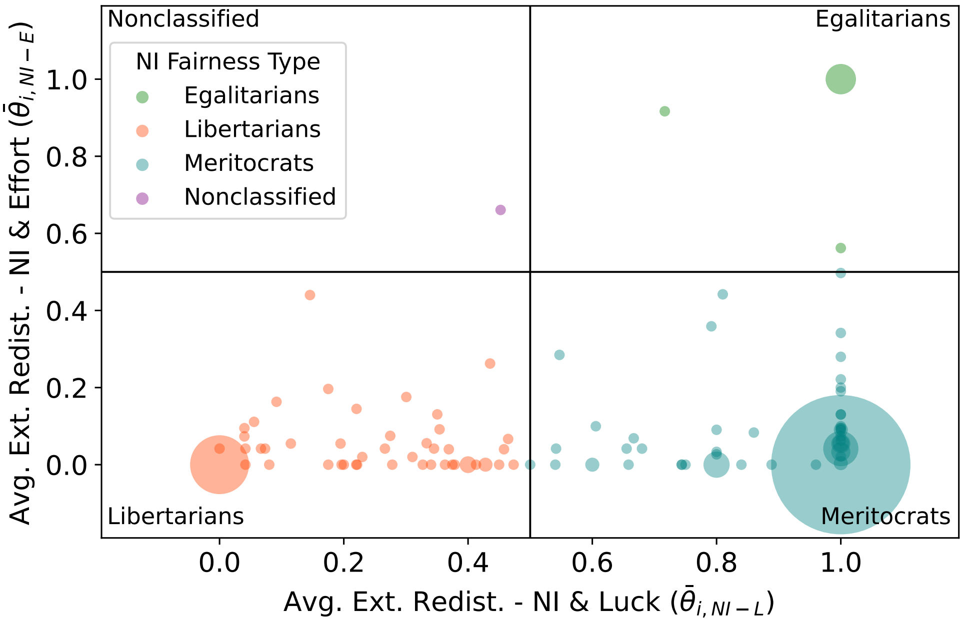



**Figure 1.2.** Classification into Fairness Types – Noninherited Inequality

*Notes:* Circles correspond to subjects in the spectator role of the experiment. The horizontal axis describes the share of inequality that the individual equalized on average in the Noninherited Inequality & Luck condition. The vertical axis describes the share of inequality that the individual equalized on average in the Noninherited Inequality & Effort condition. Circle size is proportional to the number of spectators at the corresponding position. Subjects were classified according to the label names in the four quadrants, and colors indicate the respective classes.

Two aspects of the plot attract particular attention. First, the majority of spectators (76%) fall into the bottom right quarter and are, therefore, classified as meritocrats. A much smaller fraction of spectators (21%) are classified as libertarians, and only a few (3%) are classified as egalitarians. Only a single spectator in the restricted sample remains unclassified. Second, spectators in general behave very consistently: most of them make either perfectly meritocratic (59%), libertarian (10%), or egalitarian (3%) decisions.

As a second step, in analogy to the noninherited inequality classification, we define a spectator's redistribution pattern in situations with inherited inequality, $\tau_{i,II}$:

$$\tau_{i,II} = \begin{cases} E & \text{if } \bar{\theta}_{i,II-L} \geq 0.5 \text{ and } \bar{\theta}_{i,II-E} \geq 0.5 \\ M & \text{if } \bar{\theta}_{i,II-L} \geq 0.5 \text{ and } \bar{\theta}_{i,II-E} < 0.5 \\ L & \text{if } \bar{\theta}_{i,II-L} < 0.5 \text{ and } \bar{\theta}_{i,II-E} < 0.5 \\ NC & \text{else.} \end{cases} \tag{1.10}$$

Figure 1.3 shows, in the familiar fashion, where spectators are positioned in the $\bar{\theta}_{i,II-L} \times \bar{\theta}_{i,II-E}$ space. To relate spectators' redistribution patterns across situations

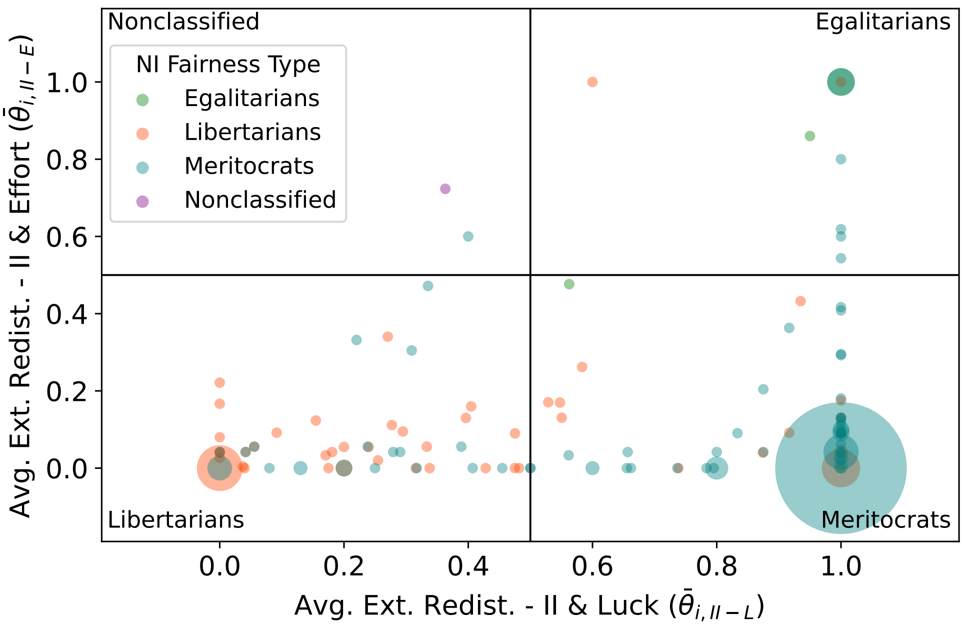



**Figure 1.3.** Classification by Redistribution Patterns – INHERITED INEQUALITY

*Notes:* Circles correspond to subjects in the spectator role of the experiment. The horizontal axis describes the share of inequality that the individual equalized on average in the INHERITED INEQUALITY & LUCK condition. The vertical axis describes the share of inequality that the individual equalized on average in the INHERITED INEQUALITY & EFFORT condition. Circle size is proportional to the number of spectators at the corresponding position. Subjects were classified according to the labels in the four quadrants. Colors indicate how spectators were classified in the NONINHERITED INEQUALITY situations.

with noninherited and inherited inequality, spectators' noninherited inequality fairness type is indicated by the color of the corresponding circle. Recall from Section 1.2 that we would not expect subjects who were classified as egalitarians and libertarians to display differential redistribution patterns if inequality is inherited. Hence, we should observe that green dots ($\tau_{i,NI} = E$) are situated in the upper right quarter of the figure, and that orange dots ($\tau_{i,NI} = L$) are situated in the lower left quarter. For meritocrats (teal circles), the theoretical prediction is vague: depending on $\alpha$—how they weigh fairness toward workers versus beneficiaries—they should either behave meritocratically ($\alpha > 0.5$, lower right quarter) or in an egalitarian way ($\alpha < 0.5$, upper right quarter).

The figure shows that, just like before, many spectators behave very consistently and are either placed on a corner or on an edge. Most spectators "remain in their quarter", that is, display similar redistribution patterns in situations featuring inherited and noninherited inequality. Focusing on those spectators who have been classified as meritocrats under noninherited inequality, we see that only a few switch to an egalitarian redistribution pattern when inequality is inherited. This indicates that most of them prioritize fairness toward the workers ($\alpha > 0.5$). In contrast to our expectations, we observe some switching between meritocrats and libertarians.

These observations are quantified in the moving matrix displayed in Figure 1.4, which shows the distribution of two-dimensional redistribution patterns in a more condensed way. The position on the vertical axis describes spectators' fairness type

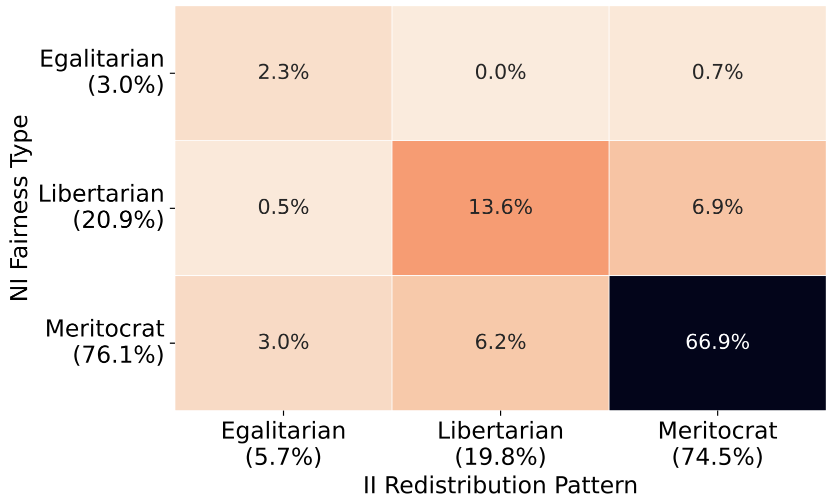



**Figure 1.4.** Two-Dimensional Redistribution Patterns

*Notes:* This moving matrix displays the distribution of spectators over two-dimensional redistribution patterns. Fairness types under noninherited inequality are shown on the vertical axis. Redistribution patterns under inherited inequality are shown on the horizontal axis.

under noninherited inequality, and the position on the horizontal axis describes their redistribution pattern under inherited inequality.[8] Marginal distributions are reported with the axis labels. The figure shows that most spectators are "on the diagonal", that is, they display the same redistribution pattern under both inherited and noninherited inequality. Only 3% of all spectators in the restricted sample switch from meritocratic to egalitarian, meaning that they prioritize fairness toward beneficiaries ($\alpha < 0.5$ in the theoretical framework). Between 6% and 7% of spectators each switch from meritocratic to libertarian or vice versa, which is not consistent with our theoretical framework and suggests that this may be more

8. The figure disregards two spectators who are nonclassified in at least one dimension.

than just noise. Besides that, there are only very few "inconsistent" spectators. Overall, more than 85% of spectators are classified in a way that is consistent with our theoretical framework, which—together with the observation that spectators make very consistent observations *within* each condition—indicates that the framework explains spectators' behavior well.

As shown theoretically in Section 1.2, the fact that the money is distributed between passive stakeholders who differentially profit from their friends' effort in the Inherited Inequality conditions should only matter for meritocrats, and only if the initial distribution reflects relative effort. To formally test whether this is the case, we estimate regression Equation 1.8 using OLS and clustering standard errors on the spectator level. We are particularly interested in the triple interaction of the Inherited Inequality and Effort indicators ($II_\sigma$ and $E_\sigma$) with spectators' (noninherited inequality) fairness type.

The results are displayed in Table 1.4, in which a number of coefficients are suppressed for increased readability.[9] The estimates in column (1), which corresponds to Equation 1.8 and uses egalitarians as the reference fairness type, show that the triple interaction effect amounts to 24.3%p and is significant for meritocrats. This indicates that, relative to egalitarians, the fact that inequality is inherited nudges meritocrats more strongly to redistribute more if inequality is based on effort instead of luck. As the triple interaction effect for meritocrats is also significantly higher than that for libertarians (Wald test, $p < 0.0001$), the data formally yields strong support for Hypothesis 1.4.

**Result 1.4.** *The fact that inheritance increases the extent of redistribution more strongly if inequality is based on effort instead of luck is driven by meritocrats.*

Considering columns (2)–(4), where Equation 1.7 is estimated separately for the three fairness types, it becomes apparent that the data do not perfectly fit the story behind Hypothesis 1.4, though. While the interaction effect of Inherited Inequality and Effort amounts to almost 10%p for meritocrats and is highly significant, in the Luck domain they redistribute on average about 6%p less if inequality is inherited, which is a significant difference as well. Conversely, libertarians redistribute on average about 27%p more if inequality is inherited in the Luck domain, while the interaction effect largely offsets this difference (−23%p) for the Effort domain, and both coefficients are highly significant again.

9. For a regression table that reports the same regression equations but does not omit coefficients, please refer to Table 1.C.2 in Section 1.C.

**Table 1.4.** Treatment Effects on the Extent of Redistribution $\theta_{i,\sigma}$ by Fairness Type

| | Restricted Sample | | | |
|---|---|---|---|---|
| | (1) Pooled | (2) Egalitarians | (3) Meritocrats | (4) Libertarians |
| EFFORT ($E_\sigma$) | -0.025 | -0.025 | -0.960*** | -0.109*** |
| | (0.036) | (0.038) | (0.006) | (0.018) |
| INHERITED ($II_\sigma$) | -0.018 | -0.017 | -0.059*** | 0.268*** |
| | (0.031) | (0.032) | (0.012) | (0.042) |
| EFFORT ($E_\sigma$) × INHERITED ($II_\sigma$) | -0.144 | -0.144 | 0.099*** | -0.232*** |
| | (0.103) | (0.108) | (0.015) | (0.044) |
| EFFORT ($E_\sigma$) × INHERITED ($II_\sigma$) × Meritocrat | 0.243** | | | |
| | (0.104) | | | |
| EFFORT ($E_\sigma$) × INHERITED ($II_\sigma$) × Libertarian | -0.088 | | | |
| | (0.112) | | | |
| Initial Inequality ($\Delta_\sigma$) | 0.031** | -0.052 | -0.004 | 0.175*** |
| | (0.014) | (0.101) | (0.012) | (0.045) |
| Constant | 0.977*** | 1.001*** | 0.977*** | 0.084*** |
| | (0.015) | (0.036) | (0.006) | (0.019) |
| Clusters | 437 | 13 | 332 | 91 |
| Observations | 8399 | 249 | 6403 | 1731 |
| $R^2$ | 0.817 | 0.106 | 0.864 | 0.228 |

*Notes:* This table reports results from OLS regressions of the extent of redistribution implemented by spectator $i$ in situation $\sigma$ on treatment indicators and spectator i's fairness type, controlling for the initial extent of inequality in situation $\sigma$. Column (1) corresponds to Equation 1.8. Columns (2)–(4) correspond to Equation 1.7 but are estimated on subsets of spectators who share the corresponding fairness type. Standard errors (in parentheses) are clustered on the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

### 1.5.3 Potential Channels

#### 1.5.3.1 Spectators' Explanations for Their Redistribution Decisions

Why do spectators redistribute so little when they face the dilemma of meritocracy? To develop an understanding of how people reason about the dilemma and to generate hypotheses for potential channels, we analyze the open-ended explanations subjects gave for their redistribution decisions. Most spectators use the opportunity to write open-ended explanations after each decision block. For all open-ended explanation fields, more than 98% of spectators make an entry. Figure 1.A.3 in Section 1.A shows that responses correspond well to treatment arms and fairness types. Hence, open-ended responses seem to provide useful information.

To get an overview of how spectators explain their decisions, we sort all mentioned explanations by hand into categories. Table 1.C.3 shows the complete list of categories and gives examples of the kind of explanations they encompass. Most spectators state specific rationales for their behavior. Yet, 49 spectators do not explain their decisions or use explanations like "I just tried to be fair", which cannot be assigned to a meaningful category. Consequently, our analysis excludes these spectators and is based on the remaining 388 subjects, who comprise about 89% of the spectators in the restricted sample.

Figure 1.5 depicts the frequencies with which explanations for redistribution decisions in Inherited Inequality & Effort are given by the explanation category. The plurality of spectators mentions that they implemented final allocations

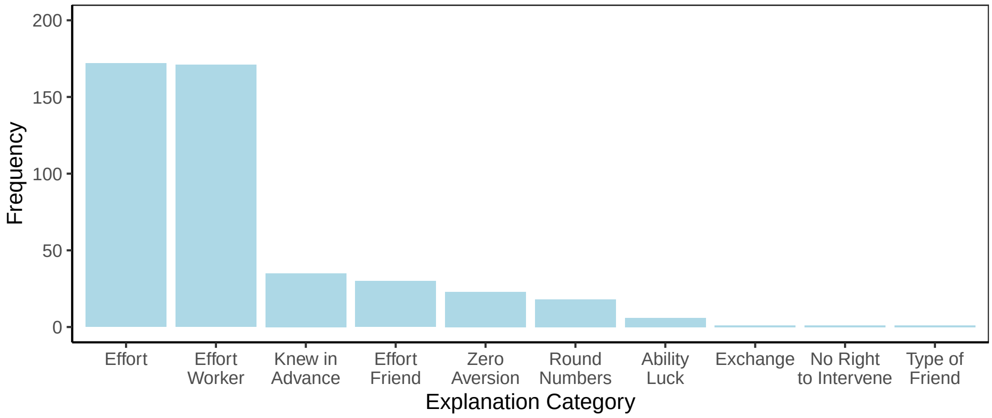



**Figure 1.5.** Spectators' Explanations for their Decisions in Inherited Inequality & Effort

*Notes:* This figure displays the frequency of explanations spectators gave for their redistribution decisions in Inherited Inequality & Effort by explanation category. Results are based on up to 3 arguments made by the 388 spectators from the restricted sample who gave specific explanations for their behavior. We included up to 3 arguments per spectator.

proportional to relative efforts without specifying whether that refers to the efforts of the workers or the efforts of the friends. Of those who specify this, most refer to the workers' efforts and few to the friends' efforts, which is consistent with our results for the redistribution decisions. The three corresponding categories contain nearly 82% of all explanations. Hence, relative effort levels appear to be the main theme behind redistribution decisions.

Alternative explanations are much less frequently mentioned by spectators. For instance, it is conceivable that a worker's effort changes the spectators' belief about what kind of person the respective friend is. However, only a single spectator mentions this as relevant to his decision. Similarly, only one spectator mentions being influenced by the thought that workers and their friends might exchange money after the experiment. Slightly more frequently mentioned explanation categories include that subjects "Knew in Advance" and agreed to the rules of the study, such that redistribution would mean an unfair ex-post rule adjustment[10]; an aversion to giving people zero or very little money; a preference for round numbers; the idea that some people might have been less able to perform the task due to bad luck; and the belief that one must not intervene in the affairs of others. Figure 1.A.5, Figure 1.A.4 and Figure 1.A.6 in the appendix show similar results for the other 3 treatment conditions. Consistent with our other results, most spectators in each condition argue that earnings should be based on effort but not on luck.

Why do most spectators base their decisions on the relative efforts of the workers rather than on the relative efforts of the friends? To examine this question, we focus on the explanations of spectators in Inherited Inequality & Effort who acknowledge the dilemma of meritocracy, because they consciously think about fairness toward the workers versus fairness toward the friends. We consider a spectator to acknowledge the dilemma of meritocracy if he provides arguments for and against redistribution based on the meritocratic fairness ideal in his explanation. Due to this strong selection requirement, this includes only 25 spectators who provide 34 arguments collectively.

Figure 1.6 shows the frequencies of explanation categories spectators use to rationalize their decisions. About 82% of all explanations belong to two categories:

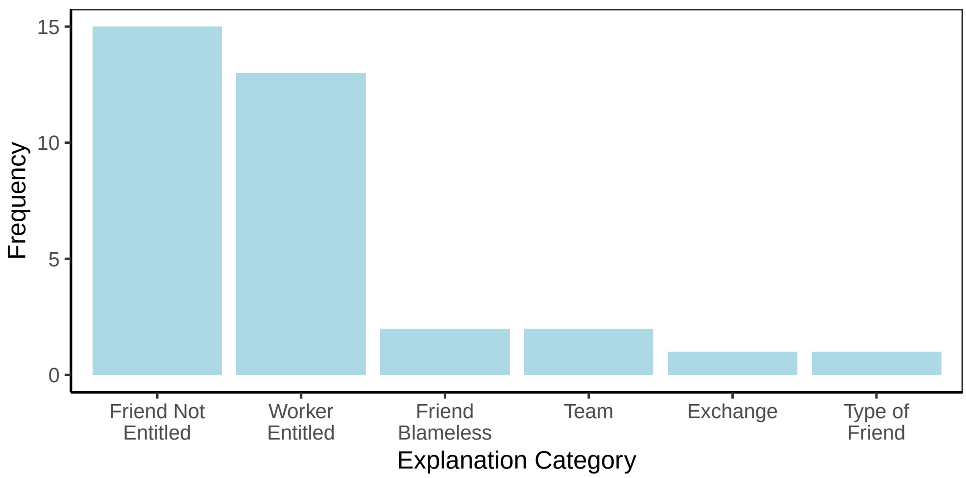



**Figure 1.6.** Spectators' Explanations for Resolving the Dilemma of Meritocracy

*Notes:* This figure displays the frequency of explanations spectators gave for resolving the dilemma of meritocracy in the way they did by explanation category. Results are based on up to 3 arguments made by 25 spectators from the restricted sample who mentioned the dilemma of meritocracy in their explanations.

explanations in the "Worker Entitled" category argue that the workers are entitled

10. As described in Section 1.3 workers were informed that their (or their friend's) payoff could be affected by the decision of a third person, and spectators knew that. Spectators who refer to this issue apparently still consider altering the initial distribution an unfair rule adjustment.

to the fruits of their labor. Conversely, explanations in the “Friend Not Entitled” category state that, in contrast to workers, friends are not entitled to the bonus payment because they did not earn it through effort. Both explanation categories refer to the same asymmetry between workers and friends: workers work for the bonus while friends do not. In the view of most spectators who mentioned the dilemma of meritocracy, this makes the entitlement of workers stronger than the entitlement of friends. This can explain why most spectators prefer to be fair toward the workers rather than toward their friends.

Again, alternative explanations are mentioned much less frequently. About 6% of the respondents mention that priority should be given to friends precisely because they did not work and are therefore blameless for the initial distribution. Another 6% view a worker and his friends as one team and argue that resources that were earned by the team should remain within the team. One respondent expects the friend to return some of his earnings to his associated worker and another respondent argues that a friend who is not worked for is not worth the work.

Hence, most spectators seem to believe workers earned the right to distribute a monetary amount that is proportional to their relative effort levels. While spectators might at the same time find it unfair that some passive friends receive less than others even though neither of them worked themselves, the former consideration might be perceived as more important. These considerations suggest that in the Effort conditions (meritocratic) spectators' redistribution decisions should depend on their belief about workers' preferred distributions. For example, a spectator might equalize the distribution between passive friends based on the belief that workers prefer a 50/50 split. Conversely, a spectator who believes that workers only care about their own friends might not redistribute to respect workers' preferences.

#### 1.5.3.2 Redistribution Decisions and Spectators' Beliefs about Workers' Preferences

To pursue this potential explanation, we make use of spectators' beliefs about how workers would distribute money in a dictator game between a) themselves and another worker and b) their own friend and the friend of another worker, elicited subsequent to the redistribution blocks.[11] If spectators indeed make merit judgments based on workers' relative effort and then try to respect their distributional preferences (in particular: those of the more industrious worker), we should observe that these beliefs are associated with the average extent of redistribution

11. Histograms of these beliefs and the individual-level differences in these beliefs are shown in Figure 1.B.2 and Figure 1.B.3 in Section 1.B.

implemented by spectators. We should further observe that these associations are stronger in the Effort conditions and driven by meritocrats.

To test these predictions, we proceed in two steps. First, we regress subjects' average extent of redistribution in a given condition on the corresponding belief about workers' preferred distribution. To make estimates comparable across conditions, we standardize both the dependent variable (across spectators but within conditions) as well as the independent variable. Formally, we estimate the following regression equation using OLS:

$$std(\bar{\theta}_{i,c}) = \alpha + \beta_{c,k} \cdot std(\mu_{i,k}) + \varepsilon_{i,c,k}. \tag{1.11}$$

As usual, $\bar{\theta}_{i,c}$ is the average extent of redistribution implemented by spectator $i$ in condition $c \in$ {NI-L, NI-E, II-L, II-E}. $\mu_{i,k}$ describes the belief of spectator $i$ about workers' preferred distributions in case $k$, with $k$ indicating which dictator decision is used: for $c \in$ {NI-L, NI-E} we use spectators' beliefs about workers' preferred distribution between themselves and the other worker, and for $c \in$ {II-L, II-E} we use spectators' beliefs about workers' preferred distribution between their own friend and the friend of the other worker.

The coefficients from these regressions are displayed in Figure 1.7. In Non-

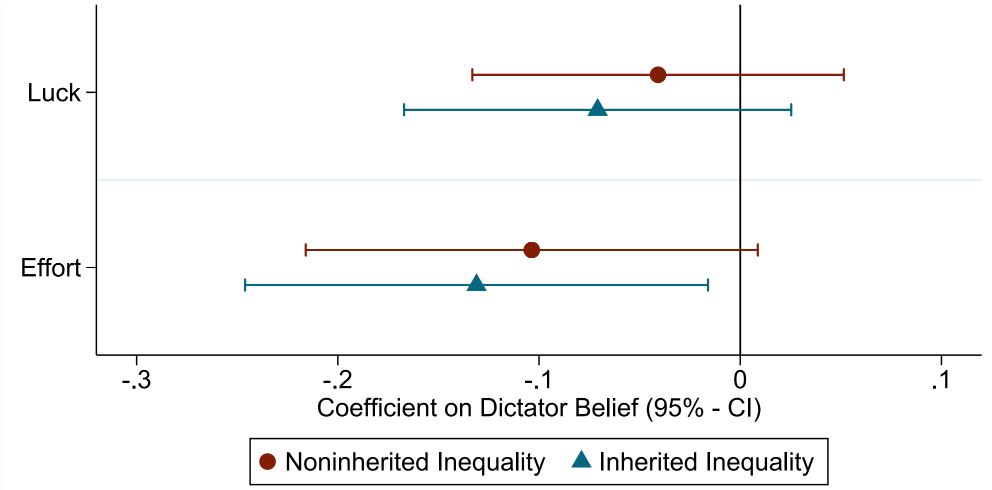



**Figure 1.7.** Association between Beliefs about Workers' Preferences and Redistribution Decisions

*Notes:* This figure displays coefficients on spectators' beliefs about workers' preferred distributions, obtained from separate regressions of redistribution levels (standardized across spectators but within conditions) on the corresponding standardized beliefs (see Equation 1.11). The corresponding regression results are reported in Table 1.C.4 in Section 1.C.

inherited Inequality & Luck, an increase of one standard deviation (SD) in the belief about the share of the \$10 workers on average keep for themselves is associated with a 0.04 SD reduction in the average extent of redistribution ($p = 0.39$). With a 1 SD increase in the same belief being associated with a 0.10

decrease in the average extent of redistribution, the estimate for the Noninherited Inequality & Effort conditions is more than twice as large and weakly significant ($p = 0.07$). In the Inherited Inequality domain, the pattern is very similar but estimated coefficients a bit larger in terms of absolute value. In Inherited Inequality & Luck, a 1 SD increase in the belief about the share of the \$10 workers on average give to their own friends is associated with a 0.07 SD decrease in the average extent of redistribution ($p = 0.15$). Again, with a 1 SD increase in the belief being associated with a 0.13 SD decrease in the average extent of redistribution, the same estimate for the Inherited Inequality & Effort condition is about twice as large and statistically significant ($p = 0.03$). These patterns indicate that spectators' beliefs about workers' preferred distributions are, in particular in the Effort case, indeed associated with their redistribution decisions in the expected way.

As a second step, we test the more nuanced prediction that these associations are most pronounced for spectators classified as meritocrats in the Noninherited Inequality domain. We estimate the same regression equation as before, but separately for the three fairness types and, to increase comparability of effects across types, standardizing the belief (redistribution) variable not across all spectators (and within a given condition), but across spectators of a given type (and within a given condition). The results for the Effort domain, reported in Table 1.5, are mixed.[12] While our sample includes too few egalitarians to consider the corre-

**Table 1.5.** Association between Beliefs and Redistribution Decisions by Fairness Type

| | Noninherited Inequality | | | Inherited Inequality | | |
|---|---|---|---|---|---|---|
| | (1) Egalitarians | (2) Meritocrats | (3) Libertarians | (4) Egalitarians | (5) Meritocrats | (6) Libertarians |
| Guess Self/Other | 0.244* | 0.043 | −0.089 | | | |
| | (0.134) | (0.045) | (0.089) | | | |
| Guess Own Friend/Other's Friend | | | | −0.246 | −0.115 | 0.036 |
| | | | | (0.291) | (0.075) | (0.136) |
| Observations | 13 | 332 | 91 | 13 | 332 | 91 |
| $R^2$ | 0.060 | 0.002 | 0.008 | 0.060 | 0.013 | 0.001 |

*Notes:* This table reports results from OLS regressions of spectators' average extent of redistribution in the two Effort conditions, standardized across spectators of a given (Noninherited Inequality) fairness type and within experimental conditions, on their beliefs about workers' preferred distributions, standardized across spectators of the same fairness type. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

sponding estimates reliable (columns (1) and (4)), the estimates for meritocrats (columns (2) and (5)) and libertarians (columns (3) and (6)) are insignificant. Focusing on meritocrats, we observe that in the Noninherited Inequality &

12. For completeness, a similar regression table reporting the results for the Luck domain can be found here: Table 1.C.5.

Effort condition, the association goes in the wrong direction ($p = 0.34$). In the Inherited Inequality & Effort condition, a 1 SD increase in the belief about the share workers on average keep for their own friends is associated with a 0.12 SD decrease in the average extent of redistribution among meritocrats. This effect, however, does not reach statistical significance ($p = 0.13$).

Overall, our observations on the relation of spectators' beliefs about workers' preferences and their redistribution decisions suggest that spectators making merit judgments and then seeking to respect (the more diligent) workers' preferences may be a part of what is behind our results. However, the associations documented in the first step seem to be driven to some extent by differentially distributed beliefs across different fairness types, and this potential explanation requires a more thorough investigation.[13]

### 1.5.4 Heterogeneity between Demographic Groups

The previous analysis has shown that most people do not redistribute in the Inherited Inequality & Effort treatment. To investigate whether this result masks heterogeneity between sociodemographic groups, we construct binary sample splits along a variety of dimensions and test whether spectators on different sides of these sample splits make different redistribution decisions. We consider the following sociodemographic characteristics: age, voting frequency (below vs. above median); sex (female vs. male); education (college degree vs. no college degree); income (below vs. above $68,000); wealth (below vs. above $124,000); party identification (republican vs. democrat); perceived social class (above vs. below middle class); and economic ideology (state- vs. market-oriented).[14] Because we have not preregistered any hypotheses regarding heterogeneity, we rely on the main sample for this exercise.

For the different sample splits, Figure 1.8 displays subgroup averages (with equal weights) of spectators' average extent of redistribution in Inherited Inequality & Effort. Heterogeneity is most pronounced along the wealth dimension. This is consistent with the notion that inherited inequality can be considered just from the perspective of those who bequest but unjust from the perspective of those who inherit — the key idea behind the dilemma of meritocracy. High-wealth

13. The average beliefs about the share workers on average keep for themselves (when they distribute between themselves and the worker they are matched to) are $4.98 (Egalitarians), $6.14 (Meritocrats), and $6.35 (Libertarians). The average beliefs about the share workers on average give to their own friends (when they distribute between their own friend and the friend of the worker they are matched to) are $5.20 (Egalitarians), $6.13 (Meritocrats), and $6.22 (Libertarians).

14. When spectators reported their political affiliation, perceived social class, and economic ideology, they could select a middle option; when we consider these sociodemographic dimensions, we drop spectators who selected this middle option.

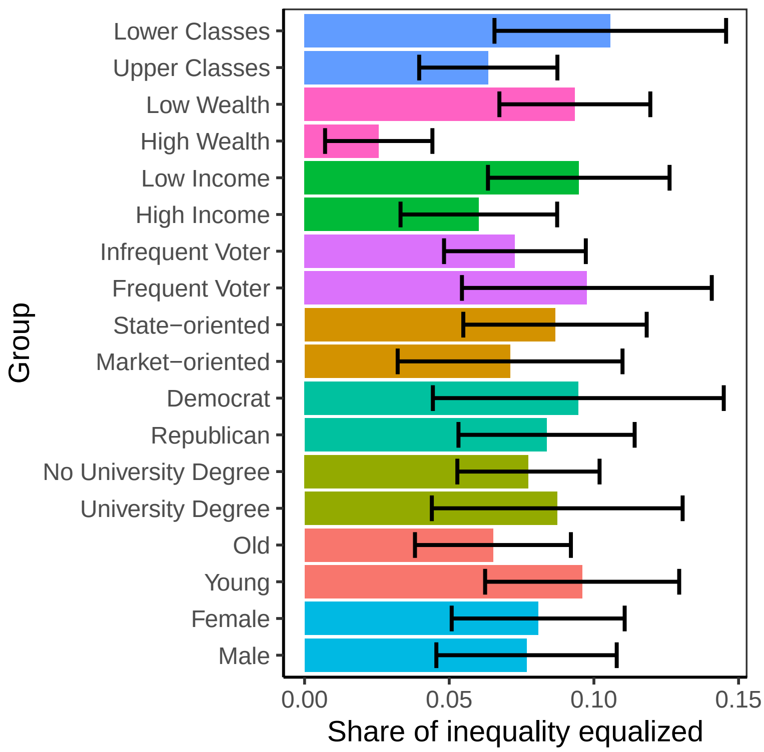



**Figure 1.8.** Average Equalization in Condition Inherited Inequality & Effort by Demographic Group

*Notes:* Shares of inequality equalized for a group are calculated by averaging over the average extent of redistribution in the Inherited Inequality & Effort condition for all spectators in the main sample who belong to the group. 95% confidence intervals around the averages based on standard errors of the mean.

individuals might be more likely to take the benefactors' perspective while for low-wealth individuals the beneficiaries' perspective might be more salient. Similarly, those from the upper classes tend to redistribute less than those from the lower classes.[15]

Yet, there is not much heterogeneity overall; in particular, Democrats and Republicans redistribute to a similar extent on average, and no subgroup equalizes more than $12 of the initial inequality on average. As shown in Figure 1.B.4, Figure 1.B.5 and Figure 1.B.6 in Section 1.B, the patterns in Noninherited Inequality & Effort closely resemble those in Inherited Inequality & Effort displayed here, and heterogeneity in the two Luck conditions is even less pronounced.

15. A potential explanation for heterogeneity along the wealth/socio-economic status dimension could be that individuals take perspectives, endorse fairness ideals, and form beliefs in a self-serving way (Konow, 2000; Rodriguez-Lara and Moreno-Garrido, 2012; Deffains, Espinosa, and Thöni, 2016; Cassar and Klein, 2019; Valero, 2022).

To test formally whether there is heterogeneity in the treatment effects across any of the binary splits in the Inherited Inequality & Effort condition, we run the following OLS regression:

$$\begin{aligned}\theta_{i,\sigma} \;=\; & \alpha + \alpha^D D_i + \alpha_E E_\sigma + \alpha_E^D E_\sigma D_i + \beta II_\sigma + \beta^D II_\sigma D_i \\ & + \beta_E E_\sigma II_\sigma + \beta_E^D E_\sigma II_\sigma D_i + \delta \Delta_\sigma + \epsilon_{i,\sigma}\end{aligned} \tag{1.12}$$

where $D_i$ indicates whether spectator i belongs to a certain sociodemographic subgroup. We cluster standard errors on the spectator level. Figure 1.B.7 in the appendix plots estimates for $\beta^D$ and $\beta_E^D$ by demographic variable, which describe the differences across the sample split in a) the effect of inequality being inherited in the luck domain and b) the "difference-in-differences" effect of inequality being inherited in the effort versus luck domain. Table 1.C.6 and Table 1.C.7 in Section 1.C also report estimated coefficients on other variables. Few estimates for $\beta^D$ and $\beta_E^D$ are significant before controlling for multiple hypothesis testing, and after applying the Benjamini-Hochberg procedure none of the coefficients differs significantly from zero. Hence, resolving the dilemma of meritocracy in favor of those who bequest is common across sociodemographic groups.

To explore whether the distribution of redistribution patterns differs by socioeconomic characteristics, we calculate for each demographic subgroup the distribution over the two-dimensional redistribution patterns $(\tau_{NI}, \tau_{II}) \in$ {(Egalitarian, Egalitarian), (Libertarian, Libertarian), (Meritocrat, Meritocrat), (Meritocrat, Egalitarian)}, which are consistent with our theoretical framework, and a residual type which encompasses all remaining spectators. Figure 1.B.8 in Section 1.B shows the resulting distribution of redistribution patterns by demographic subgroups. There is no notable variation between demographic subgroups. In each subgroup, most spectators can be classified into one of the four main patterns, and in each subgroup more than half of all spectators display a meritocratic redistribution pattern in both dimensions. Using Fisher's exact test, we do not detect any significant differences in the distribution between any two subgroups of the same demographic variable.

### 1.5.5 External Validity

As a next step, we investigate to what extent our experimental measures of redistributional preferences are associated with preferences over real-world policies elicited in the post-experimental questionnaire. Because spectators' average extent of redistribution is highly correlated both within the Luck and Effort domain ($\rho_{\bar{\theta}_{i,NI-L},\bar{\theta}_{i,II-L}} = 0.64$ and $\rho_{\bar{\theta}_{i,NI-E},\bar{\theta}_{i,II-E}} = 0.60$), we apply a factor analysis on the four variables that capture an individual's tendency to redistribute in the four conditions, retaining two factors (eigenvalues equal to 1.11 and 0.91; −0.21 for the third factor). $\bar{\theta}_{i,NI-L}$ and $\bar{\theta}_{i,II-L}$ load heavily on the first factor (0.73 in both cases) but not the second one (0.02 and 0.03). Conversely $\bar{\theta}_{i,NI-E}$ and $\bar{\theta}_{i,II-E}$ load heavily on the second factor (0.69 in both cases) but not the first one (0.02 and 0.04). Hence, we conclude that the first factor captures an individual's preference for redistribution if inequality is based on luck ("Redistribution (Luck)"), while the second factor captures the preference for redistribution if inequality is the result of differential effort ("Redistribution (Effort)").

In the questionnaire, we elicited preferences regarding six inequality-related policies. First, we asked spectators to indicate their preferred maximum marginal income and estate tax rates on scales from 0%–100%. Second, we used 7-point Likert scales to elicit their support for disability insurance, unemployment insurance, and equal opportunity programs, with options ranging from "[the policy] should be significantly reduced" to "significantly extended". Finally, we asked to what extent spectators find intergenerational transmission fair, eliciting responses by means of a 6-point Likert scale from "clearly unfair" to "clearly fair". To facilitate the analysis, we reverse-coded the last variable such that higher values always indicate stronger support for redistribution. Further, we standardized all policy variables and the two factor variables.

Figure 1.9 displays coefficients from OLS regressions of the policy variables on the two factor variables. Without exception, the estimated coefficients are positive,

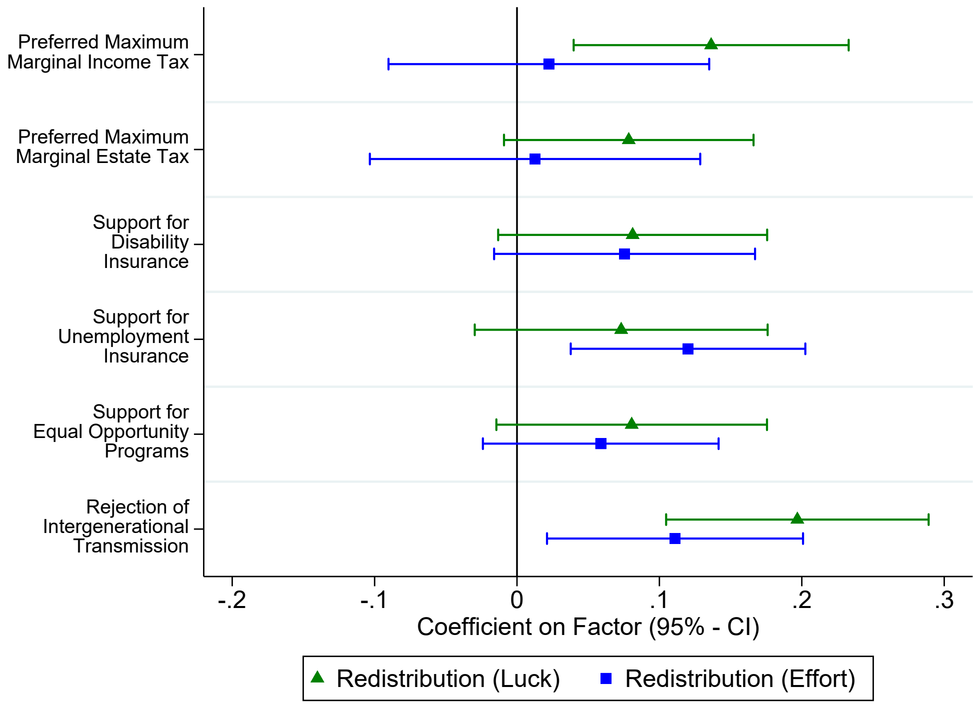



**Figure 1.9.** Association between Experimental Measures and Policy Preferences

*Notes:* This figure plots coefficients from OLS regressions of spectators' (standardized) policy preferences on (standardized) factor variables based on the average extent of redistribution in the four treatment conditions. 95% confidence intervals are based on robust standard errors. The corresponding regressions are reported in Table 1.C.8. Results are based on the main sample.

indicating that more redistribution in the impartial spectator experiment is associated with stronger support for redistributive policies. A 1SD increase in one of the factor variables is often associated with an increase in support for the respective policy by about 0.1SD. Given that recent research has shown that preferences over real-world (redistributive) policies are strongly influenced by factors other than inequality preferences such as views on government efficiency (Stantcheva, 2021), it is perhaps unsurprising that the associations are not too strong. However, for all policy variables, at least one of the two factor variables is significant at the 10%-level. In sum, the results suggest that the experimental measures capture meaningful information about individuals' fairness preferences, and that these preferences are associated with preferences over real-world (redistributive) policies.

### 1.5.6 Validation of Survey Items

Sometimes it may be infeasible to elicit incentivized experimental measures of fairness preferences in a survey. To test whether short nonincentivized survey measures can be employed as substitutes, we asked spectators to what extent they find luck-based and effort-based inequality between two individuals fair. Responses were elicited by means of 6-point Likert scales ranging from "clearly unfair" to "clearly fair".[16]

To assess how closely the experimental and survey measures are related, we run OLS regressions with the average extent of redistribution in either the Noninherited Inequality & Luck or the Noninherited Inequality & Effort condition as the dependent variable and the (standardized) survey measures as the independent variable(s). The results are reported in Table 1.6 and indicate that the experimental measures of redistributional preferences are strongly related to the corresponding survey measure, but not related to the non-corresponding survey measure. Columns (1)-(3) refer to the average extent of redistribution in the Noninherited Inequality & Luck condition. We observe that a 1SD increase in the luck survey measure is associated with a decrease in the average extent of redistribution by almost 15%p. In contrast, there is no association at all between the experimental measure for this condition and the effort survey measure. Conversely, focusing on the Noninherited Inequality & Effort case in columns (4)-(6), a 1SD increase in the effort survey measure is associated with a $6-7$%p decrease in the average extent of redistribution, but there is no association between the experimental measure for this condition and the luck survey measure. These observations are corroborated by the fact that at least 15% of the variance

16. The survey questions asked spectators to complete the sentences "If one person receives more than another due to having better luck, I find that …" and "If one person receives more than another due to exerting higher effort, I find that …" by selecting the option on the Likert scale that corresponded most closely to their view. Figure 1.B.9 in Section 1.B show cumulative distribution functions for the two survey questions.

**Table 1.6.** Association between Experimental and Survey Measures of Redistributional Preferences

| | $\bar{\theta}_{i,NI-L}$ | | | $\bar{\theta}_{i,NI-E}$ | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Luck Survey Measure | −0.148*** | | −0.148*** | 0.004 | | 0.003 |
| | (0.017) | | (0.017) | (0.010) | | (0.010) |
| Effort Survey Measure | | 0.008 | 0.006 | | −0.067*** | −0.066*** |
| | | (0.016) | (0.018) | | (0.018) | (0.018) |
| Constant | 0.799*** | 0.799*** | 0.799*** | 0.048*** | 0.048*** | 0.048*** |
| | (0.016) | (0.017) | (0.016) | (0.008) | (0.008) | (0.008) |
| Observations | 437 | 437 | 437 | 437 | 437 | 437 |
| $R^2$ | 0.172 | 0.000 | 0.172 | 0.000 | 0.147 | 0.147 |

*Notes:* This table reports results from OLS regressions of the average extent of redistribution in the Noninherited Inequality & Luck ($\bar{\theta}_{i,NI-L}$) and Noninherited Inequality & Effort ($\bar{\theta}_{i,NI-E}$) conditions on the respective (standardized) survey measures. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

in the average extent of redistribution is explained if the regression includes the "right" survey measure, but none of the variance is explained if only the "wrong" survey measure is included as a regressor. Overall, our results suggest that if researchers have to economize on survey content these nonincentivized survey measures constitute decent alternatives to elicit fairness preferences and even allow to differentiate between different sources of inequality.

## 1.6 Conclusion

Human beings tend to more altruistic toward their family members, friends, and compatriots than toward non-relatives, strangers and foreigners (Bernhard, Fischbacher, and Fehr, 2006; Cappelen, Enke, and Tungodden, 2022). In many instances the underlying relationships are accidental; for example, we do not choose to which parents or in which country we are born. In meritocratic societies where inequality is accepted if it is based on factors within individuals' control but rejected if it is based on factors outside individuals' control, this creates a fundamental dilemma: unequal outcomes between individuals who differentially profit from other people's efforts are at the same time within the benefactors' control (and therefore just) but outside the beneficiaries' control (and therefore unjust). This paper studied US citizens' fairness preferences in situations with such inherited inequality and how they deal with this dilemma.

Our results show that most US citizens prioritize the benefactors' efforts and accept inherited inequality, which can help to explain why many people accept high levels of inequality and unequal starting positions within and across societies. It is not that they find it fair that some people have better opportunities than others; rather, they weigh this concern against another—in their view stronger—fairness argument. For example, creating equal opportunities among children requires preventing parents from channeling extra resources to their children, even if they themselves earned them fairly. When meritocrats have to decide whether to accept unequal opportunities or prevent families or friends from endowing their loved ones with extra endowments, our results suggest that they choose the former.

Since we find that individuals clearly prioritize rewarding the benefactors' efforts over equalizing payoffs between the non-working beneficiaries when facing the dilemma of meritocracy, a natural avenue for future research is to explore how much the decision environment has to be tweaked for spectators to redistribute more. Our setup is ideally suited to do so because it admits controlled variation in a variety of dimensions.

One potentially relevant dimension is the relationship between benefactor and beneficiary, which varies between outside-the-lab contexts. For example, people usually bequest their resources to their children, and the parent-child relationship is usually stronger than the relationship between friends (Cappelen, Enke, and Tungodden, 2022). In light of our finding that spectators tend to redistribute less if they think that workers tend to prioritize their own friends more strongly, it seems unlikely that the results would differ if we had used family ties instead of friendships, where redistribution levels are already low. Instead, redistribution in the friends-case likely poses an upper bound to redistribution in the family case. Still, spectators might view kinship differently from friendships because people can choose their friends but not their kin. To examine this possibility, researchers could combine our experimental design with a subject sample containing pairs of relatives.

The size of the stakes involved constitutes a second dimension that might be relevant for fairness judgments. High stakes may not only induce individuals to make considerate decisions but, in the context of redistribution, also call into play different motivations such as taking into account individuals' needs (Konow, 2001). Further, employing high stakes may also enable researchers to study preferences over more nuanced (e.g., progressive) redistribution schemes. While the correlation between spectators' behavior in our experiment and their policy preferences indicates that a lot can be learned also from small-stakes settings, it might be worthwhile to study how the stake size affects the relevance of different fairness motives and overall fairness judgments.

Third, our Effort and Luck treatments make it very clear that the initial distribution is either exclusively determined by workers' relative efforts or by luck, whereas resource distributions are usually determined by a combination of the two that is hard to disentangle. Recent research has documented in the context of noninherited inequality that if inequality is based on both effort and luck, this affects redistribution behavior in a non-trivial way. For example, spectators prioritize rewarding effort when the relative contribution of effort and luck can be decomposed (Cappelen and Tungodden, 2017), but uncertainty induces meritocrats to behave in a more egalitarian way (Cappelen et al., 2022). Similarly, uncertainty allows individuals to form biased beliefs about the source of inequality (Konow, 2000; Rodriguez-Lara and Moreno-Garrido, 2012; Deffains, Espinosa, and Thöni, 2016; Cassar and Klein, 2019; Valero, 2022). Hence, it might be interesting to study how uncertainty about the source of inequality affects preferences for redistribution in the context of inherited inequality.

Fourth, individuals may not only inherit differential amounts of resources that can be consumed but also differential opportunities to generate resources themselves. Some papers investigate preferences for redistribution under unequal opportunities, albeit in settings where those unequal opportunities arise exogenously (Eisenkopf, Fischbacher, and Föllmi-Heusi, 2013; Alesina, Stantcheva, and Teso, 2018; Andre, 2022; Schwaiger et al., 2022). Our setup could easily be extended to accommodate the inheritance of unequal opportunities by introducing a second production stage in which the beneficiaries' returns to effort depend on their benefactors' efforts in the first production stage. This would introduce a dilemma similar to the one studied in this paper because a meritocrat should reject unequal opportunities but welcome that higher effort in the first stage pays off for beneficiaries in the second stage, leading to a very different decision problem for individuals making fairness judgments as compared to those in the papers mentioned above.

Finally, we have provided suggestive evidence for a potential mechanism behind individuals' fairness judgments in the context of inherited inequality. Our observations — and also the results from Cohen, Maltz, and Ofek-Shanny (2022) — are consistent with the idea that individuals determine entitlements based on the benefactors' merits and then try to take into account the benefactors' preferences over resource distributions between potential beneficiaries when making fairness judgments. Devising a causal test of this mechanism seems to be a promising endeavor.

## Appendix 1.A Data Quality

In this section, we detail how we tried to promote high-quality responses in the spectator survey and report various data quality checks. The data reveal that a) very few spectators fail attention checks, b) the vast majority states that the instructions were comprehensible, c) spectators make few errors on control questions, d) most spectators write detailed and thoughtful responses to open-ended questions, and e) few spectators perceive the survey to have been biased in either political direction.

**Attention Checks.** The survey features two attention checks, and participants are informed on the first page that they will be rejected if they fail both of them. In line with Prolific's attention check policy, the first attention check instructs subjects to select prespecified options, and the second attention check is a nonsensical question for which only two options are objectively correct. Attention checks are placed strategically: one is administered right at the start of the survey, and the other one is administered as part of the policy preferences questionnaire and resembles the other questions at first glance. None of the 543 subjects who completed the spectator survey failed both attention checks, such that we do not have to exclude anyone in the main sample to follow our pre-analysis plan. Generally, few spectators failed attention checks at all: among the 543 spectators in the main sample, 2 failed the first attention check, and 15 failed the second attention check. Considering only the 437 spectators in the restricted sample (see Section 1.4.2), only one failed the first attention check, and 11 failed the second attention check.

**Comprehensibility.** We attach great importance to not confronting spectators with walls of text. For example, we introduce them to each condition of the earnings stage and how they can make their redistribution decisions with the help of individual slideshows. Each slideshow displays graphical representations of the different steps in the earnings stage with only minimal text, and spectators can go back and forth within each slideshow. The slideshow and the combination of visual and text information are designed to make the survey as engaging and easy to digest as possible.

At the end of the survey, we ask spectators how comprehensible they find the instructions. On a 7-point Likert scale, subjects can choose options from "not comprehensible at all" to "perfectly comprehensible". For spectators in the restricted sample, Figure 1.A.1 shows the distribution of the responses (the figure for the main sample looks very similar). We observe that spectators judge the instructions very favorably. The vast majority (58%) say that the instructions were "perfectly comprehensible," and 89% assess the instructions as at least "fairly comprehensible." It is particularly reassuring that less than 1% of the spectators perceive the instructions as "not very comprehensible," and no one chooses the lowest two options.

**Control Questions.** To check more directly whether spectators understand the instructions, they have to answer two control questions each after they were introduced to a particular type of situation by means of the slideshow. They can proceed to the corresponding block of decisions only if they answered both questions correctly; otherwise, they are referred to the slideshow again. Control questions ask about the most crucial features of the situation: whether workers worked for

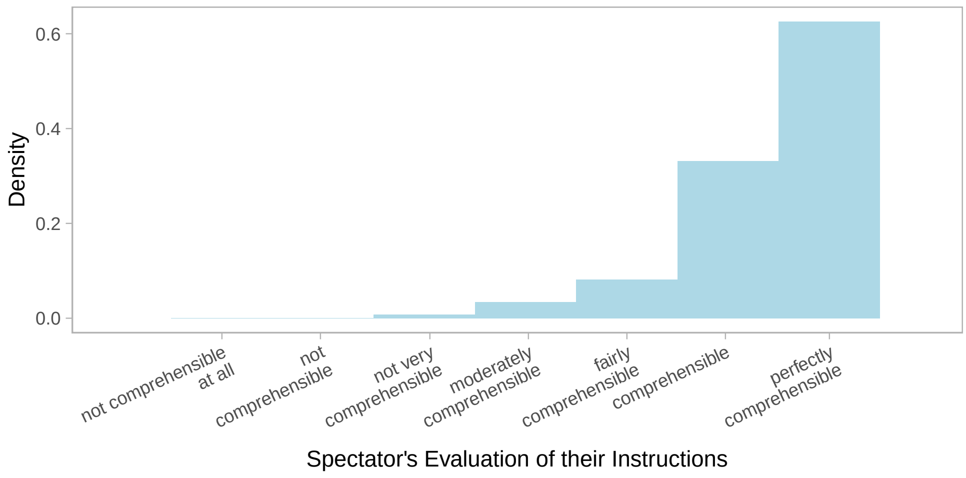



**Figure 1.A.1.** Spectators' Assessment of the Instructions

*Notes:* Histogram showing how spectators in the restricted sample chose to complete the sentence "Overall, I found the instructions ..." on a 7-point Likert scale from "not comprehensible at all." to "perfectly comprehensible.".

themselves or friends and whether the initial allocation of the $10 would be based on a random draw or the relative number of completed tasks. In total, each spectator responds to 8 control questions. Figure 1.A.2 depicts a histogram of the total number of errors spectators in our sample made. We observe that most spectators

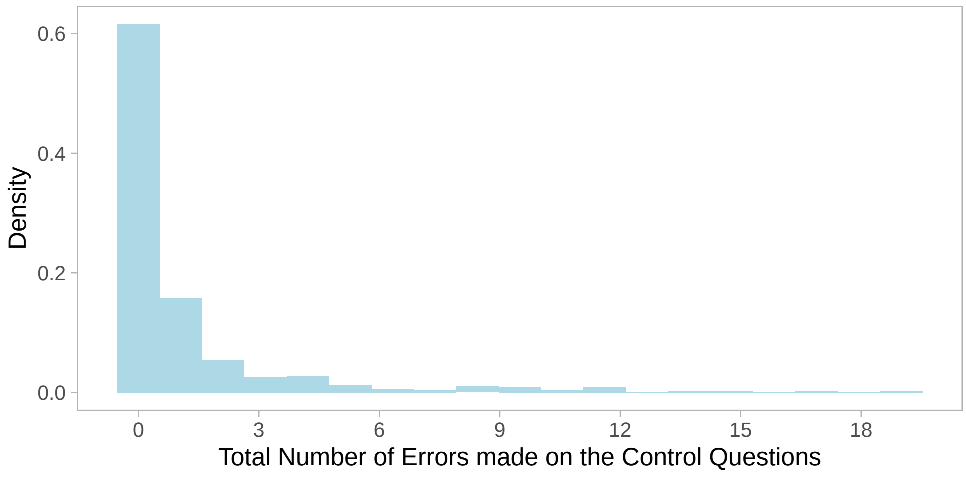



**Figure 1.A.2.** Control Question Errors

*Notes:* Histogram of the total number of errors that spectators in the restricted sample made when responding to the 8 control questions.

made few errors, which indicates that they usually understood the instructions well. About 65% of spectators made no error, and only about 13% made more than 2 errors in total.

**Open-Ended Questions.** The spectator survey features several open-ended questions. After spectators have made all redistribution decisions within a particular block, we ask them to describe their considerations regarding these decisions. Further, at the end of the survey, subjects can leave a final comment on the general topic, the instructions, whether they experienced difficulties or anything else they have on their mind. Most open-ended responses are quite detailed and thoughtful. Only one spectator in the restricted sample (four spectators in the main sample) did not write any open-ended response during the study, suggesting that spectators generally put considerable effort into the study.

Figure 1.A.3 summarizes responses in four word clouds, one for each treatment. To generate these word clouds, we remove all numbers from the open-ended responses, transform all words to lowercase and remove punctuation and stop words. Finally, we reduce all words to their base word (stem). The size of words in Figure 1.A.3 indicates the frequency with which that word was used. The term "work" was among the most often used terms in all conditions, consistent with the large share of meritocrats in our sample. In the Luck conditions, the term "equal" was also used very frequently, while it was nearly absent in the Effort conditions. Similarly, the term "friend" belongs to the most commonly used terms in the Inherited Inequality conditions but is rarely used in the Noninherited Inequality treatments. This suggests that subjects understood the conditions and gave thoughtful explanations.

Figure 1.A.4, Figure 1.A.5 and Figure 1.A.6 show the frequencies of explanations that spectators give for their decisions by explanation category. Table 1.C.3 provides an overview of all categories with definitions and examples. Figure 1.A.4 shows that, consistent with their redistribution decisions, most spectators state to redistribute in the Noninherited Inequality & Effort condition based on the workers' efforts. Figure 1.A.5 reveals that most spectators rationalize their behavior in the Noninherited Inequality & Luck condition with a preference for a distribution based on effort too. However, many also mention that they find distributions based on luck unfair, while a few argue that the random allocation of resources is a fair method of distribution. Similarly, Figure 1.A.6 shows that many spectators justify their behavior in the Inherited Inequality & Luck treatment with arguments based on luck. Moreover, many spectators specifically refer to the effort of the workers or their friends. Hence, the explanations spectators give for their decisions correspond reasonably to the treatment conditions, which suggests that they had a good understanding of the study setup.

Finally, Figure 1.A.7 shows a word cloud of final comments spectators could make at the end of the survey. Again, to generate this word cloud, we remove all numbers from the open-ended responses, transform all words to lowercase and remove punctuation and stop words. Finally, we stem all words. Most comments are positive. Many spectators mention that they found the study interesting and understandable.


**(a)** Noninherited & Effort

**(b)** Inherited & Effort

**(c)** Noninherited & Luck

**(d)** Inherited & Luck

**Figure 1.A.3.** Word clouds of terms subjects used to explain their considerations when making redistribution decisions by treatment condition.

**Political Bias.** For surveys on highly politicized topics such as redistribution, it may be particularly important to phrase instructions and questions in a neutral way. We tried to keep this caveat in mind when we decided on the formulations used in the survey. Additionally, we ask subjects at the end of the survey whether they have the impression that the survey is biased toward a particular political stance, using a 7-point Likert scale with options from "strong left bias" to "strong right bias." Figure 1.A.8 displays how spectators' responses in the restricted sample are distributed (again, the figure for the main sample looks very similar). Less than 5% of the spectators perceive a strong bias in either direction. About 23% perceive a left-wing bias of any strength, whereas about 6% perceive a right-wing

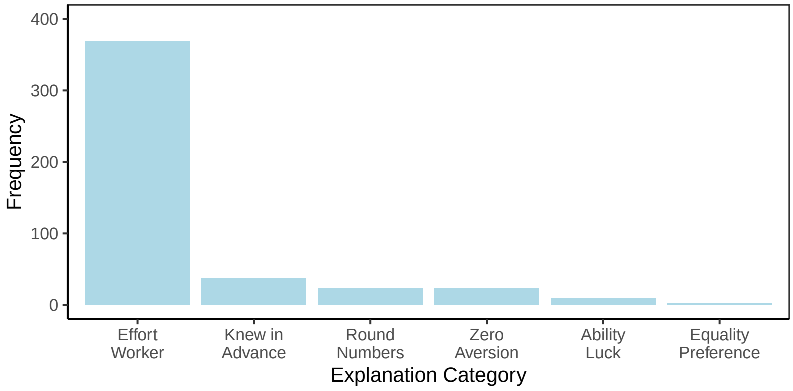



**Figure 1.A.4.** Spectators' Explanations for their Decisions in NONINHERITED INEQUALITY & EFFORT

*Notes:* This figure displays the frequency of explanations spectators gave for their redistribution decisions in NONINHERITED INEQUALITY & EFFORT by explanation category. Results are based on up to 3 arguments made by 432 spectators from the restricted sample. We included up to 3 arguments per spectator.

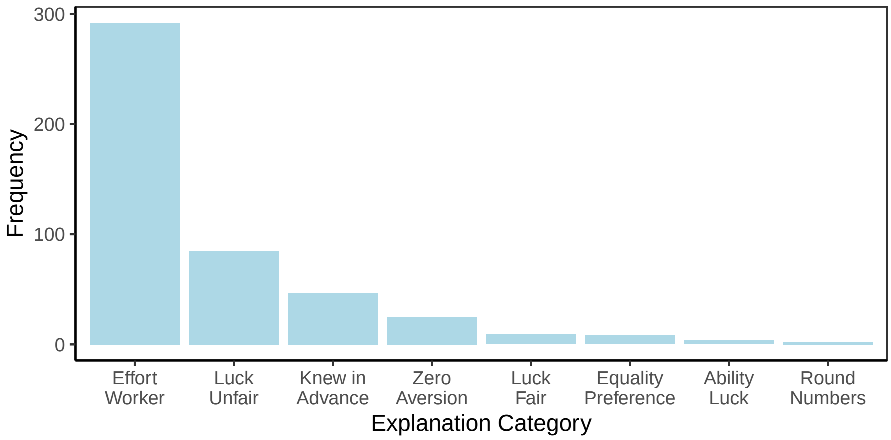



**Figure 1.A.5.** Spectators' Explanations for their Decisions in NONINHERITED INEQUALITY & LUCK

*Notes:* This figure displays the frequency of explanations spectators gave for their redistribution decisions in NONINHERITED INEQUALITY & LUCK by explanation category. Results are based on up to 3 arguments made by 435 spectators from the restricted sample. We included up to 3 arguments per spectator.

bias of any strength. More than 70% of the spectators in the restricted sample respond with "No or almost no bias," which is remarkable given that the theme of the survey is redistribution.

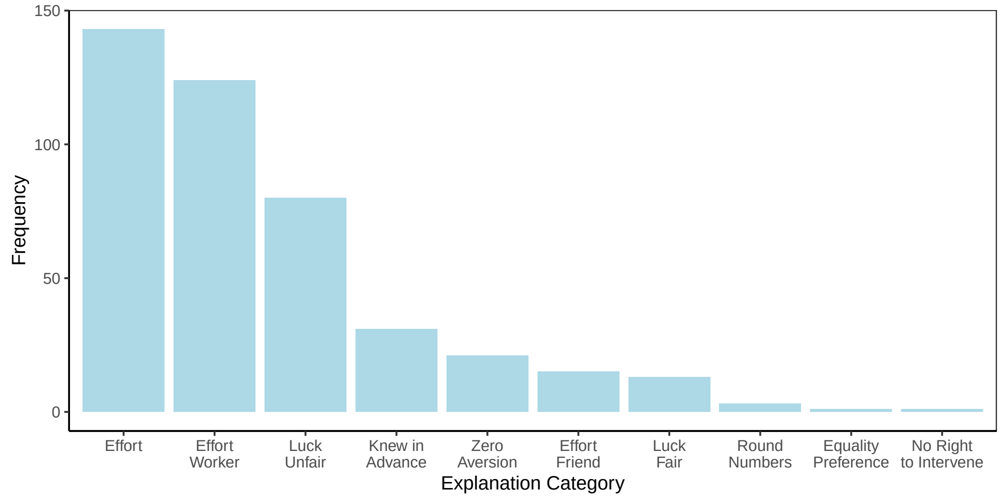



**Figure 1.A.6.** Spectators' Explanations for their Decisions in INHERITED INEQUALITY & LUCK

*Notes:* This figure displays the frequency of explanations spectators gave for their redistribution decisions in INHERITED INEQUALITY & LUCK by explanation category. Results are based on up to 3 arguments made by 432 spectators from the restricted sample. We included up to 3 arguments per spectator.


**Figure 1.A.7.** Word Cloud of Final Comments

*Notes:* A word cloud relating to final comments spectators could make at the end of the survey.

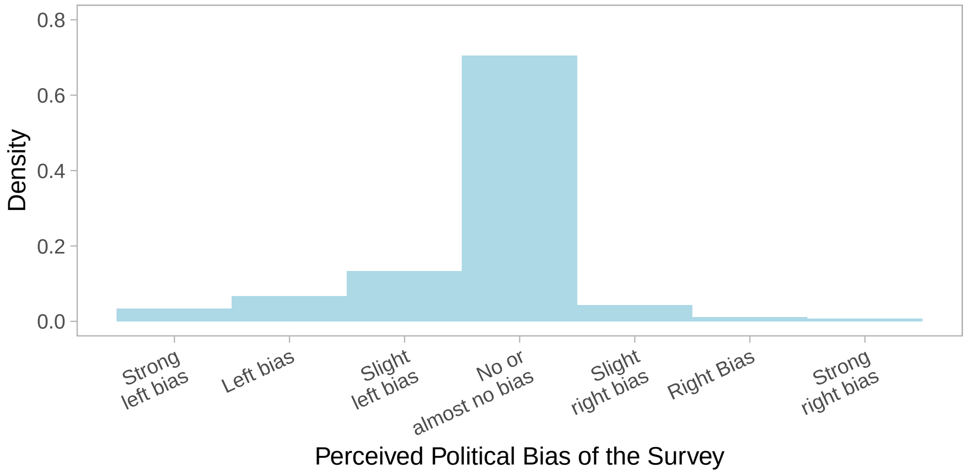



**Figure 1.A.8.** Spectators' Perception of the Survey's Political Bias

*Notes:* Histogram of how subjects in the restricted sample respond to the question "Do you think this survey was biased toward a certain political stance?", asked at the end of the survey using a 7-point Likert scale from "strong left bias" to "strong right bias".

## Appendix 1.B Supplementary Figures

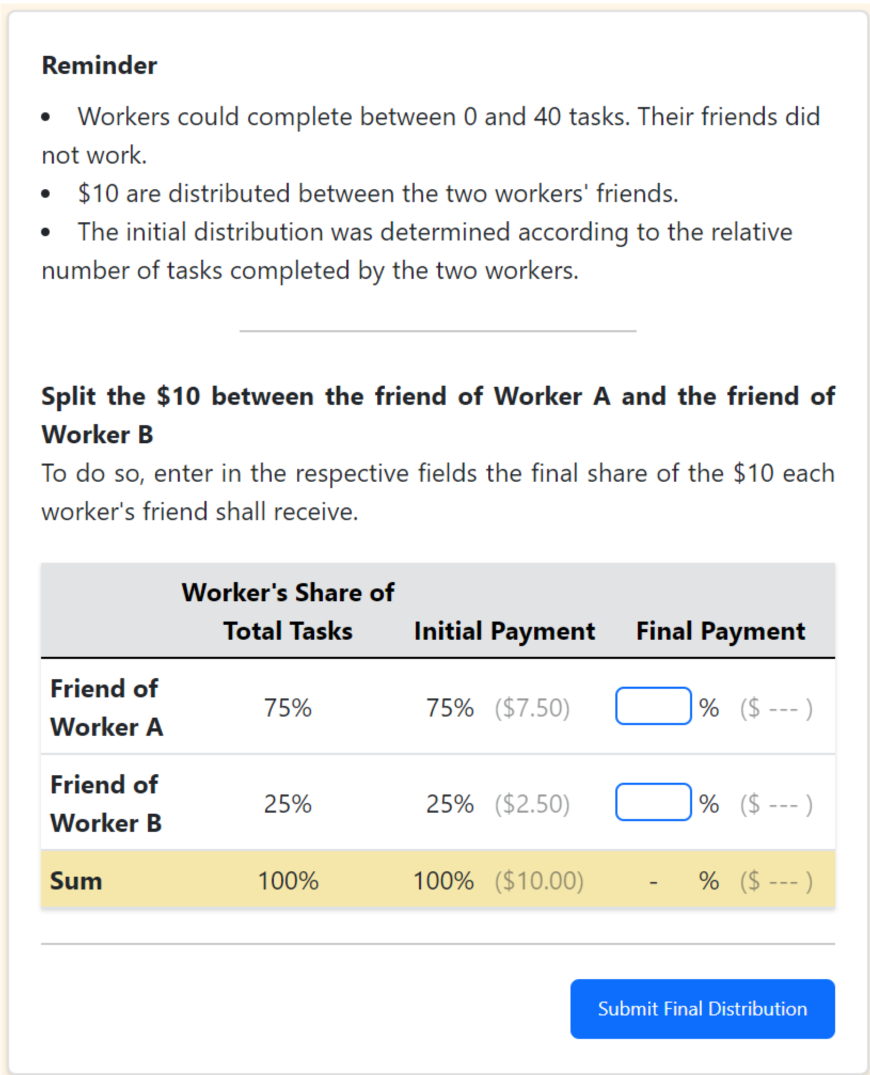
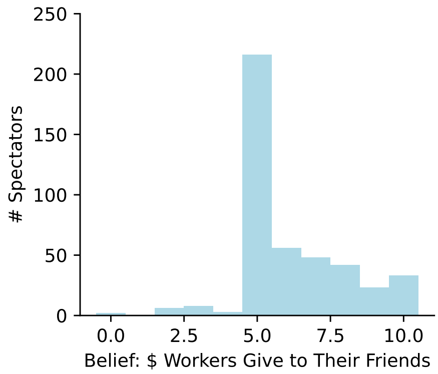



**Figure 1.B.1.** Screenshot of the Decision Screen for Spectator's Redistribution Decisions

*Notes:* This decision screen corresponds to the Inherited Inequality & Merit condition. The decision screens for the other conditions had the same structure.


**(a)** Self vs. Other Worker




**(b)** Own Friend vs. Other Worker's Friend

**Figure 1.B.2.** Spectators' Beliefs about Workers' Preferred Distributions

*Notes:* Figure 1.B.2a displays a histogram of spectators' incentivized beliefs about the share of the $10 workers on average keep for themselves when they are asked how they would like to distribute $10 between themselves and the worker they are matched to in the first incentivized dictator decision. Figure 1.B.2b displays a histogram of spectators' incentivized beliefs about the share of the $10 workers on average give to their own friends when they are asked how they would like to distribute $10 between their own friend and the friend of the worker they are matched to in the second incentivized dictator decision.

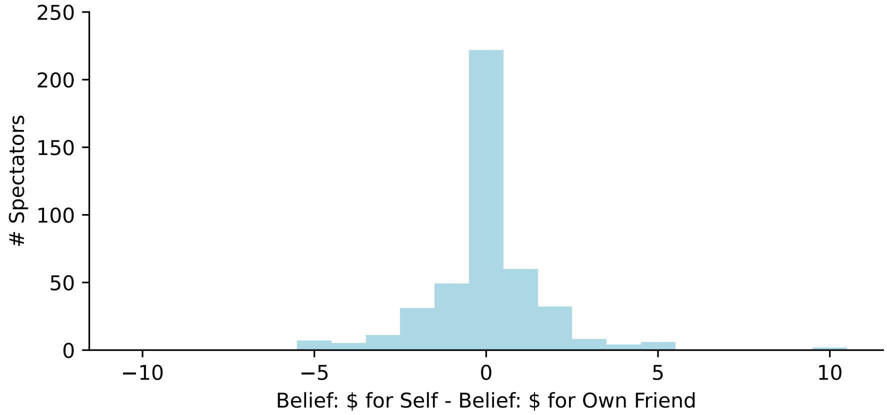



**Figure 1.B.3.** Differences in Spectators' Beliefs about Workers' Preferred Distributions

*Notes:* This figure displays a histogram of the individual differences in spectators' beliefs about workers' preferred distributions in the dictator decisions for a) themselves vs. the worker they are matched to and b) their own friend vs. the friend of the worker they are matched to. For example, if a spectator indicated a belief that workers on average keep $8 for themselves when they are asked how they would like to distribute $10 between themselves and the worker they are matched to, and that workers on average give $7 to their own friend when they are asked how they would like to distribute $10 between their own friend and the friend of the worker they are matched to, this would yield a difference of $1.

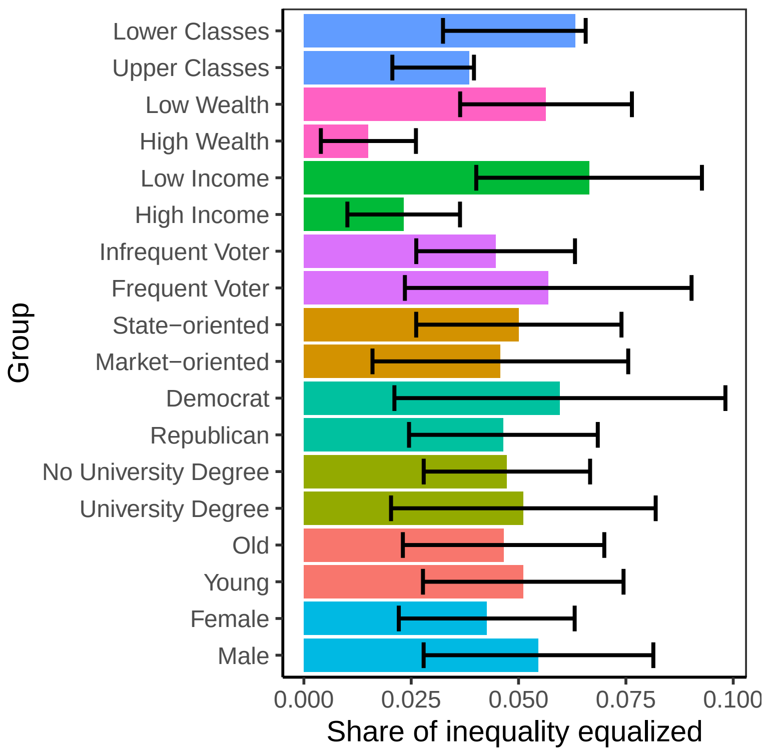



**Figure 1.B.4.** Average Equalization in Condition NONINHERITED INEQUALITY & EFFORT by Demographic Group

*Notes:* Shares of inequality equalized for a group are calculated by averaging over the average extent of redistribution in the Noninherited Inequality & Effort condition for all spectators in the main sample who belong to the group. 95% confidence intervals around the averages based on standard errors of the mean.

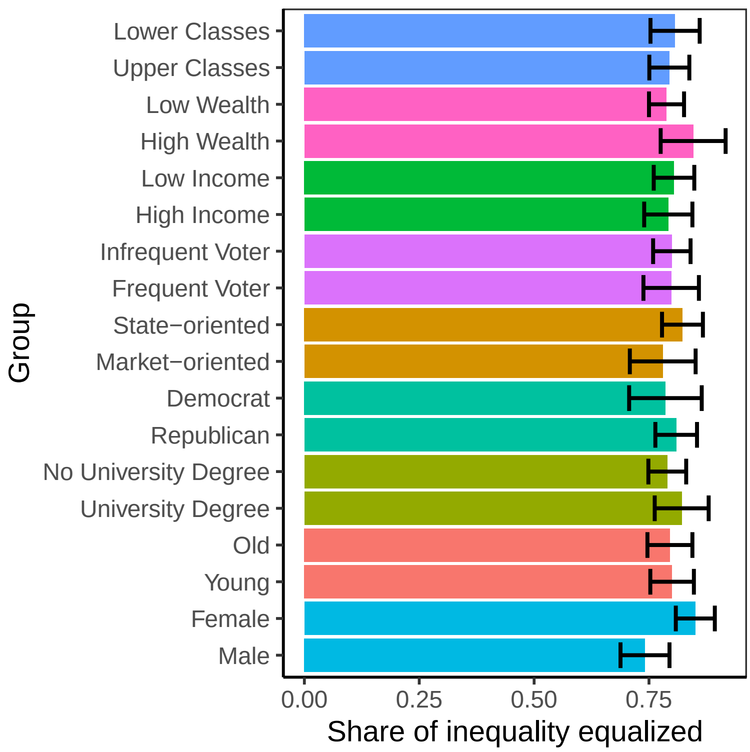



**Figure 1.B.5.** Average Equalization in Condition NONINHERITED INEQUALITY & LUCK by Demographic Group

*Notes:* Shares of inequality equalized for a group are calculated by averaging over the average extent of redistribution in the NONINHERITED INEQUALITY & LUCK condition for all spectators in the main sample who belong to the group. 95% confidence intervals around the averages based on standard errors of the mean.

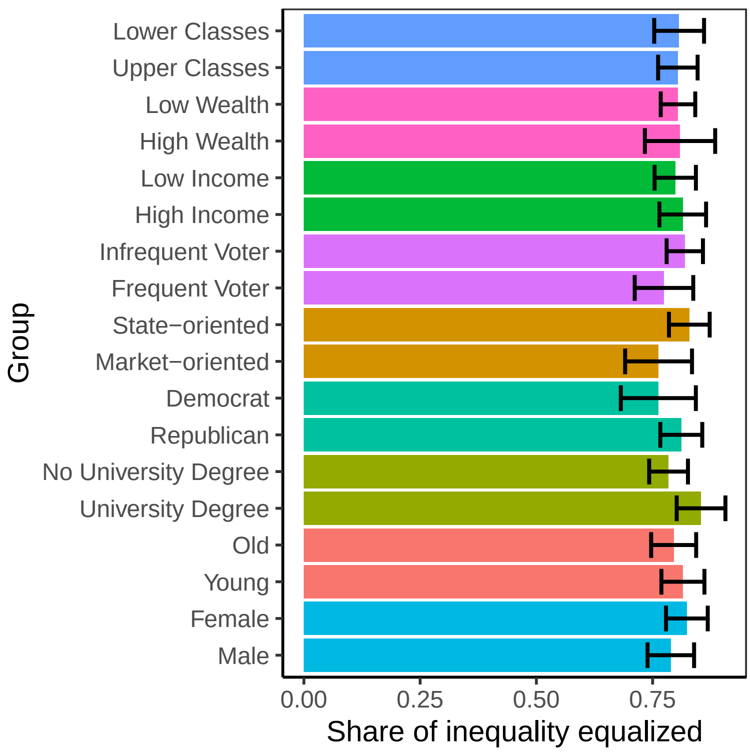



**Figure 1.B.6.** Average Equalization in Condition Inherited Inequality & Luck by Demographic Group

*Notes:* Shares of inequality equalized for a group are calculated by averaging over the average extent of redistribution in the Inherited Inequality & Luck condition for all spectators in the main sample who belong to the group. 95% confidence intervals around the averages based on standard errors of the mean.

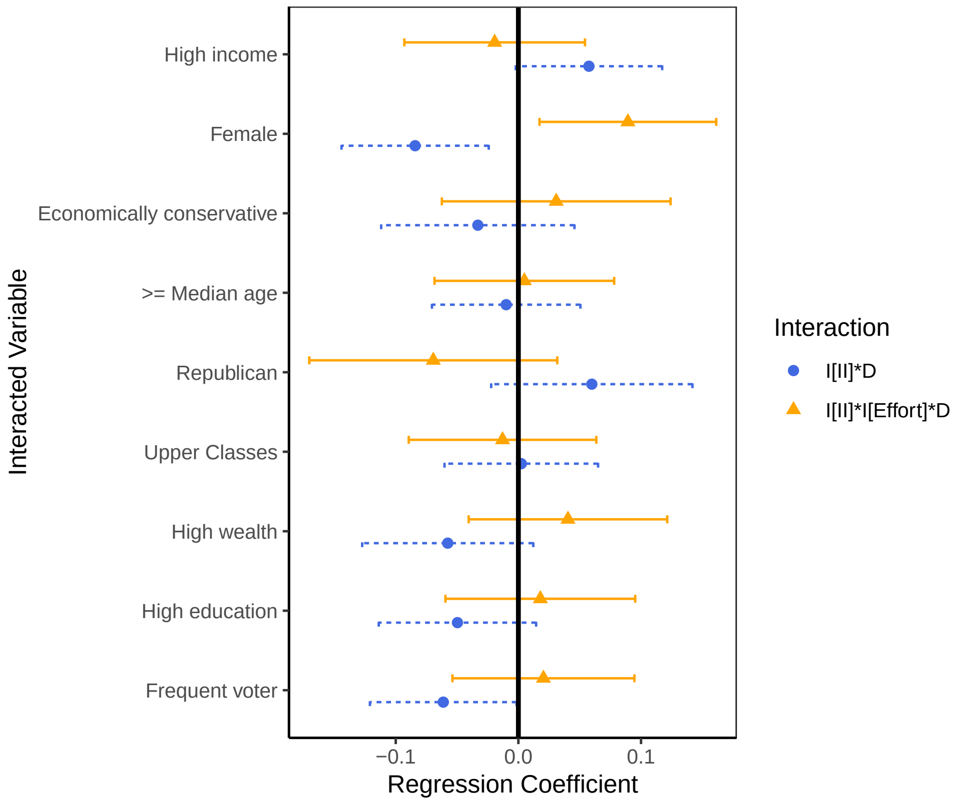



**Figure 1.B.7.** Heterogeneity in Treatment Effects between Demographic Groups

*Notes:* This figure shows coefficients and 95% confidence intervals. The vertical axis shows demographic variables. These variables were interacted with two other terms in Equation 1.12. The blue points show the coefficient on the interaction term of each demographic variable ($D_i$) with the indicator for the INHERITED INEQUALITY conditions ($II_o$). The orange points visualize the interaction of $D_i$ with INHERITED INEQUALITY and an indicator for the EFFORT conditions ($E_o$). Results are based on the main sample.

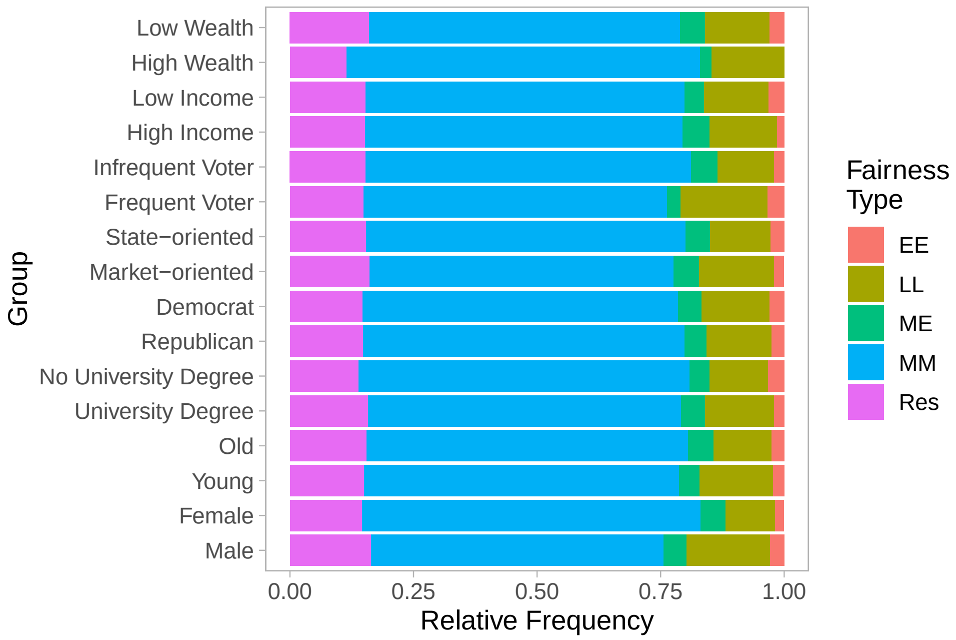
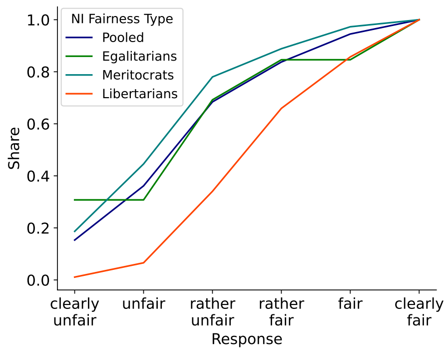
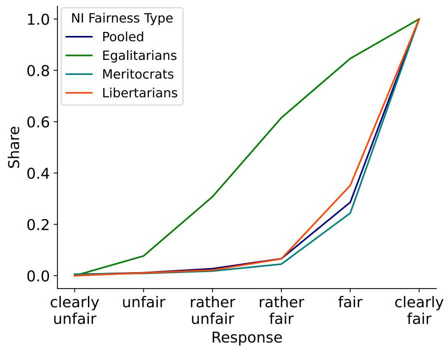



**Figure 1.B.8.** Distribution of Fairness Types by Demographic Group

*Notes:* The vertical axis depicts demographic subgroups. Colors indicate 5 fairness types based on redistribution decisions under noninherited and inherited inequality. The horizontal axis shows the relative frequency with which these fairness types appear within the demographic subgroups. The fairness type ME stands for spectators who are classified as meritocrats under noninherited inequality and as egalitarians under inherited inequality. Likewise, EE, LL, and MM stand for egalitarian/egalitarian, liberterian/libertarian, and meritocrat/meritocrat, respectively. All spectators who do not belong to either of these types are summarized in the residual category “Res”.




**(a)** Luck




**(b)** Effort

**Figure 1.B.9.** CDFs of the Responses to the Inequality Acceptance Survey Measures

*Notes:* This figure shows cumulative redistribution functions of spectators' responses to the inequality acceptance survey questions. Figure 1.B.9a corresponds to the question "If one person receives more than another due to having better luck, I find that ..." and Figure 1.B.9b corresponds to the question "If one person receives more than another due to exerting higher effort, I find that ...". Included are the responses of spectators in the restricted sample.

## Appendix 1.C Supplementary Tables

**Table 1.C.1.** Descriptives and Representativeness

| | Spectator Sample | | US Population |
|---|---|---|---|
| | Full/Main Sample | Restricted Sample | |
| Female | 50.6 % | 50.6 % | 50.5 % |
| Age Groups | | | |
| 18-19 | 1.5 % | 1.6 % | 3.4 % |
| 20-24 | 9.9 % | 8.5 % | 8.3 % |
| 25-29 | 11.7 % | 9.5 % | 8.6 % |
| 30-34 | 9.3 % | 8.8 % | 8.9 % |
| 35-39 | 10.8 % | 9.5 % | 8.7 % |
| 40-44 | 8.9 % | 9.0 % | 8.3 % |
| 45-49 | 6.9 % | 7.2 % | 7.7 % |
| 50-54 | 8.4 % | 8.8 % | 8.1 % |
| 55-59 | 10.8 % | 11.3 % | 8.2 % |
| 60-64 | 9.1 % | 10.6 % | 8.4 % |
| 65-69 | 7.1 % | 8.5 % | 7.1 % |
| 70-74 | 3.2 % | 3.7 % | 6.0 % |
| 75-79 | 2.2 % | 2.5 % | 3.8 % |
| 80-84 | 0.4 % | 0.5 % | 2.4 % |
| 85+ | 0.0 % | 0.0 % | 2.3 % |
| Education Groups | | | |
| No Highschool | 0.4 % | 0.2 % | 10.6 % |
| High School Diploma Equivalent | 30.4 % | 30.0 % | 45.6 % |
| Bachelor's or Associate's Degree | 51.7 % | 51.3 % | 30.0 % |
| Master's Degree or Higher | 17.5 % | 18.5 % | 13.8 % |
| Income Groups | | | |
| $< \$34,000$ | 26.7 % | 27.5 % | 25.0 % |
| $\$34,000 - \$68,000$ | 30.0 % | 30.9 % | 25.0 % |
| $\$68,000 - \$125,000$ | 30.0 % | 28.4 % | 25.0 % |
| $> \$125,000$ | 13.3 % | 13.3 % | 25.0 % |
| Race | | | |
| White | 72.6 % | 73.5 % | 75.8 % |
| Black | 12.6 % | 12.9 % | 13.6 % |
| Asian | 7.2 % | 6.3 % | 6.1 % |
| Mixed | 4.0 % | 3.7 % | 2.9 % |
| Other | 3.6 % | 3.5 % | 1.6 % |
| Observations | 543 | 437 | |

*Notes:* This table reports descriptive statistics for our spectator sample and how they compare to the US general population. The survey company did not provide us with information on a spectator's age in two cases, gender in one case, and ethnicity in 13 cases. Shares in these groups are relative to the sample of spectators for which this information is available. Data for the US population are obtained from the 2021 American Community Survey, S0101 Age and Sex, via the United States Census Bureau (https://data.census.gov/table?tid=ACSST1Y2021.S0101, last accessed: January 9th, 2023; age and gender), the 2021 American Community Survey, S1501 Educational Attainment, via the United States Census Bureau (https://data.census.gov/table?tid=ACSST1Y2021.S1501, last accessed: January 9th, 2023; education groups), the United States Census Bureau QuickFacts table (https://www.census.gov/quickfacts/fact/table/US/PST045221, last accessed: January 16th, 2023; race), and https://dqydj.com/2020-household-income-percentile-calculator/, last accessed: January 9th, 2023; household income groups. Population data on educational attainment is based on citizens aged 25 years or older because for younger citizens the reported education groups did not match those we used in our survey. Likewise, we used the data on household income referenced above because they provided quartile household income group thresholds which we used in our survey.

**Table 1.C.3.** Categories of Explanations That Spectators Give for Their Redistribution Decisions

| Category Name | Argument Made by Spectator | Example |
|---|---|---|
| Effort | The final distribution should be based on the relative amount of tasks done (The spectator does not mention whether he means the tasks done by the workers or the tasks done by their friends). | The money should be based on the percentage of work each one did. |
| Effort Workers | The final distribution should be based on the relative amount of tasks done by the workers. | I made the payment based on the amount of work that each worker produced. It made no difference to me where the money ended up going, I just wanted to make sure that payments were made according to the amount of work produced. |
| Effort Friends | The final distribution should be based on the relative amount of tasks done by the friends. | I think it is fair to split the money evenly between the friends of the participants. They did not do any work. |
| Knew in Advance | All subjects knew the rules of the experiment in advance and agreed by participating. Changing rules after decisions have been made is unfair. | It was an easy task, and all participants were aware of what they were working towards - it would be unethical to change that agreement after the fact. |
| Zero Aversion | Every subject should receive something (of the bonus)/should at least receive a certain amount (e.g., $1). | i tried to be fair and also give 10% to those that completed 0 |
| Round Numbers | Spectator has a preference for round numbers. | i prefer even numbers. even percentages. |
| Ability Luck | Some workers were more able to perform on the task than other workers due to lucky circumstances. | … I did want to move it back closer to an even split a little bit in case one worker had an advantage that made the task easier for them |
| Equality Preference | Money should always be distributed equally (no specific reasons stated). | No matter how much work I do, I think everyone has the right to about the same amount of money. |
| Luck Unfair | Outcomes that result from luck are unfair. | Just because your luck ran out on certain examples shouldn't be a cause to distribute that way |
| Luck Fair | Distributing based on luck is a fair procedure. | A random drawing is about as fair as it gets so I kept the same numbers. The workers just needed to cross their fingers that day. |
| No Right to Intervene | Spectator has no right to intervene in the affairs of others. | … If the Friend was lucky, why should I change things for them so that I make things fair for everyone within my own sense of justice or fairness. I can't play God. I believe it is contingent upon the person who has been lucky to give off his/her/they/their wealth to others who were less fortunate. |
| Exchange | The workers should earn what they worked for and the spectator expects the friends to share with their workers after the study. | … I think the people who did the work deserve to get the outcome they expected. Some of them probably selected a friend who would give them the money. |

**Table 1.C.3 Continued:** Categories of Explanations That Spectators Give for Their Redistribution Decisions

| Category Name | Argument Made by Spectator | Example |
|---|---|---|
| Type of Friend | The worker working for his friend means that the friend is a good person, and a good person should be rewarded. | ... If Bill felt like knocking out a lot of tasks for his friend, who am I to take some of that and give it to James' friend when James did not think his friend was worth it? |
| Friend Not Entitled | The friends did not work for the money. Hence, they are not entitled to receive nay money. | These "friends" should feel lucky to be receiving anything at all. Neither friend is entitled to anything — especially more so for, that which the friend did *not* work for, ze'mself |
| Worker Entitled | The workers worked for the money. Hence, each worker is entitled to the amount he earned through his work. | The participants worked for and earned their share of the money. Even though the friends had no choice, the participants should receive (for their friend) a payment equivalent to how hard they worked |
| Friend Blameless | The friends did not work and are therefore not to blame for the distribution, in contrast to the workers. Hence, it is unfair that one friend gets less than another. | I had to make a decision between honoring the initiative of the workers or the making the receipts more equitable. Since the friends were "blameless" (and unconscious?) regarding the amount of labor involved, I elected to honor that side of the exercise with a 50-50 split |
| Team | Worker and friend are one team. What the team earns should stay with the team. | Even though friends did not work, he is a part of the team regardless and should be paid equally |
| NA | Comment without any explanation for the spectators' decisions. | Now is the time for the communist revolution! No more can these capitalist pigs turn us against one another! Throw off your chains, comrades, and let us create a world where no one goes hungry and we are truly free to pursue our passions! |

**Table 1.C.2.** Treatment Effects on the Extent of Redistribution $\theta_{i,\sigma}$ by Fairness Type

| | Restricted Sample | | | |
|---|---|---|---|---|
| | (1) Pooled | (2) Egalitarians | (3) Meritocrats | (4) Libertarians |
| Effort ($E_\sigma$) | −0.025 | −0.025 | −0.960*** | −0.109*** |
| | (0.036) | (0.038) | (0.006) | (0.018) |
| Inherited ($II_\sigma$) | −0.018 | −0.017 | −0.059*** | 0.268*** |
| | (0.031) | (0.032) | (0.012) | (0.042) |
| Effort ($E_\sigma$) × Inherited ($II_\sigma$) | −0.144 | −0.144 | 0.099*** | −0.232*** |
| | (0.103) | (0.108) | (0.015) | (0.044) |
| Meritocrat | −0.010 | | | |
| | (0.015) | | | |
| Libertarian | −0.850*** | | | |
| | (0.023) | | | |
| Nonclassified | −0.532*** | | | |
| | (0.014) | | | |
| Effort ($E_\sigma$) × Meritocrat | −0.935*** | | | |
| | (0.036) | | | |
| Effort ($E_\sigma$) × Libertarian | −0.083** | | | |
| | (0.040) | | | |
| Effort ($E_\sigma$) × Nonclassified | 0.234*** | | | |
| | (0.036) | | | |
| Inherited ($II_\sigma$) × Meritocrat | −0.042 | | | |
| | (0.034) | | | |
| Inherited ($II_\sigma$) × Libertarian | 0.286*** | | | |
| | (0.052) | | | |
| Inherited ($II_\sigma$) × Nonclassified | −0.071** | | | |
| | (0.031) | | | |
| Effort ($E_\sigma$) × Inherited ($II_\sigma$) × Meritocrat | 0.243** | | | |
| | (0.104) | | | |
| Effort ($E_\sigma$) × Inherited ($II_\sigma$) × Libertarian | −0.088 | | | |
| | (0.112) | | | |
| Effort ($E_\sigma$) × Inherited ($II_\sigma$) × Nonclassified | 0.296*** | | | |
| | (0.103) | | | |
| Initial Inequality ($\Delta_\sigma$) | 0.031** | −0.052 | −0.004 | 0.175*** |
| | (0.014) | (0.101) | (0.012) | (0.045) |
| Constant | 0.977*** | 1.001*** | 0.977*** | 0.084*** |
| | (0.015) | (0.036) | (0.006) | (0.019) |
| Clusters | 437 | 13 | 332 | 91 |
| Observations | 8399 | 249 | 6403 | 1731 |
| $R^2$ | 0.817 | 0.106 | 0.864 | 0.228 |

*Notes:* This table reports results from the same regression equations as Table 1.4 but does not omit coefficients. Standard errors (in parentheses) are clustered on the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table 1.C.4.** Association between Beliefs about Workers Preferences and Redistribution Decisions

| | Noninherited Inequality | | Inherited Inequality | |
|---|---|---|---|---|
| | (1) Luck | (2) Effort | (3) Luck | (4) Effort |
| Guess Self/Other | -0.041 | -0.104* | | |
| | (0.047) | (0.057) | | |
| Guess Own Friend/Other's Friend | | | -0.071 | -0.131** |
| | | | (0.049) | (0.059) |
| Observations | 437 | 437 | 437 | 437 |
| $R^2$ | 0.002 | 0.011 | 0.005 | 0.017 |

*Notes:* This table reports results from OLS regressions of spectators' average extent of redistribution, standardized across spectators but within conditions), on their standardized beliefs about workers' preferred distributions. The coefficients are displayed in Figure 1.7. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table 1.C.5.** Association Between Beliefs and Redistribution Decisions

| | Noninherited Inequality | | | Inherited Inequality | | |
|---|---|---|---|---|---|---|
| | (1) Egalitarians | (2) Meritocrats | (3) Libertarians | (4) Egalitarians | (5) Meritocrats | (6) Libertarians |
| Guess Self/Other | 0.691 | 0.026 | 0.045 | | | |
| | (0.397) | (0.037) | (0.102) | | | |
| Guess Own Friend/Other's Friend | | | | 0.071 | 0.013 | -0.263*** |
| | | | | (0.103) | (0.058) | (0.092) |
| Observations | 13 | 332 | 91 | 13 | 332 | 91 |
| $R^2$ | 0.478 | 0.001 | 0.002 | 0.005 | 0.000 | 0.069 |

*Notes:* In analogy to Table 1.5, this table reports results from OLS regressions of spectators' average extent of redistribution in the two Luck conditions, standardized across spectators of a given (Noninherited Inequality) fairness type and within experimental conditions, on their beliefs about workers' preferred distributions, standardized across spectators of the same fairness type. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table 1.C.6.** Heterogeneity in Treatment Effects by Demographic Group (I)

| | Dependent Variable: Extent of Redistribution $\theta_{i,\sigma}$ | | | |
|---|---|---|---|---|
| | (1)<br>Social Class<br>$D_i$=1 if upper | (2)<br>Wealth<br>$D_i$=1 if high | (3)<br>Income<br>$D_i$=1 if high | (4)<br>Education<br>$D_i$=1 if high |
| Effort ($E_\sigma$) | −0.742*** | −0.728*** | −0.735*** | −0.767*** |
| | (0.033) | (0.023) | (0.026) | (0.033) |
| Inherited ($II_\sigma$) | 0.019 | 0.031* | −0.004 | 0.055** |
| | (0.026) | (0.017) | (0.021) | (0.027) |
| Effort ($E_\sigma$) × Inherited ($II_\sigma$) | 0.030 | 0.015 | 0.031 | 0.010 |
| | (0.032) | (0.021) | (0.024) | (0.032) |
| $D_i$ | −0.028 | 0.037 | −0.009 | −0.003 |
| | (0.054) | (0.039) | (0.035) | (0.037) |
| Effort ($E_\sigma$) × $D_i$ | 0.012 | −0.104** | −0.030 | 0.029 |
| | (0.062) | (0.041) | (0.039) | (0.041) |
| Inherited ($II_\sigma$) × $D_i$ | 0.033 | −0.058 | 0.058* | −0.050 |
| | (0.046) | (0.036) | (0.030) | (0.033) |
| Effort ($E_\sigma$) × Inherited ($II_\sigma$) × $D_i$ | −0.068 | 0.040 | −0.019 | 0.018 |
| | (0.060) | (0.041) | (0.038) | (0.039) |
| Initial Inequality ($\Delta_\sigma$) | 0.062** | 0.079*** | 0.079*** | 0.079*** |
| | (0.025) | (0.019) | (0.019) | (0.019) |
| Constant | 0.796*** | 0.777*** | 0.788*** | 0.786*** |
| | (0.029) | (0.022) | (0.024) | (0.033) |
| Clusters | 287 | 543 | 543 | 543 |
| Observations | 5435 | 10236 | 10236 | 10236 |
| $R^2$ | 0.480 | 0.490 | 0.489 | 0.489 |

*Notes:* This table shows reports OLS estimates corresponding to Equation 1.12 for the first set of sample splits. Sample sizes vary because for social class the middle category (“Middle Class”) is disregarded. Standard errors (in parentheses) are clustered at the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table 1.C.7.** Heterogeneity in Treatment Effects by Demographic Group (II)

| | Dependent Variable: Extent of Redistribution $\theta_{i,\sigma}$ | | | | |
|---|---|---|---|---|---|
| | (1) Voting Freq. $D_i$=1 if high | (2) Econ. Ideology $D_i$=1 if conserv. | (3) Party Ident. $D_i$=1 if Rep. | (4) Age $D_i$=1 if old | (5) Sex $D_i$=1 if female |
| Effort ($E_\sigma$) | -0.762*** | -0.750*** | -0.747*** | -0.735*** | -0.680*** |
| | (0.023) | (0.026) | (0.026) | (0.029) | (0.031) |
| Inherited ($II_\sigma$) | 0.039** | 0.031 | 0.029 | 0.020 | 0.068*** |
| | (0.019) | (0.021) | (0.022) | (0.023) | (0.022) |
| Effort ($E_\sigma$) × Inherited ($II_\sigma$) | 0.016 | 0.006 | 0.011 | 0.030 | -0.029 |
| | (0.023) | (0.025) | (0.027) | (0.027) | (0.028) |
| $D_i$ | -0.010 | -0.011 | 0.013 | 0.032 | 0.101*** |
| | (0.037) | (0.042) | (0.046) | (0.034) | (0.034) |
| Effort ($E_\sigma$) × $D_i$ | 0.050 | 0.000 | -0.014 | -0.023 | -0.125*** |
| | (0.045) | (0.047) | (0.051) | (0.039) | (0.039) |
| Inherited ($II_\sigma$) × $D_i$ | -0.061** | -0.033 | -0.060 | 0.000 | -0.085*** |
| | (0.031) | (0.040) | (0.042) | (0.031) | (0.030) |
| Effort ($E_\sigma$) × Inherited ($II_\sigma$) × $D_i$ | 0.020 | 0.031 | 0.069 | -0.015 | 0.090** |
| | (0.038) | (0.048) | (0.052) | (0.037) | (0.036) |
| Initial Inequality ($\Delta_\sigma$) | 0.079*** | 0.070*** | 0.072*** | 0.078*** | 0.080*** |
| | (0.019) | (0.022) | (0.022) | (0.019) | (0.019) |
| Constant | 0.787*** | 0.792*** | 0.783*** | 0.768*** | 0.729*** |
| | (0.023) | (0.025) | (0.025) | (0.027) | (0.028) |
| Clusters | 543 | 417 | 398 | 543 | 542 |
| Observations | 10236 | 7853 | 7485 | 10236 | 10216 |
| $R^2$ | 0.489 | 0.502 | 0.488 | 0.489 | 0.492 |

*Notes:* This table shows reports OLS estimates corresponding to Equation 1.12 for the second set of sample splits. Sample sizes vary because for economic ideology and party identification the middle categories ("Moderate" and "Neither Republican nor Democrat") are disregarded. Standard errors (in parentheses) are clustered at the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table 1.C.8.** Association between Experimental Measures and Policy Preferences

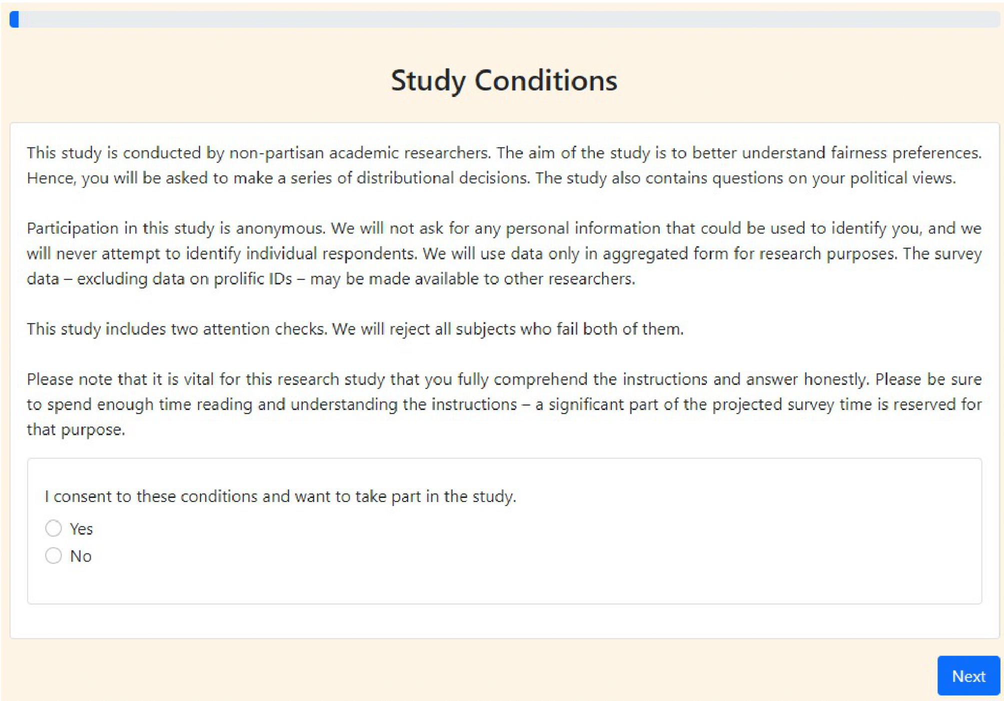
| | Preferred Max. Marg. Rate | | Support for | | | Rejection of |
|---|---|---|---|---|---|---|
| | (1) Income Tax | (2) Estate Tax | (3) Disability Ins. | (4) Unemployment Ins. | (5) Equal Opp. Prog. | (6) Interg. Transm. |
| Redistribution (Luck) | 0.136*** | 0.078* | 0.081* | 0.073 | 0.081* | 0.197*** |
| | (0.049) | (0.045) | (0.048) | (0.052) | (0.048) | (0.047) |
| Redistribution (Effort) | 0.022 | 0.013 | 0.076 | 0.120*** | 0.059 | 0.111** |
| | (0.057) | (0.059) | (0.047) | (0.042) | (0.042) | (0.046) |
| Observations | 437 | 437 | 437 | 437 | 437 | 437 |
| $R^2$ | 0.019 | 0.006 | 0.013 | 0.020 | 0.010 | 0.052 |

*Notes:* This table shows OLS estimates of (standardized) survey-based policy attitudes on (standardized) factor variables based on spectators' average extent of redistribution in the four treatment conditions. The coefficients are plotted in Figure 1.9. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

## Appendix 1.D Instructions for the Spectator Session

Below are the full instructions for the spectator session/redistribution stage.

*The following pages were shown to all subjects in the same order as presented here.*



### Study Conditions

This study is conducted by non-partisan academic researchers. The aim of the study is to better understand fairness preferences. Hence, you will be asked to make a series of distributional decisions. The study also contains questions on your political views.

Participation in this study is anonymous. We will not ask for any personal information that could be used to identify you, and we will never attempt to identify individual respondents. We will use data only in aggregated form for research purposes. The survey data – excluding data on prolific IDs – may be made available to other researchers.

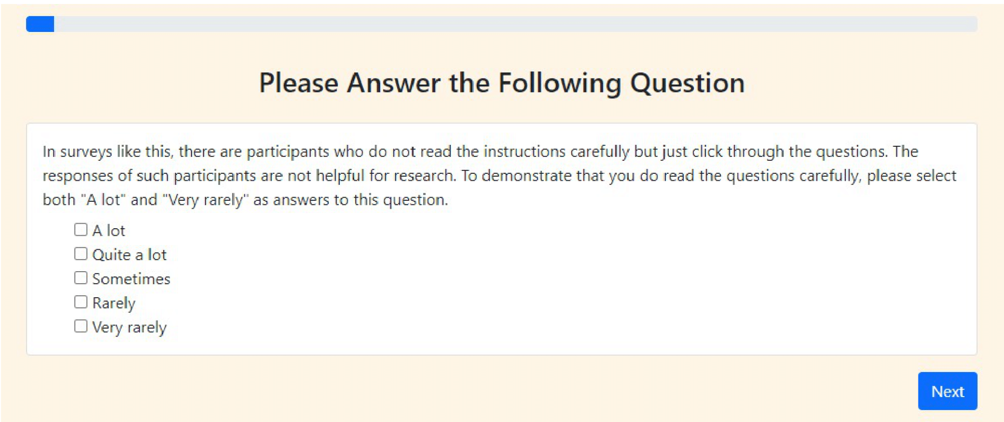
This study includes two attention checks. We will reject all subjects who fail both of them.

Please note that it is vital for this research study that you fully comprehend the instructions and answer honestly. Please be sure to spend enough time reading and understanding the instructions – a significant part of the projected survey time is reserved for that purpose.

I consent to these conditions and want to take part in the study.

- Yes
- No

Next

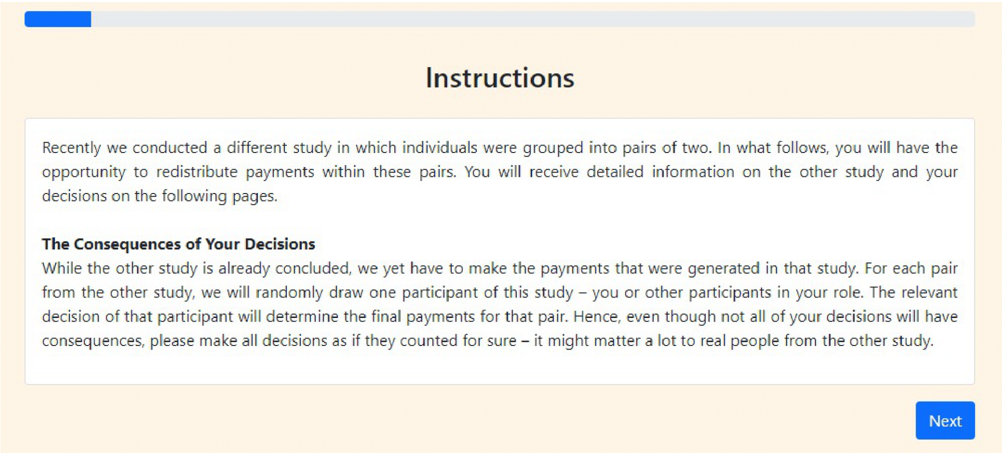


### Please Answer the Following Question

In surveys like this, there are participants who do not read the instructions carefully but just click through the questions. The responses of such participants are not helpful for research. To demonstrate that you do read the questions carefully, please select both "A lot" and "Very rarely" as answers to this question.

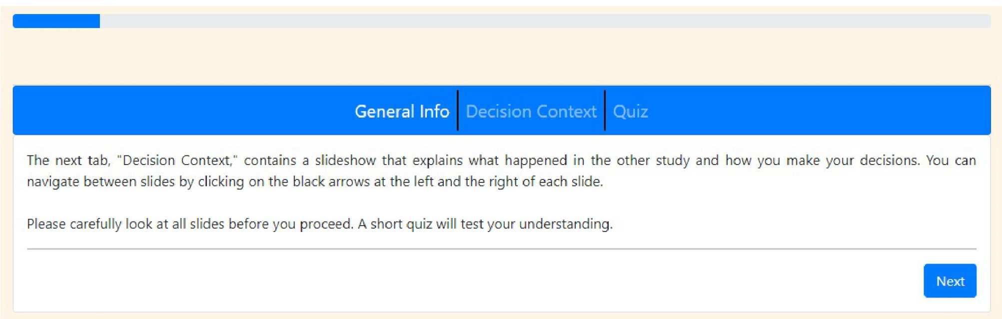
- ☐ A lot
- ☐ Quite a lot
- ☐ Sometimes
- ☐ Rarely
- ☐ Very rarely

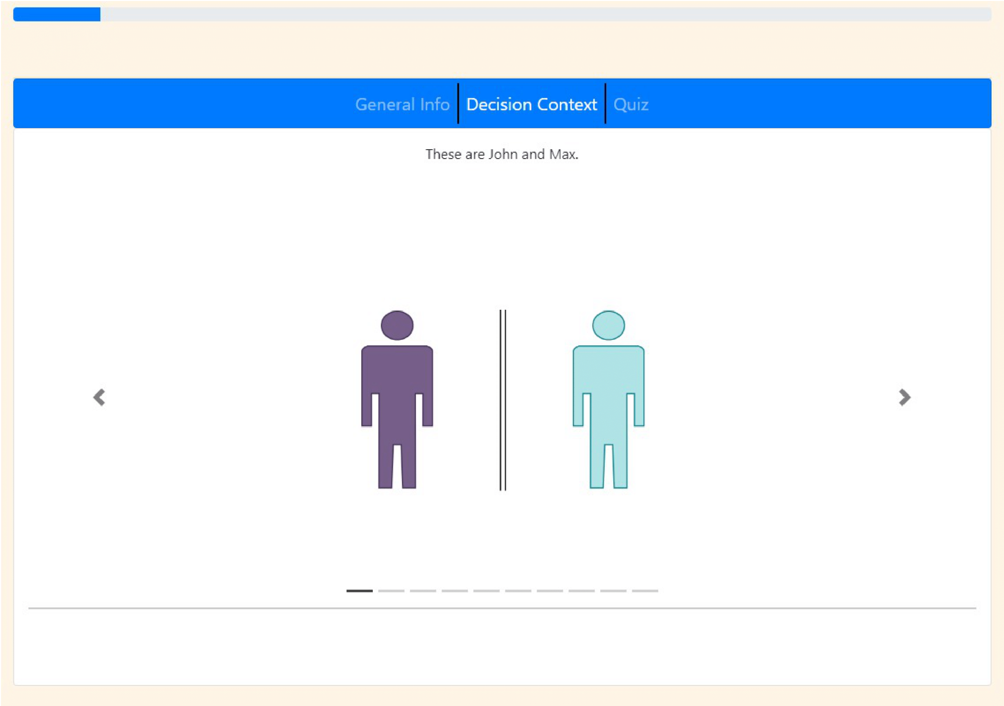
Next

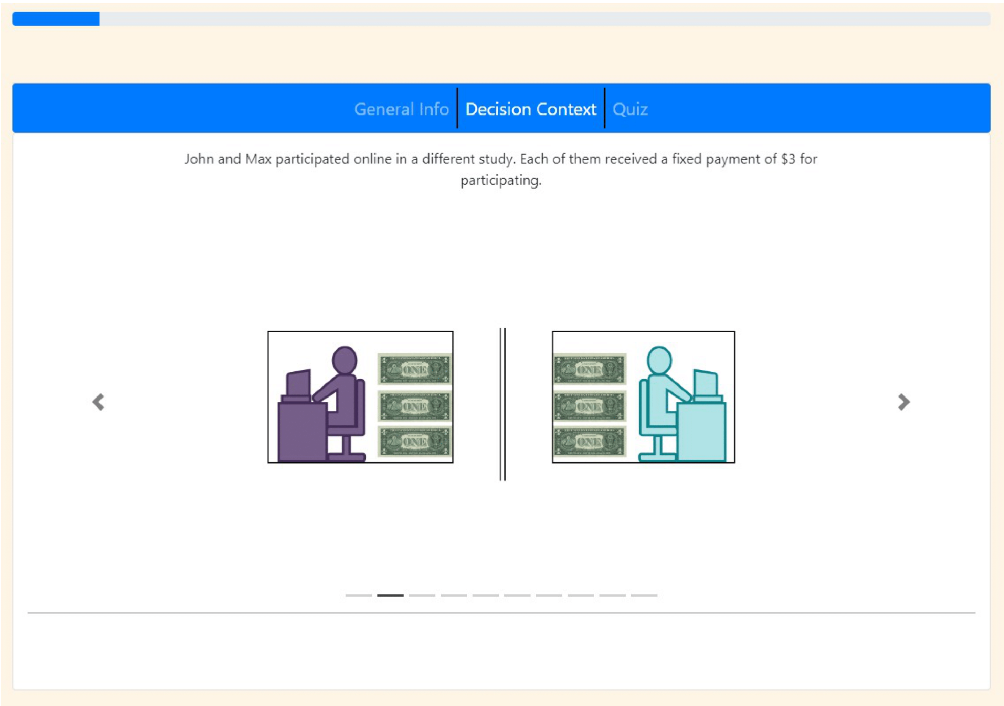


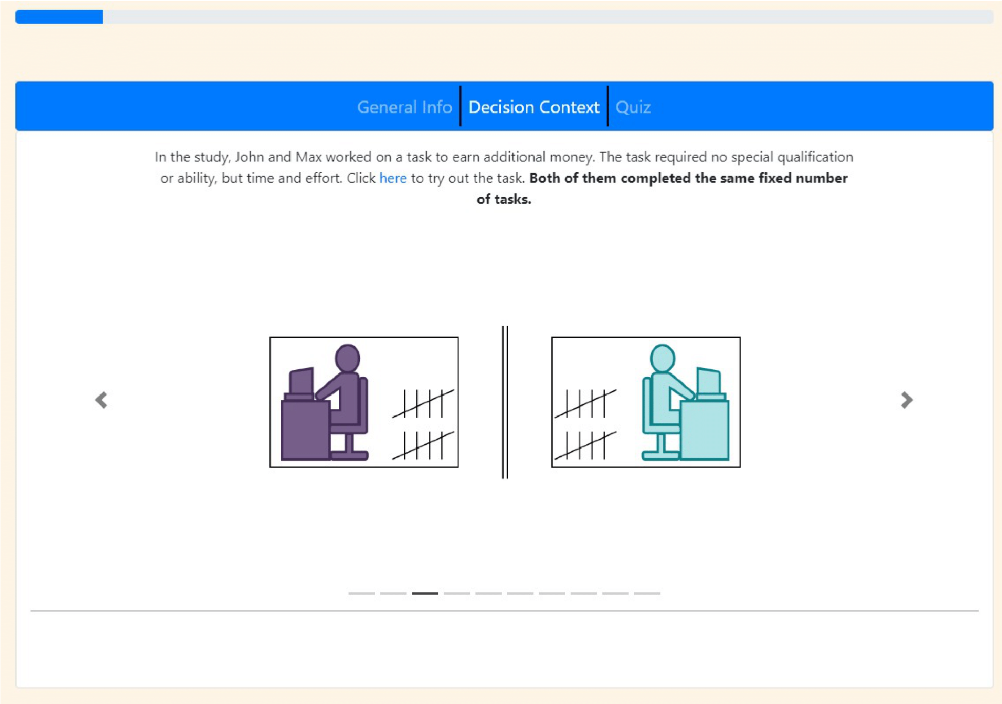

*The order of the following four blocks of pages was randomly assigned for each participant. However, the "General Info" pages were always ordered such that the first "General Info" page a subject would see referred to the first block of decisions, the second "General Info" page referred to the second block of decisions and so on. Within each block subjects made six decisions.*

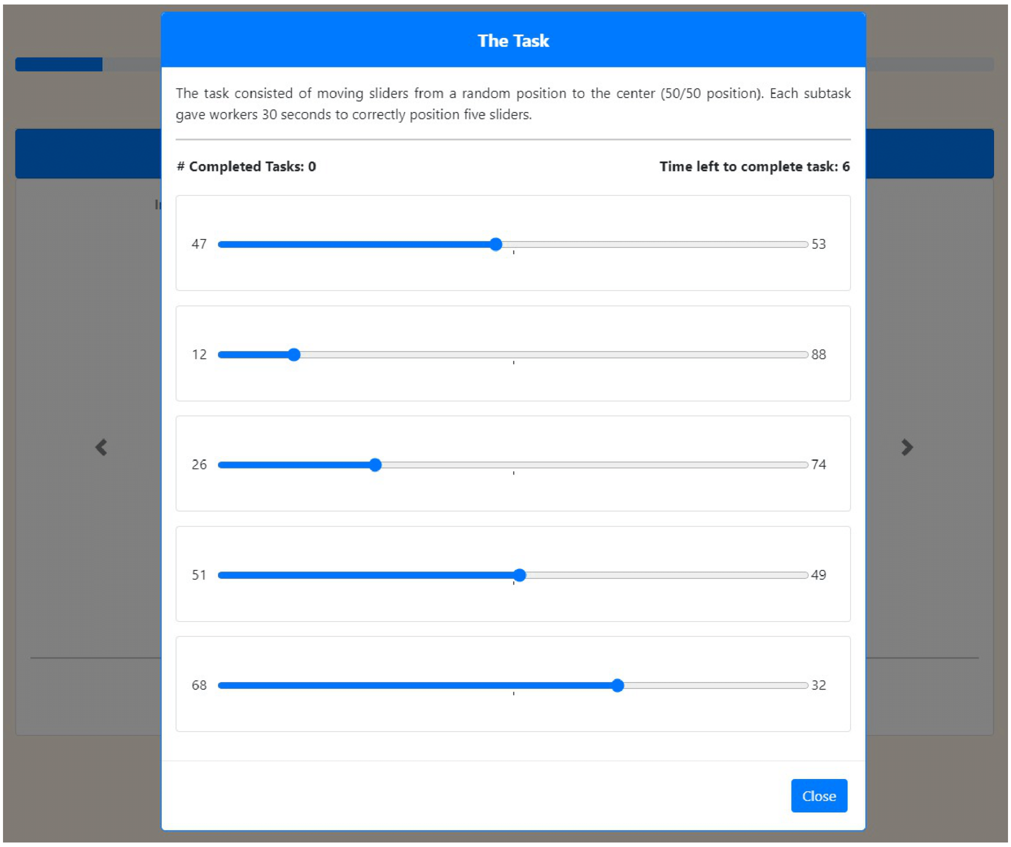
## Block 1

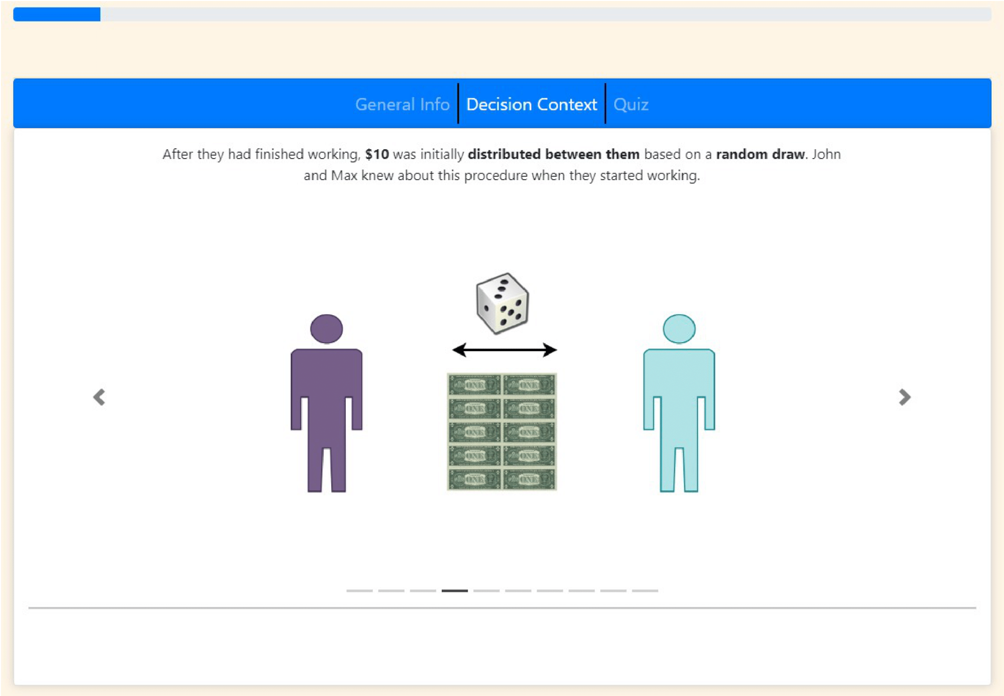


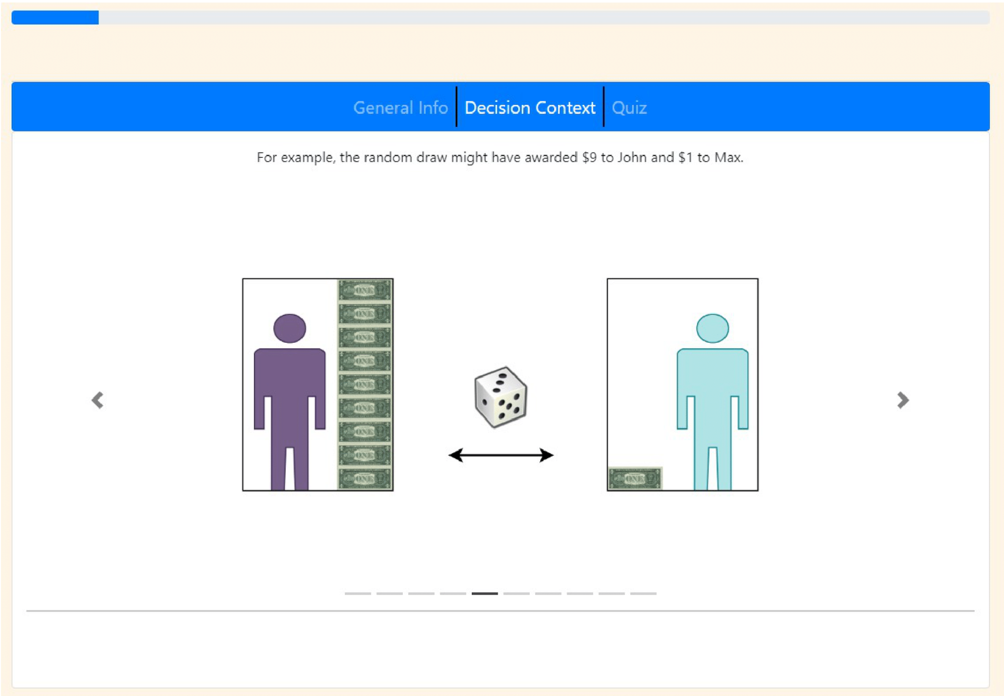

General Info
Decision Context
Quiz
These are John and Max.

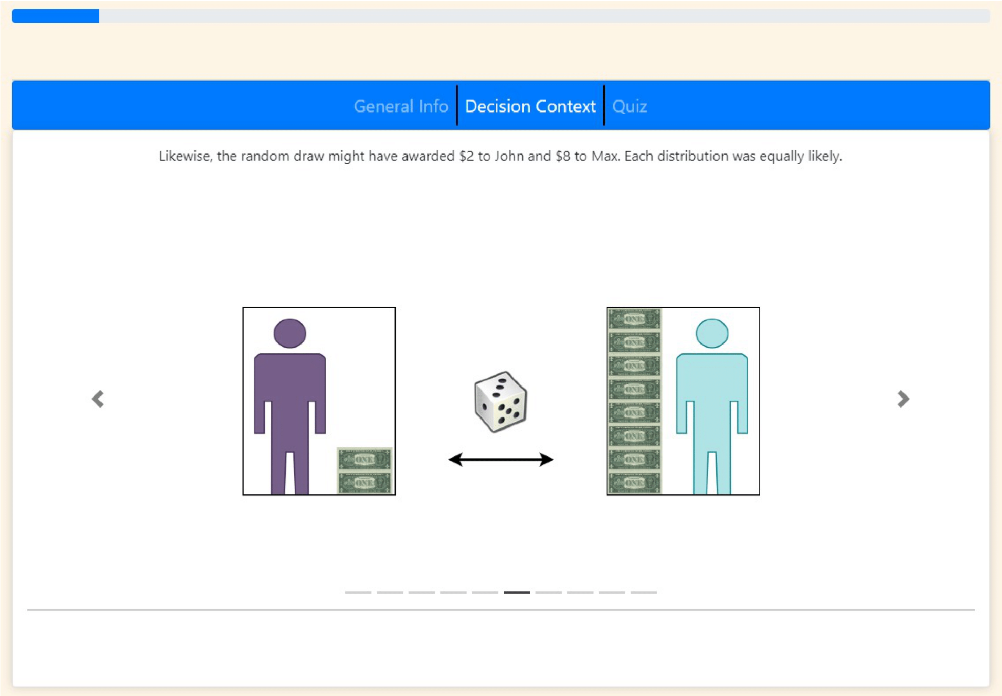

General Info
Decision Context
Quiz
John and Max participated online in a different study. Each of them received a fixed payment of $3 for participating.

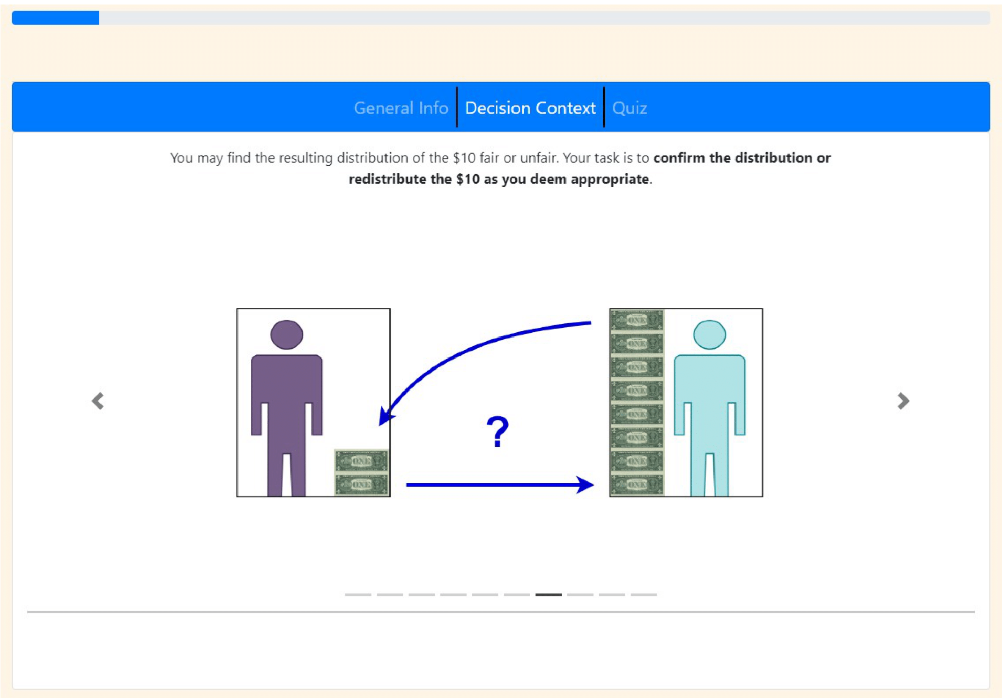

General Info
Decision Context
Quiz
In the study, John and Max worked on a task to earn additional money. The task required no special qualification or ability, but time and effort. Click here to try out the task. Both of them completed the same fixed number of tasks.

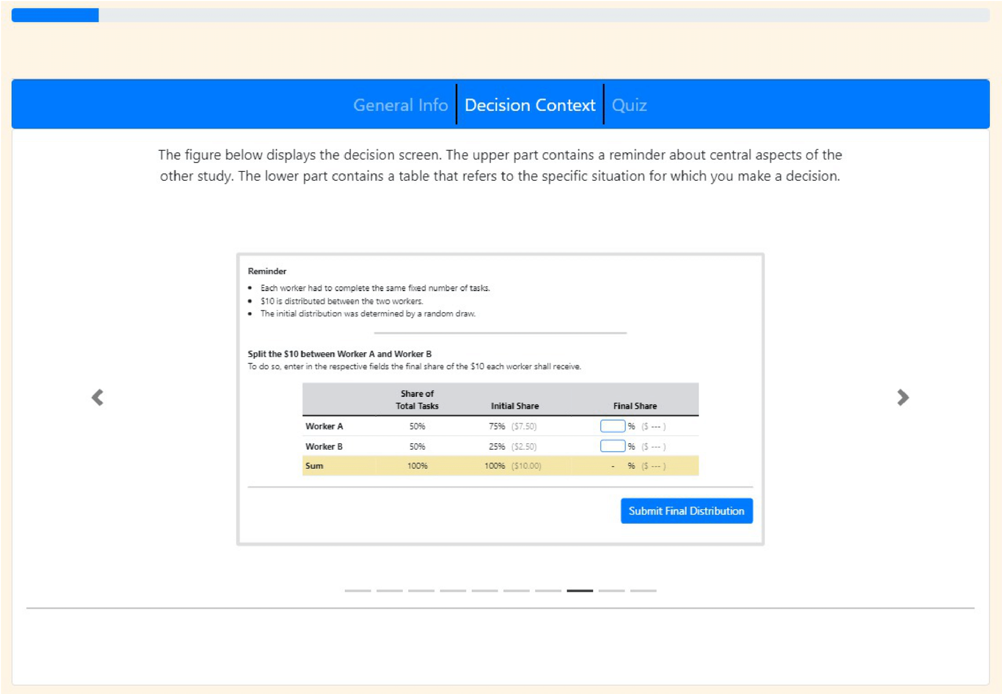

The Task
The task consisted of moving sliders from a random position to the center (50/50 position). Each subtask gave workers 30 seconds to correctly position five sliders.
# Completed Tasks: 0
Time left to complete task: 6
47
53
12
88
26
74
51
49
68
32
Close

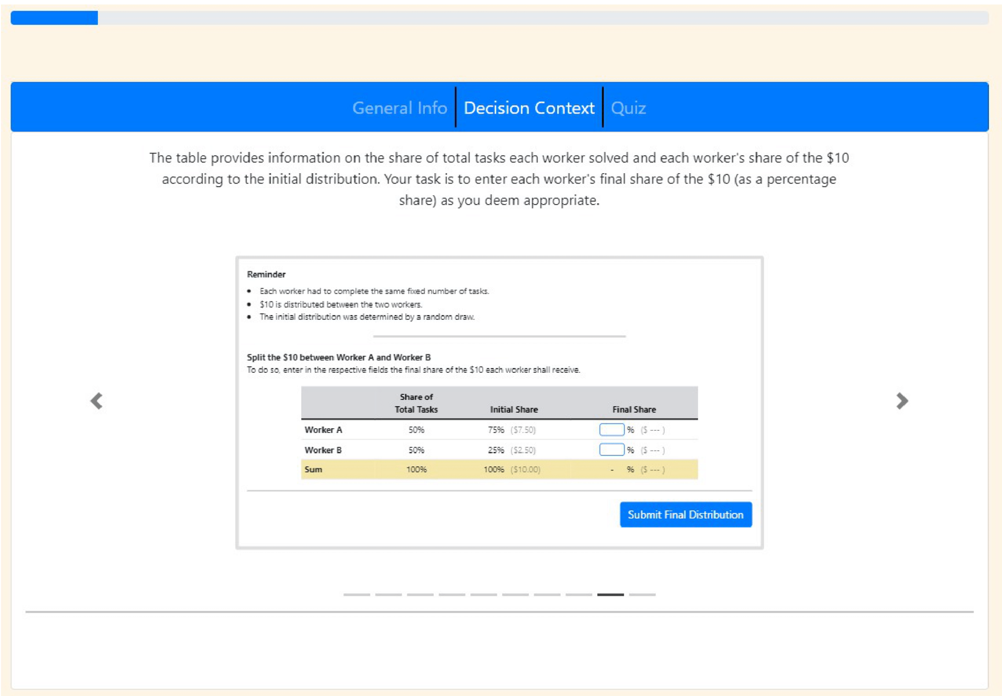

General Info
Decision Context
Quiz
After they had finished working, $10 was initially distributed between them based on a random draw. John and Max knew about this procedure when they started working.

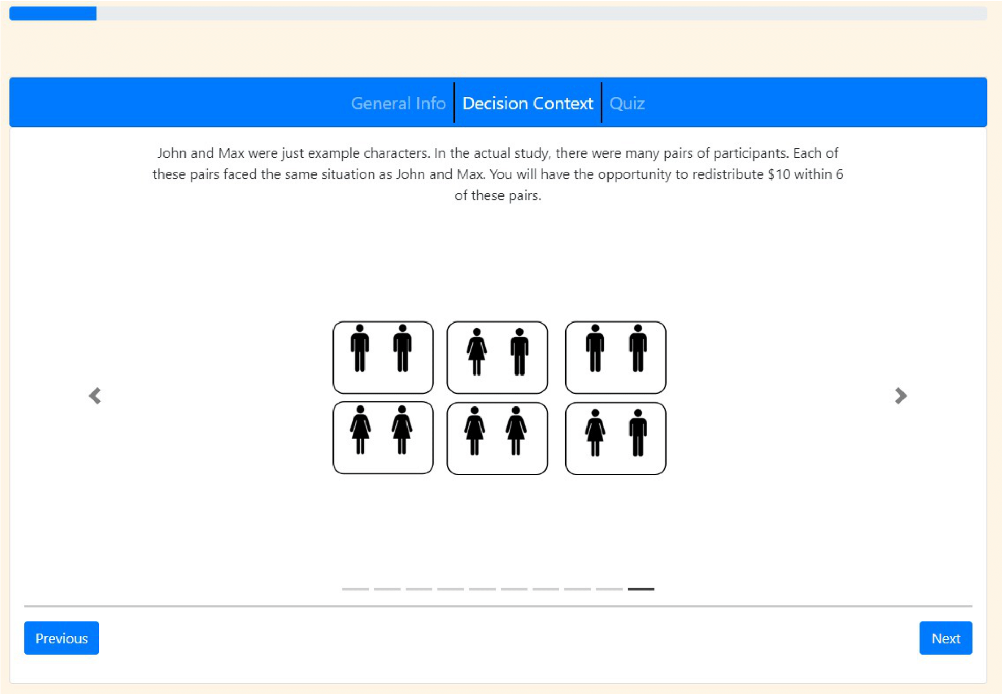

General Info
Decision Context
Quiz
For example, the random draw might have awarded $9 to John and $1 to Max.

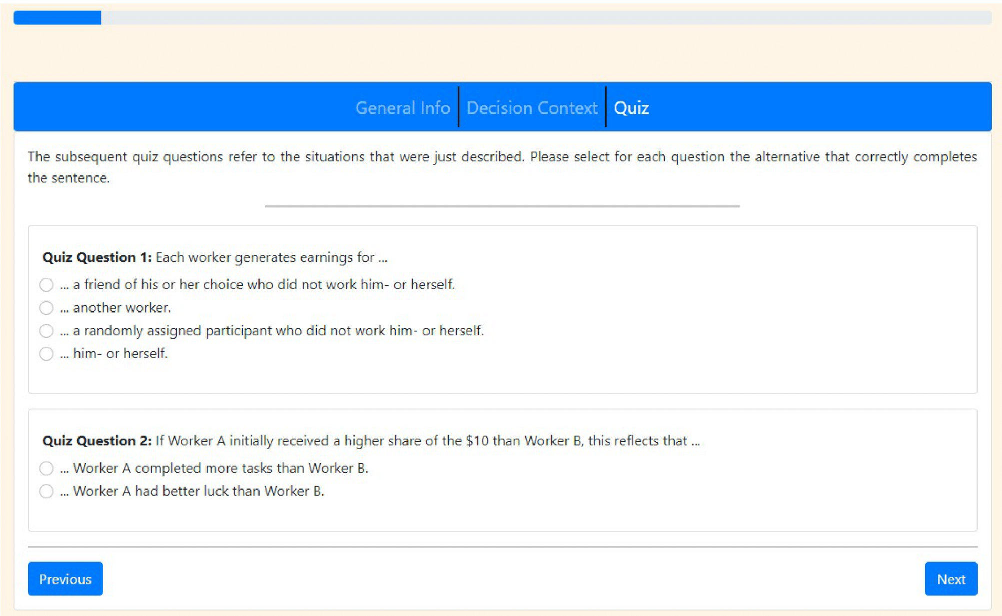

General Info
Decision Context
Quiz
Likewise, the random draw might have awarded $2 to John and $8 to Max. Each distribution was equally likely.


General Info
Decision Context
Quiz
You may find the resulting distribution of the $10 fair or unfair. Your task is to **confirm the distribution or redistribute the $10 as you deem appropriate**.
?


General Info
Decision Context
Quiz
The figure below displays the decision screen. The upper part contains a reminder about central aspects of the other study. The lower part contains a table that refers to the specific situation for which you make a decision.
Reminder
Each worker had to complete the same fixed number of tasks.
$10 is distributed between the two workers.
The initial distribution was determined by a random draw.
Split the $10 between Worker A and Worker B
To do so, enter in the respective fields the final share of the $10 each worker shall receive.
Share of Total Tasks
Initial Share
Final Share
Worker A
50%
75% ($7.50)
% ($ --- )
Worker B
50%
25% ($2.50)
% ($ --- )
Sum
100%
100% ($10.00)
- % ($ --- )
Submit Final Distribution


General Info
Decision Context
Quiz
The table provides information on the share of total tasks each worker solved and each worker's share of the $10 according to the initial distribution. Your task is to enter each worker's final share of the $10 (as a percentage share) as you deem appropriate.
Reminder
Each worker had to complete the same fixed number of tasks.
$10 is distributed between the two workers.
The initial distribution was determined by a random draw.
Split the $10 between Worker A and Worker B
To do so, enter in the respective fields the final share of the $10 each worker shall receive.
Share of Total Tasks
Initial Share
Final Share
Worker A
50%
75% ($7.50)
% ($ --- )
Worker B
50%
25% ($2.50)
% ($ --- )
Sum
100%
100% ($10.00)
- % ($ --- )
Submit Final Distribution


General Info
Decision Context
Quiz
John and Max were just example characters. In the actual study, there were many pairs of participants. Each of these pairs faced the same situation as John and Max. You will have the opportunity to redistribute $10 within 6 of these pairs.
Previous
Next


General Info
Decision Context
Quiz
The subsequent quiz questions refer to the situations that were just described. Please select for each question the alternative that correctly completes the sentence.
Quiz Question 1: Each worker generates earnings for ...
... a friend of his or her choice who did not work him- or herself.
... another worker.
... a randomly assigned participant who did not work him- or herself.
... him- or herself.
Quiz Question 2: If Worker A initially received a higher share of the $10 than Worker B, this reflects that ...
... Worker A completed more tasks than Worker B.
... Worker A had better luck than Worker B.
Previous
Next

## Proceed to Decisions

On the following six pages, you will make your redistribution decisions for six different pairs of workers.

Before you proceed, please take a short moment to deliberately think about which distribution you find appropriate in this kind of situation.

Proceed to Decisions

## Situation 1

**Reminder**

- Each worker had to complete the same fixed number of tasks.
- $10 is distributed between the two workers.
- The initial distribution was determined by a random draw.

**Split the $10 between Worker A and Worker B**
To do so, enter in the respective fields the final share of the $10 each worker shall receive.

| | Share of Total Tasks | Initial Share | Final Share |
|---|---|---|---|
| **Worker A** | 50% | 0% ($0.00) | % ($ --- ) |
| **Worker B** | 50% | 100% ($10.00) | % ($ --- ) |
| **Sum** | 100% | 100% ($10.00) | - % ($ --- ) |

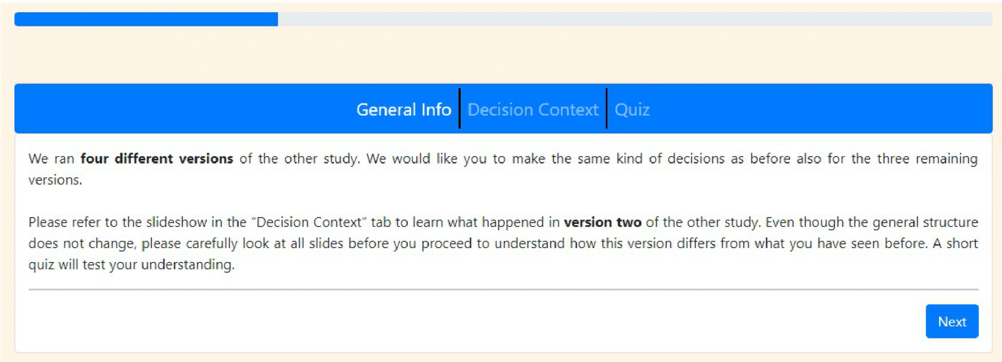
Submit Final Distribution

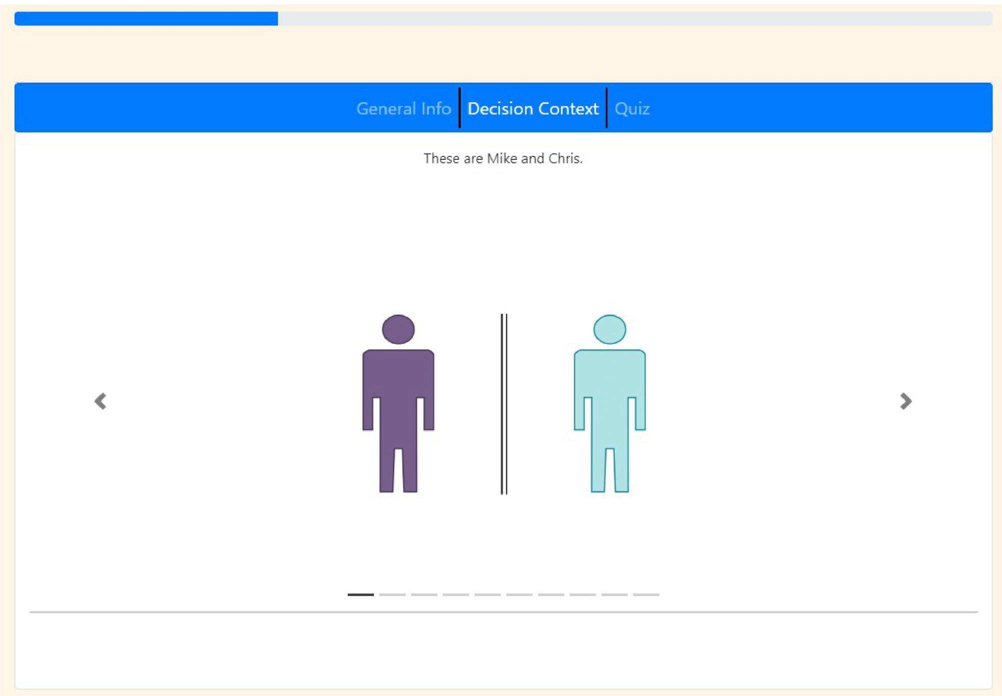
## Your Considerations

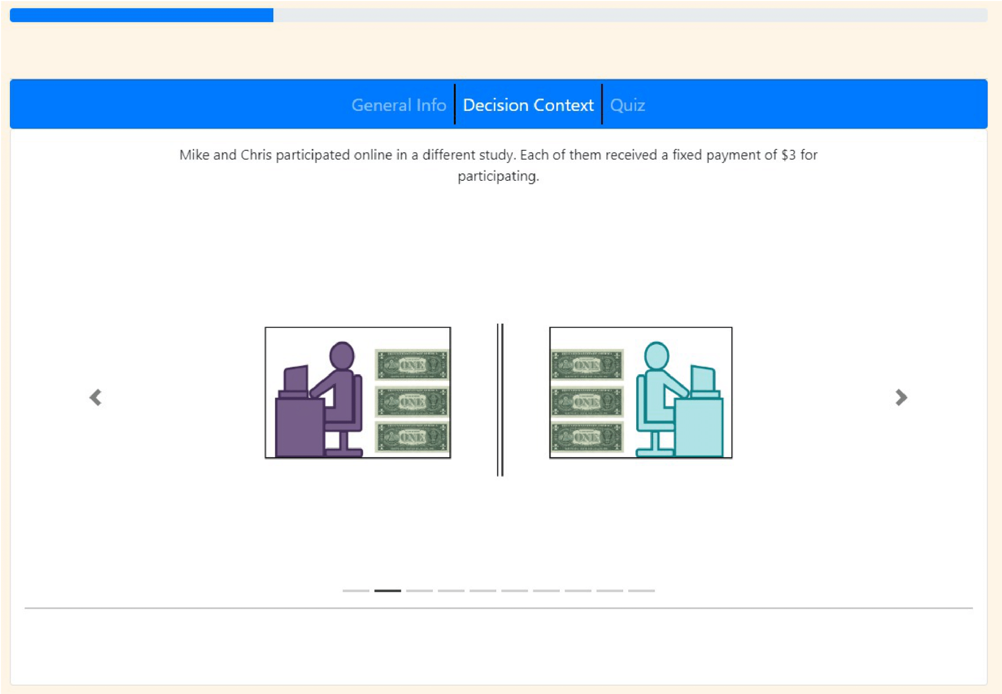
To help us understand why you consider the final payments you implemented in the past six decisions appropriate, please briefly describe the reasoning behind your decisions.

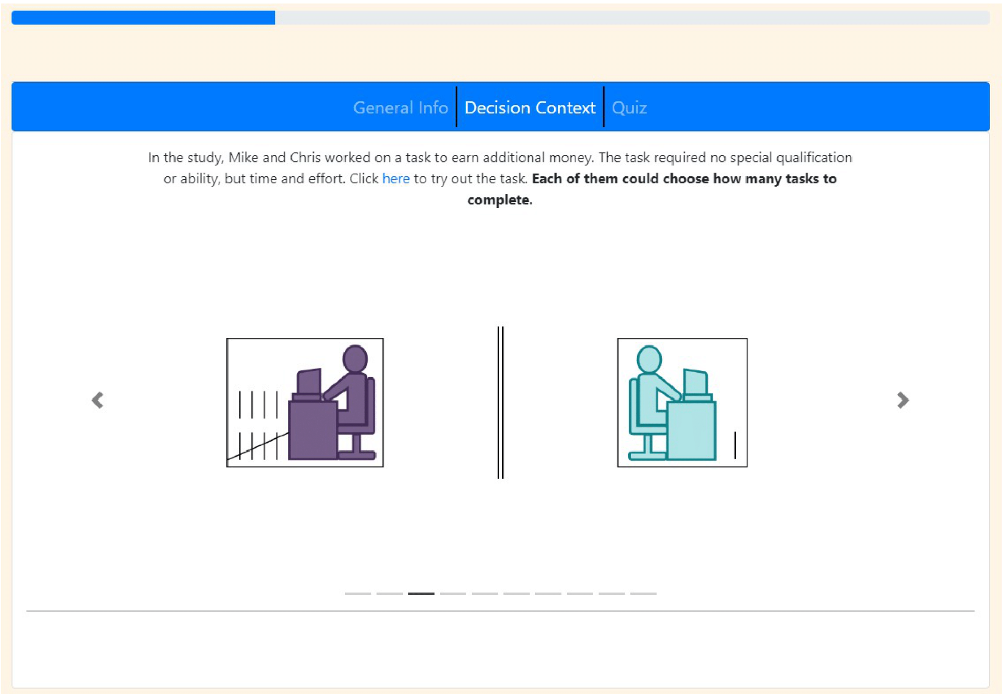
Next

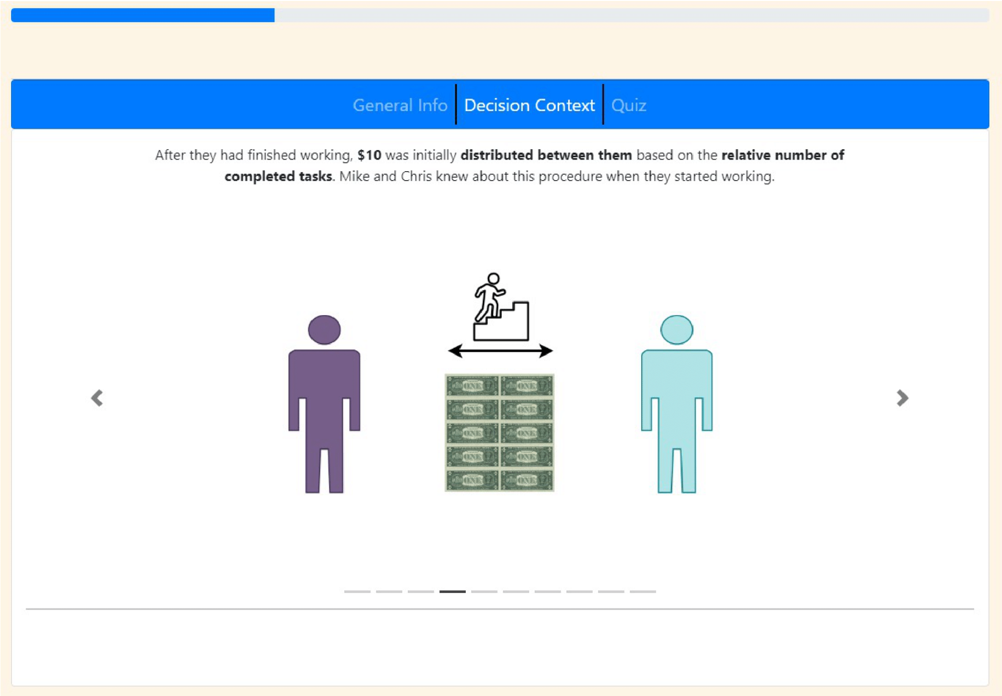
## Block 2

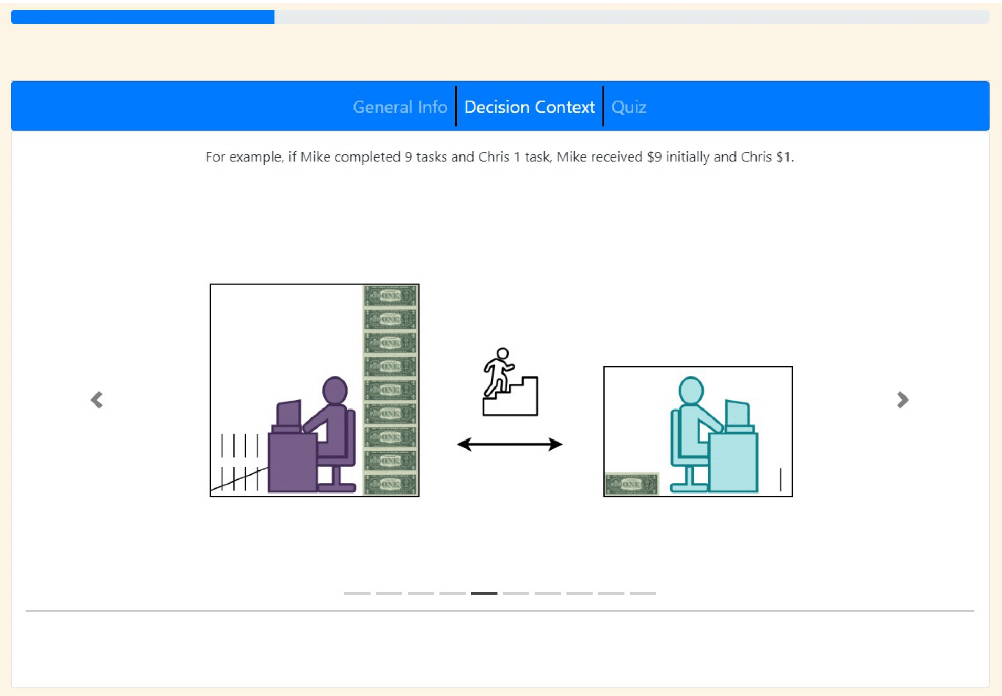

General Info
Decision Context
Quiz
We ran **four different versions** of the other study. We would like you to make the same kind of decisions as before also for the three remaining versions.
Please refer to the slideshow in the "Decision Context" tab to learn what happened in **version two** of the other study. Even though the general structure does not change, please carefully look at all slides before you proceed to understand how this version differs from what you have seen before. A short quiz will test your understanding.
Next

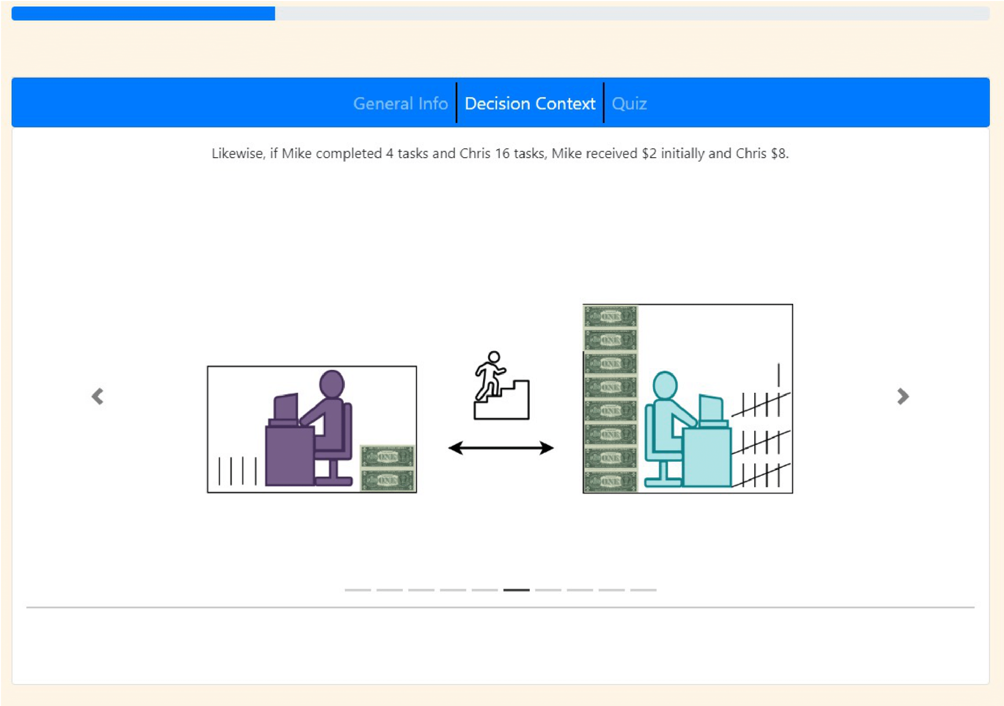

General Info
Decision Context
Quiz
These are Mike and Chris.

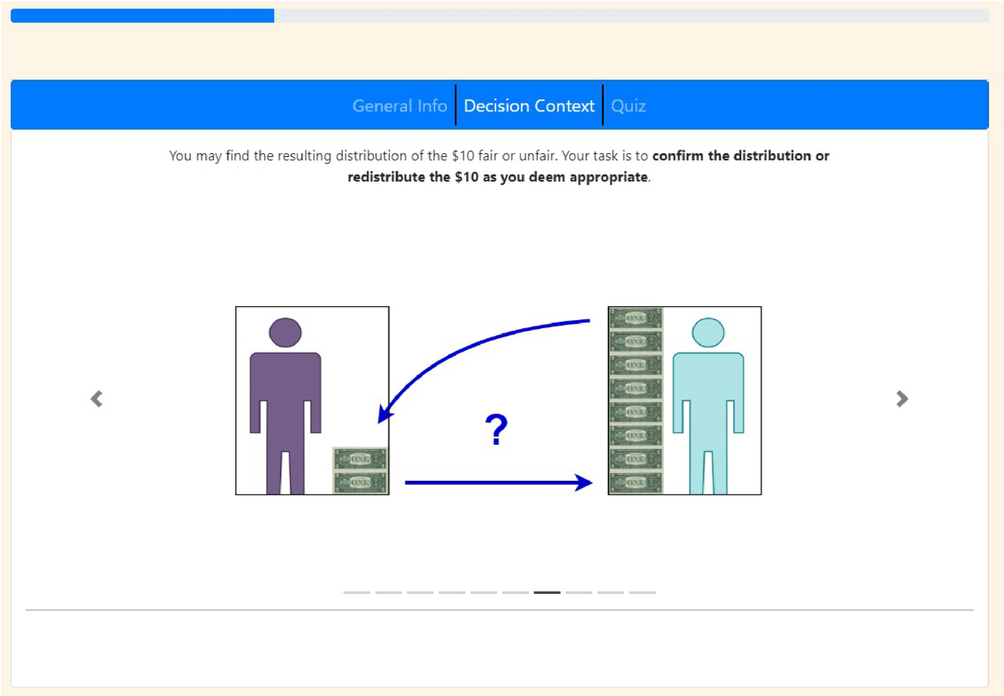

General Info
Decision Context
Quiz
Mike and Chris participated online in a different study. Each of them received a fixed payment of $3 for participating.

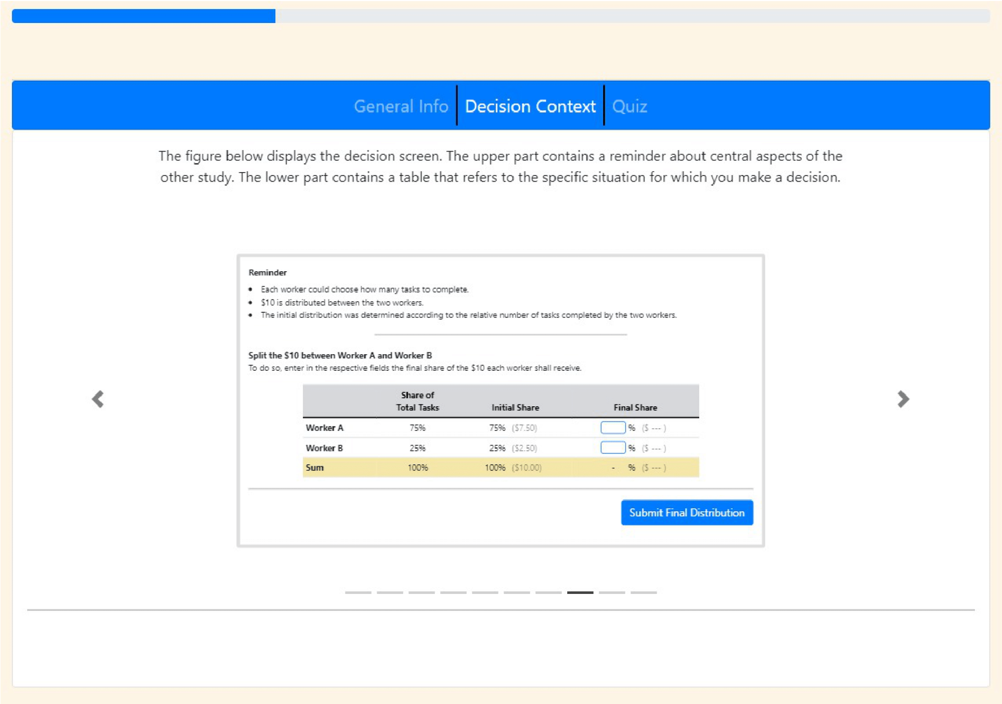

General Info
Decision Context
Quiz
In the study, Mike and Chris worked on a task to earn additional money. The task required no special qualification or ability, but time and effort. Click here to try out the task. **Each of them could choose how many tasks to complete.**

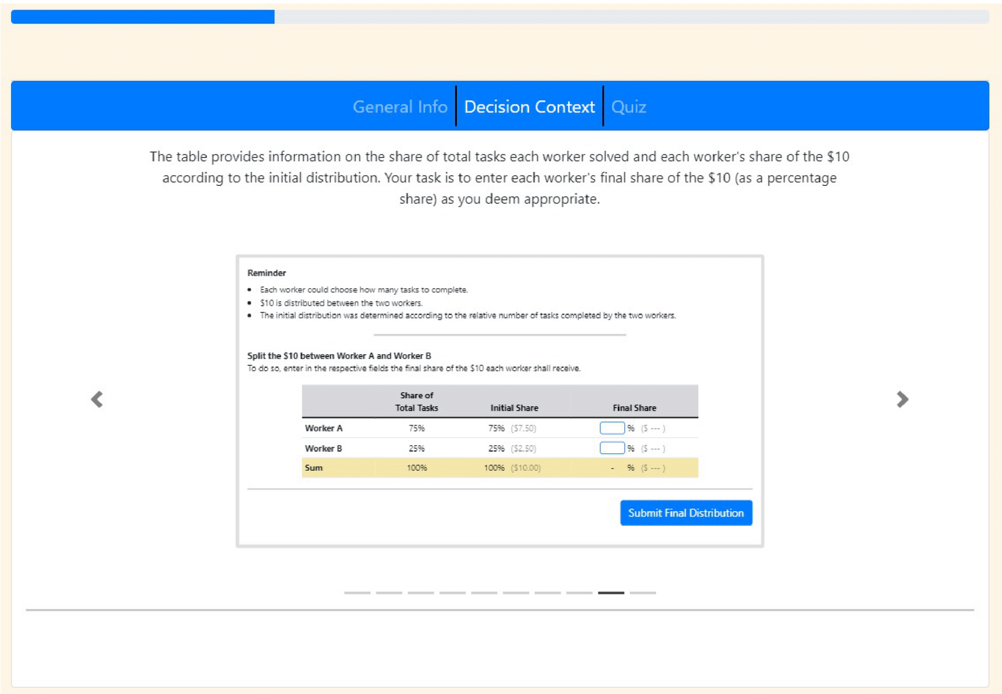

General Info
Decision Context
Quiz
After they had finished working, $10 was initially distributed between them based on the relative number of completed tasks. Mike and Chris knew about this procedure when they started working.

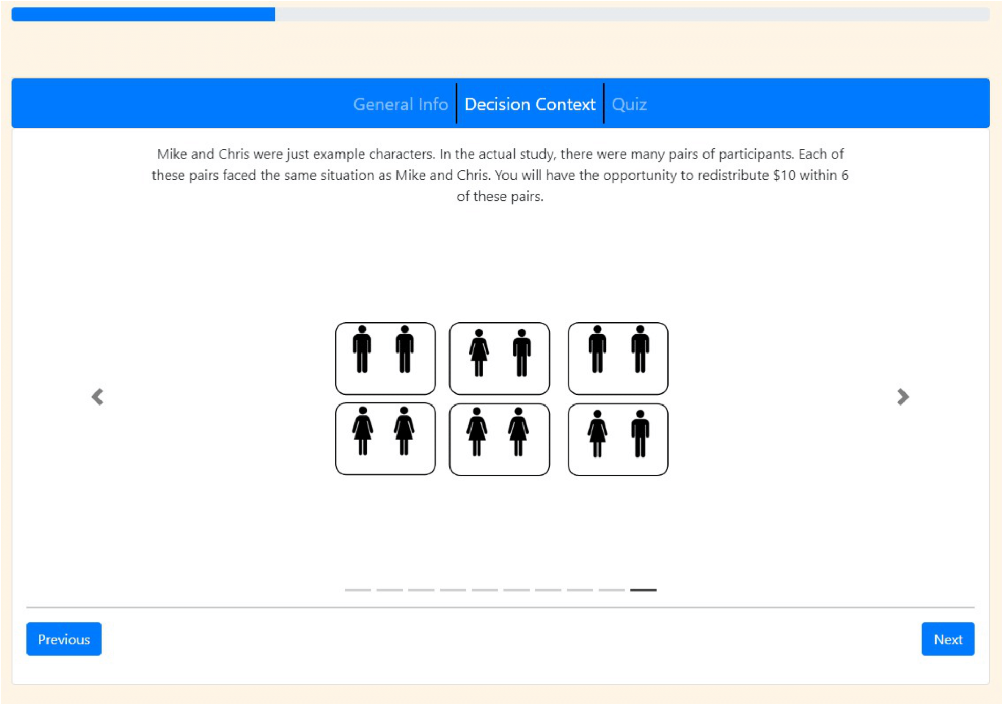

General Info
Decision Context
Quiz
For example, if Mike completed 9 tasks and Chris 1 task, Mike received $9 initially and Chris $1.

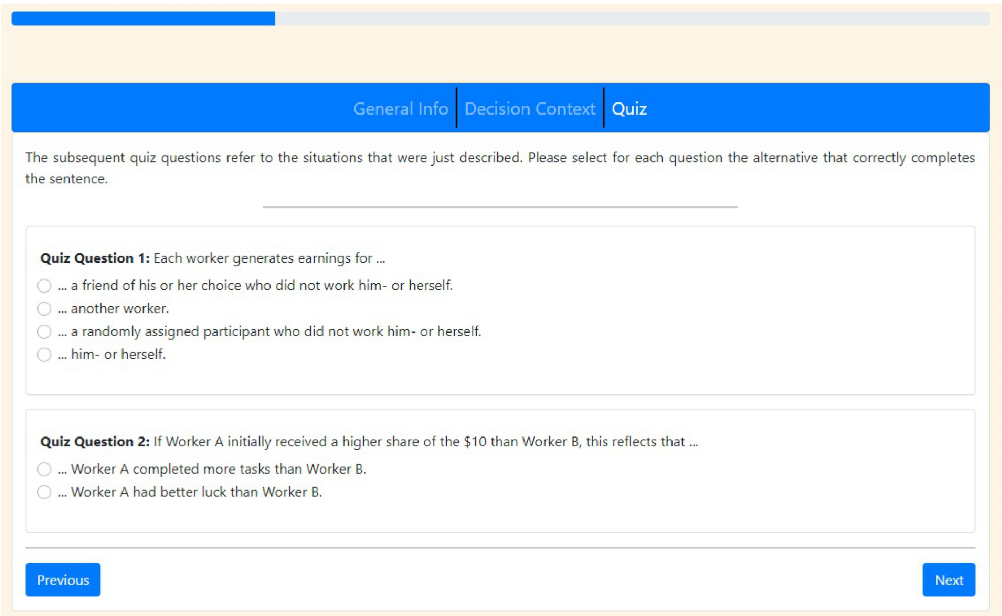

General Info
Decision Context
Quiz
Likewise, if Mike completed 4 tasks and Chris 16 tasks, Mike received $2 initially and Chris $8.


General Info
Decision Context
Quiz
You may find the resulting distribution of the $10 fair or unfair. Your task is to **confirm the distribution or redistribute the $10 as you deem appropriate**.
?


General Info
Decision Context
Quiz
The figure below displays the decision screen. The upper part contains a reminder about central aspects of the other study. The lower part contains a table that refers to the specific situation for which you make a decision.
Reminder
Each worker could choose how many tasks to complete.
$10 is distributed between the two workers.
The initial distribution was determined according to the relative number of tasks completed by the two workers.
Split the $10 between Worker A and Worker B
To do so, enter in the respective fields the final share of the $10 each worker shall receive.
Share of Total Tasks
Initial Share
Final Share
Worker A
75%
75% ($7.50)
Worker B
25%
25% ($2.50)
Sum
100%
100% ($10.00)
Submit Final Distribution


General Info
Decision Context
Quiz
The table provides information on the share of total tasks each worker solved and each worker's share of the $10 according to the initial distribution. Your task is to enter each worker's final share of the $10 (as a percentage share) as you deem appropriate.
Reminder
Each worker could choose how many tasks to complete.
$10 is distributed between the two workers.
The initial distribution was determined according to the relative number of tasks completed by the two workers.
Split the $10 between Worker A and Worker B
To do so, enter in the respective fields the final share of the $10 each worker shall receive.
Share of Total Tasks
Initial Share
Final Share
Worker A
75%
75% ($7.50)
Worker B
25%
25% ($2.50)
Sum
100%
100% ($10.00)
Submit Final Distribution


General Info
Decision Context
Quiz
Mike and Chris were just example characters. In the actual study, there were many pairs of participants. Each of these pairs faced the same situation as Mike and Chris. You will have the opportunity to redistribute $10 within 6 of these pairs.
Previous
Next


General Info
Decision Context
Quiz
The subsequent quiz questions refer to the situations that were just described. Please select for each question the alternative that correctly completes the sentence.
Quiz Question 1: Each worker generates earnings for ...
... a friend of his or her choice who did not work him- or herself.
... another worker.
... a randomly assigned participant who did not work him- or herself.
... him- or herself.
Quiz Question 2: If Worker A initially received a higher share of the $10 than Worker B, this reflects that ...
... Worker A completed more tasks than Worker B.
... Worker A had better luck than Worker B.
Previous
Next

## Proceed to Decisions

On the following six pages, you will make your redistribution decisions for six different pairs of workers.

Before you proceed, please take a short moment to deliberately think about which distribution you find appropriate in this kind of situation.

Proceed to Decisions

## Situation 1

**Reminder**

- Each worker could choose how many tasks to complete.
- $10 is distributed between the two workers.
- The initial distribution was determined according to the relative number of tasks completed by the two workers.

**Split the $10 between Worker A and Worker B**
To do so, enter in the respective fields the final share of the $10 each worker shall receive.

| | Share of Total Tasks | Initial Share | Final Share |
|---|---|---|---|
| **Worker A** | 68% | 68% ($6.80) | % ($ --- ) |
| **Worker B** | 32% | 32% ($3.20) | % ($ --- ) |
| **Sum** | 100% | 100% ($10.00) | - % ($ --- ) |

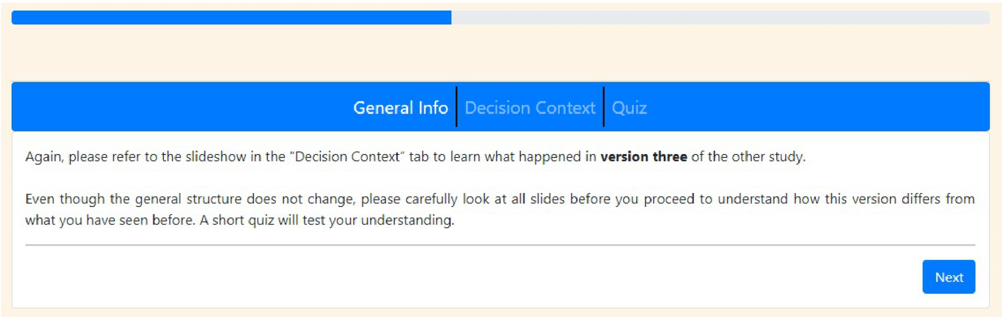
Submit Final Distribution

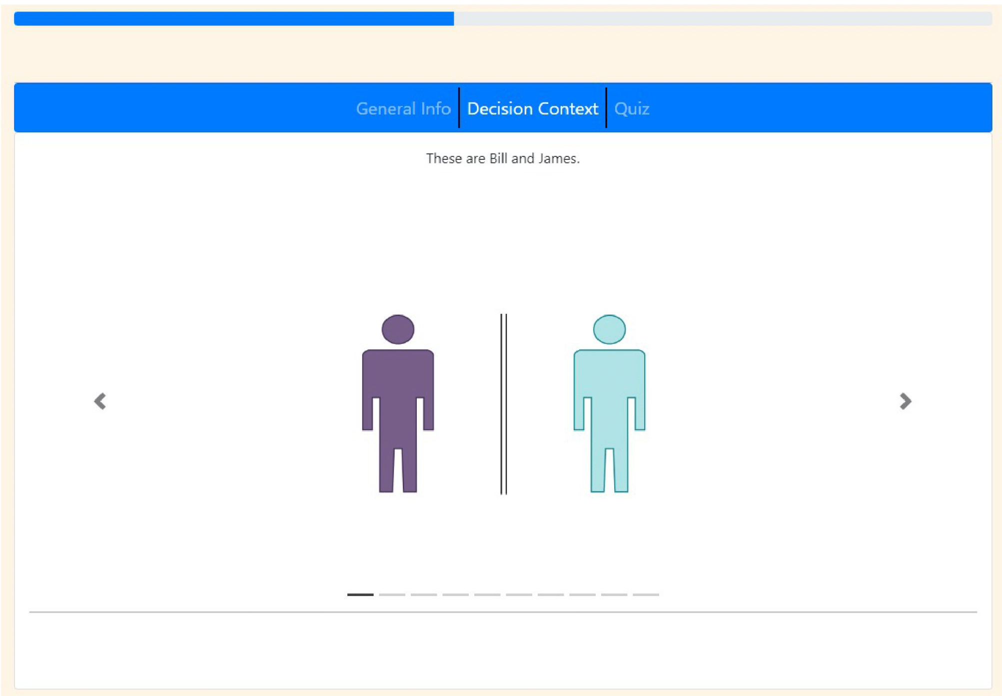
## Your Considerations

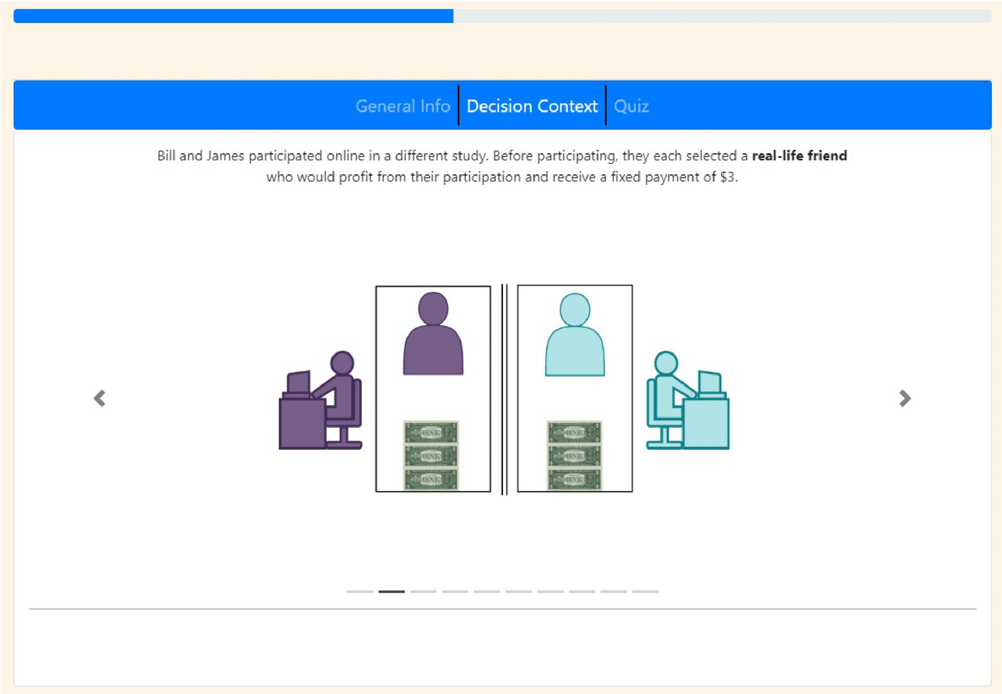
To help us understand why you consider the final payments you implemented in the past six decisions appropriate, please briefly describe the reasoning behind your decisions.

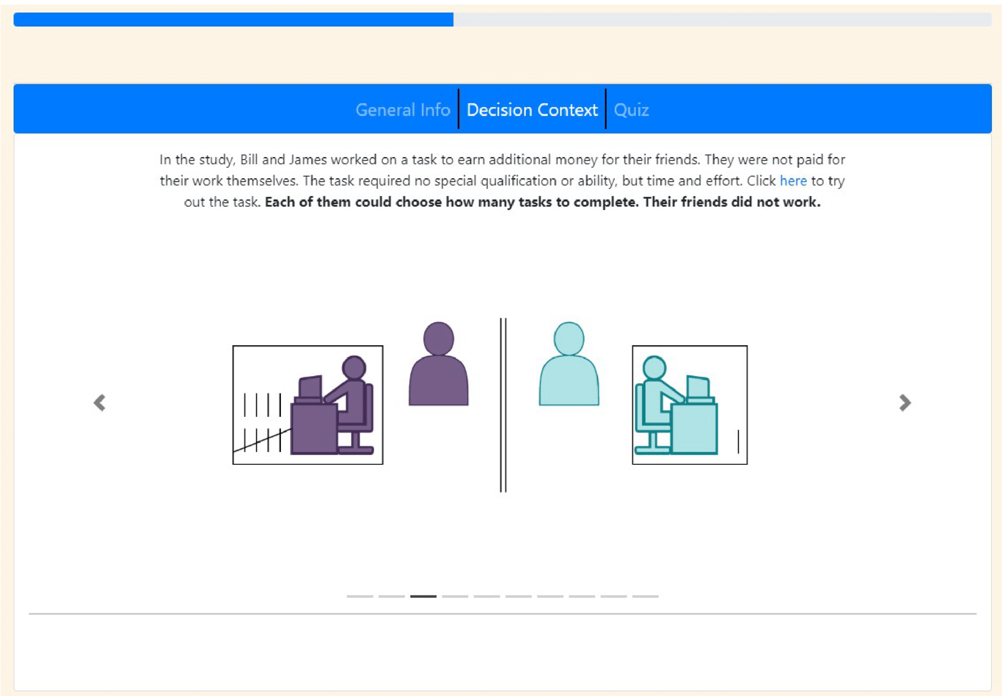
Next

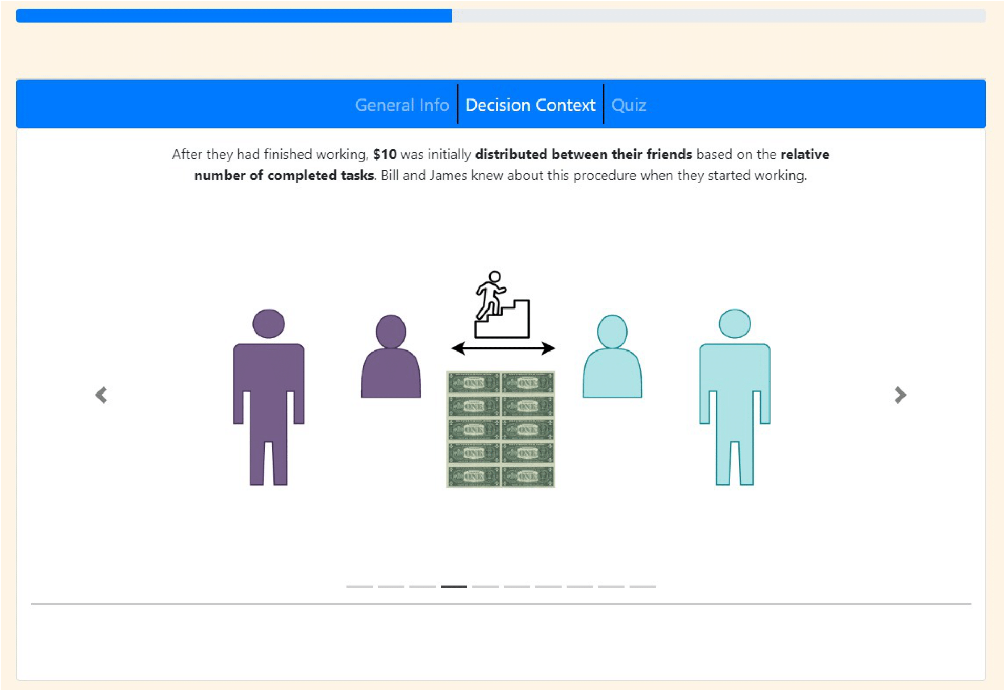
## Block 3

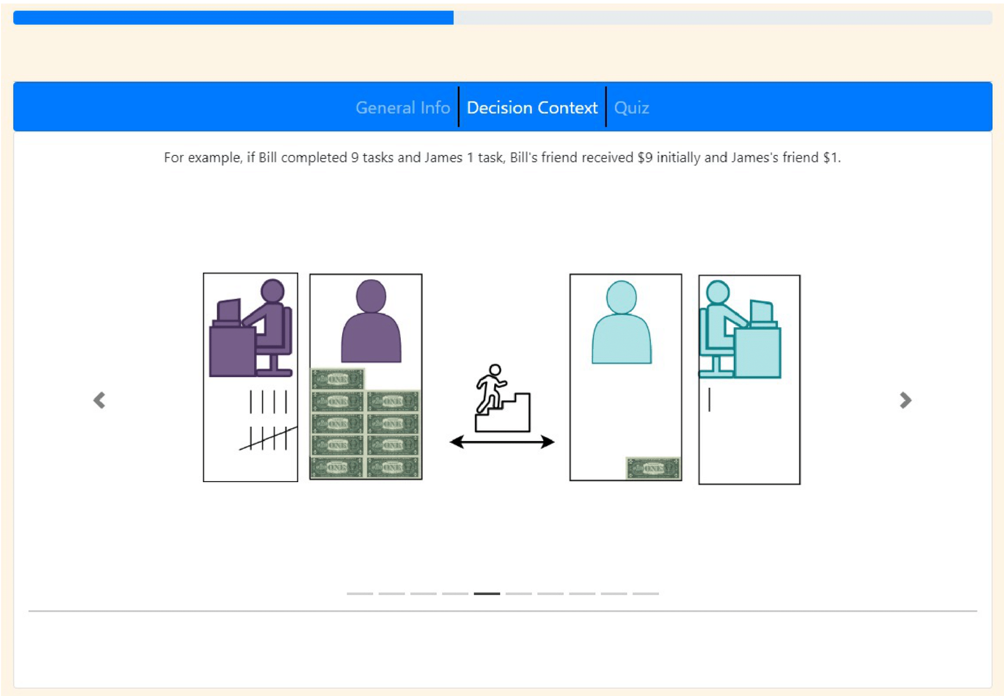

General Info
Decision Context
Quiz
Again, please refer to the slideshow in the "Decision Context" tab to learn what happened in **version three** of the other study.
Even though the general structure does not change, please carefully look at all slides before you proceed to understand how this version differs from what you have seen before. A short quiz will test your understanding.
Next

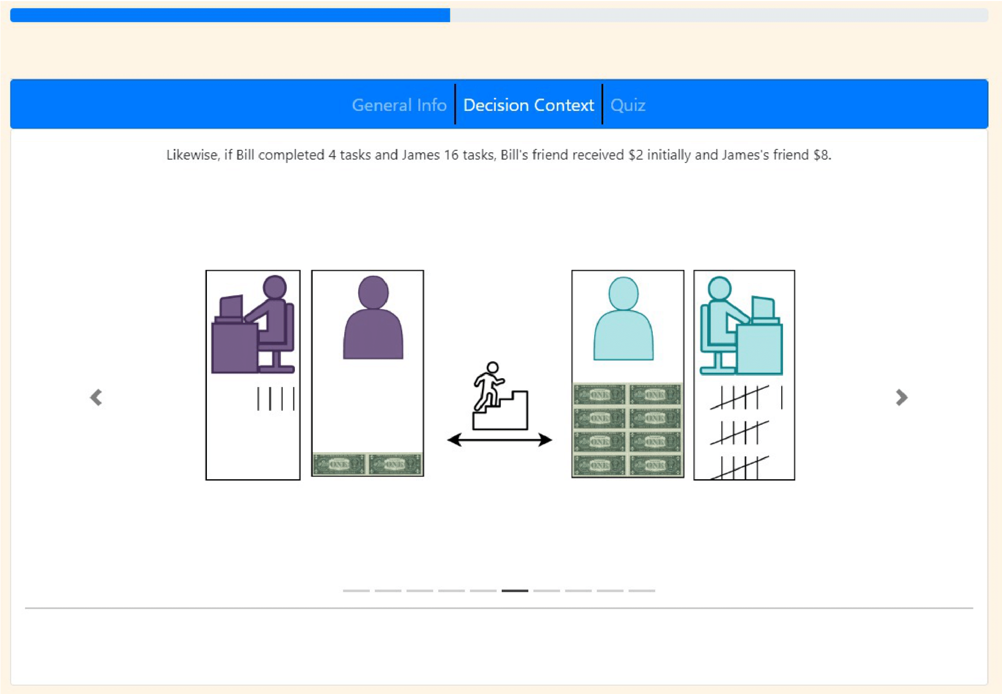

General Info
Decision Context
Quiz
These are Bill and James.

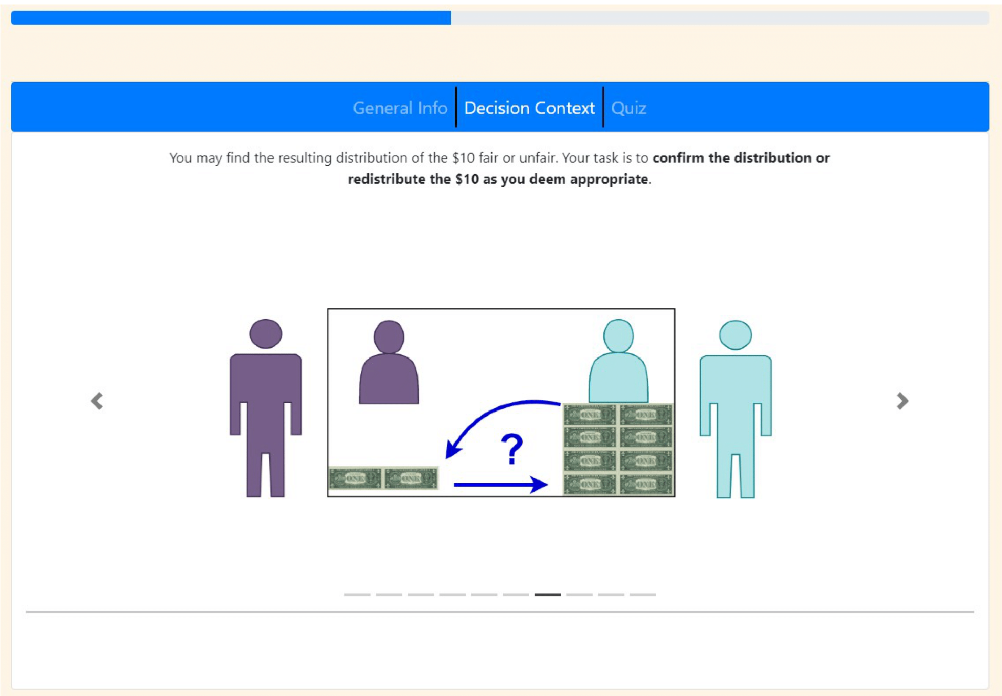

General Info
Decision Context
Quiz
Bill and James participated online in a different study. Before participating, they each selected a **real-life friend** who would profit from their participation and receive a fixed payment of $3.

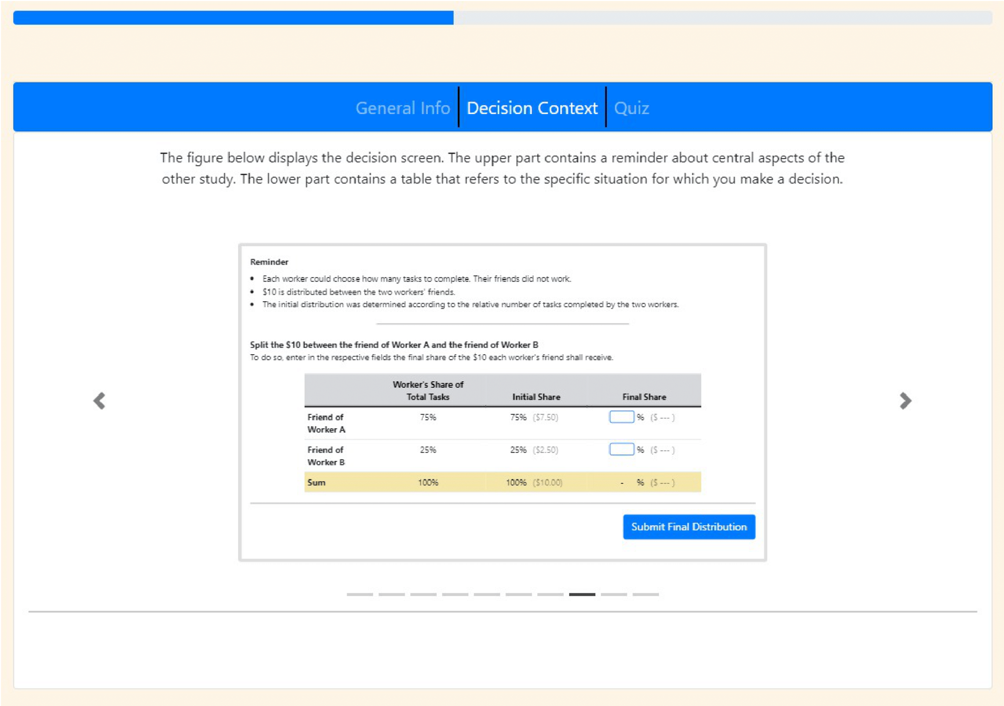

General Info
Decision Context
Quiz
In the study, Bill and James worked on a task to earn additional money for their friends. They were not paid for their work themselves. The task required no special qualification or ability, but time and effort. Click here to try out the task. **Each of them could choose how many tasks to complete. Their friends did not work.**

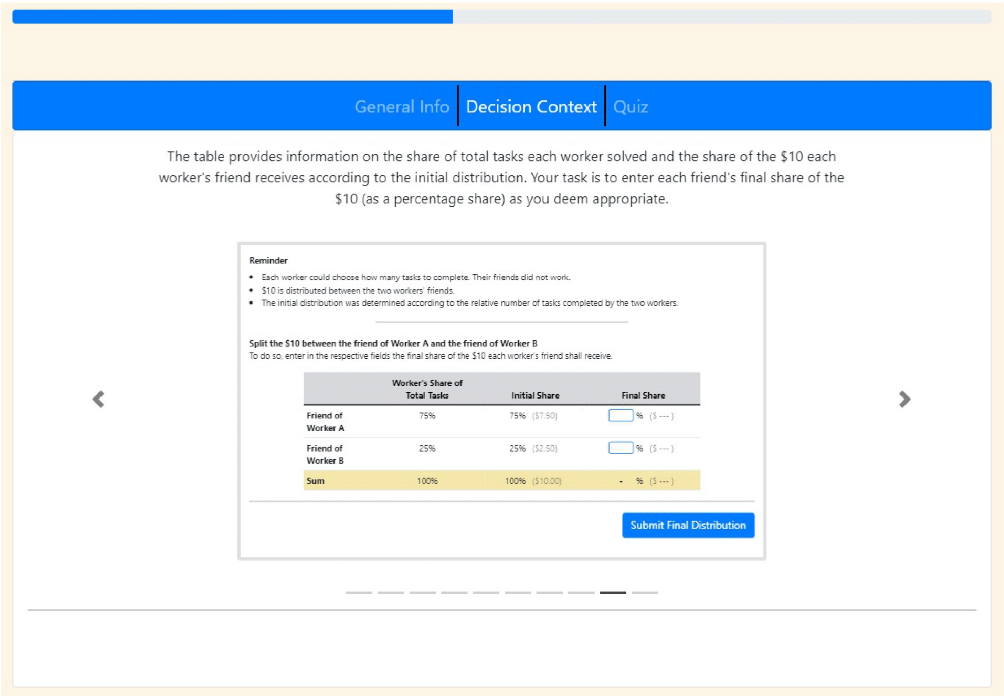

General Info
Decision Context
Quiz
After they had finished working, **$10** was initially **distributed between their friends** based on the **relative number of completed tasks**. Bill and James knew about this procedure when they started working.

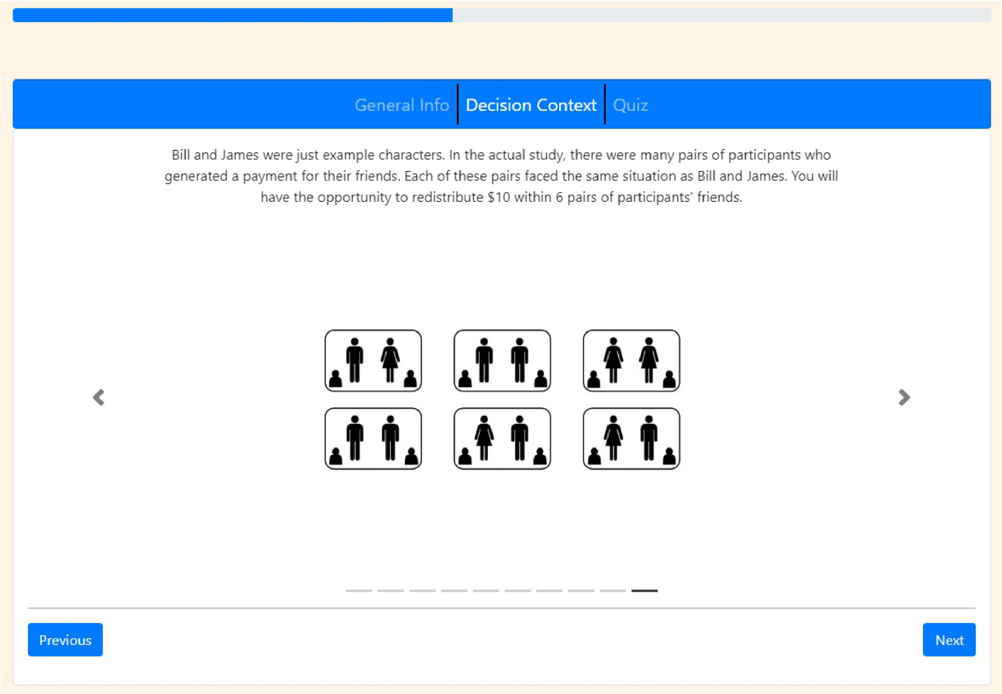

General Info
Decision Context
Quiz
For example, if Bill completed 9 tasks and James 1 task, Bill's friend received $9 initially and James's friend $1.

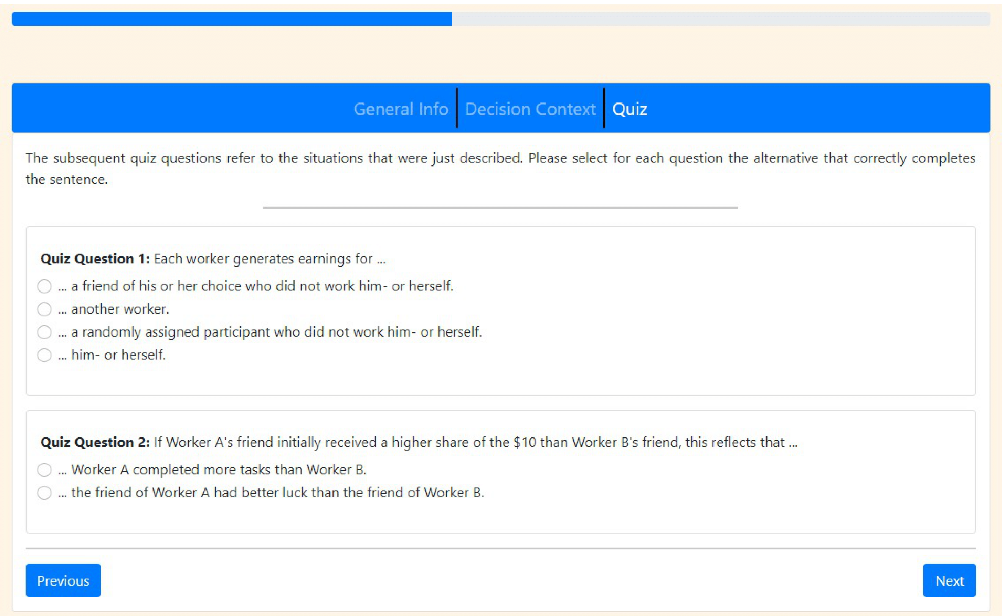

General Info
Decision Context
Quiz
Likewise, if Bill completed 4 tasks and James 16 tasks, Bill's friend received $2 initially and James's friend $8.


General Info
Decision Context
Quiz
You may find the resulting distribution of the $10 fair or unfair. Your task is to **confirm the distribution or redistribute the $10 as you deem appropriate**.
?


General Info
Decision Context
Quiz
The figure below displays the decision screen. The upper part contains a reminder about central aspects of the other study. The lower part contains a table that refers to the specific situation for which you make a decision.
Reminder
Each worker could choose how many tasks to complete. Their friends did not work.
$10 is distributed between the two workers' friends.
The initial distribution was determined according to the relative number of tasks completed by the two workers.
Split the $10 between the friend of Worker A and the friend of Worker B
To do so, enter in the respective fields the final share of the $10 each worker's friend shall receive.
Worker's Share of Total Tasks
Initial Share
Final Share
Friend of Worker A
75%
75% ($7.50)
% ($ --- )
Friend of Worker B
25%
25% ($2.50)
% ($ --- )
Sum
100%
100% ($10.00)
- % ($ --- )
Submit Final Distribution


General Info
Decision Context
Quiz
The table provides information on the share of total tasks each worker solved and the share of the $10 each worker's friend receives according to the initial distribution. Your task is to enter each friend's final share of the $10 (as a percentage share) as you deem appropriate.
Reminder
Each worker could choose how many tasks to complete. Their friends did not work.
$10 is distributed between the two workers' friends.
The initial distribution was determined according to the relative number of tasks completed by the two workers.
Split the $10 between the friend of Worker A and the friend of Worker B
To do so, enter in the respective fields the final share of the $10 each worker's friend shall receive.
Worker's Share of Total Tasks
Initial Share
Final Share
Friend of Worker A
75%
75% ($7.50)
% ($ --- )
Friend of Worker B
25%
25% ($2.50)
% ($ --- )
Sum
100%
100% ($10.00)
- % ($ --- )
Submit Final Distribution


General Info
Decision Context
Quiz
Bill and James were just example characters. In the actual study, there were many pairs of participants who generated a payment for their friends. Each of these pairs faced the same situation as Bill and James. You will have the opportunity to redistribute $10 within 6 pairs of participants' friends.
Previous
Next


General Info
Decision Context
Quiz
The subsequent quiz questions refer to the situations that were just described. Please select for each question the alternative that correctly completes the sentence.
Quiz Question 1: Each worker generates earnings for ...
... a friend of his or her choice who did not work him- or herself.
... another worker.
... a randomly assigned participant who did not work him- or herself.
... him- or herself.
Quiz Question 2: If Worker A's friend initially received a higher share of the $10 than Worker B's friend, this reflects that ...
... Worker A completed more tasks than Worker B.
... the friend of Worker A had better luck than the friend of Worker B.
Previous
Next

## Proceed to Decisions

On the following six pages, you will make your redistribution decisions for six different pairs of friends.

Before you proceed, please take a short moment to deliberately think about which distribution you find appropriate in this kind of situation.

Proceed to Decisions

## Situation 1

**Reminder**

- Each worker could choose how many tasks to complete. Their friends did not work.
- $10 is distributed between the two workers' friends.
- The initial distribution was determined according to the relative number of tasks completed by the two workers.

**Split the $10 between the friend of Worker A and the friend of Worker B**
To do so, enter in the respective fields the final share of the $10 each worker's friend shall receive.

| | Worker's Share of Total Tasks | Initial Share | Final Share |
|---|---|---|---|
| **Friend of Worker A** | 22% | 22% ($2.20) | % ($ --- ) |
| **Friend of Worker B** | 78% | 78% ($7.80) | % ($ --- ) |
| **Sum** | 100% | 100% ($10.00) | - % ($ --- ) |

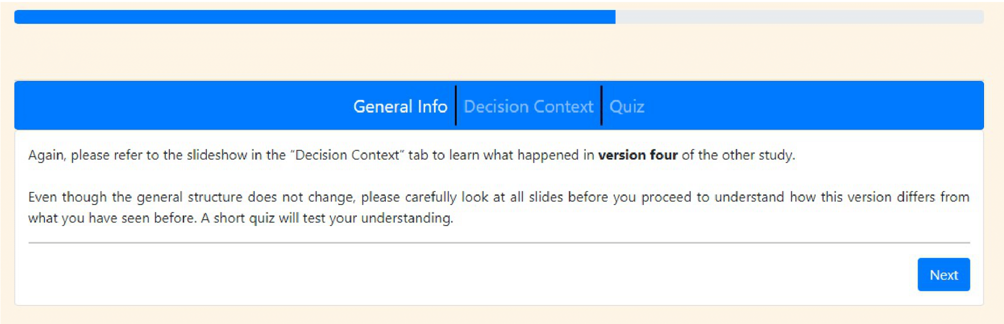
Submit Final Distribution

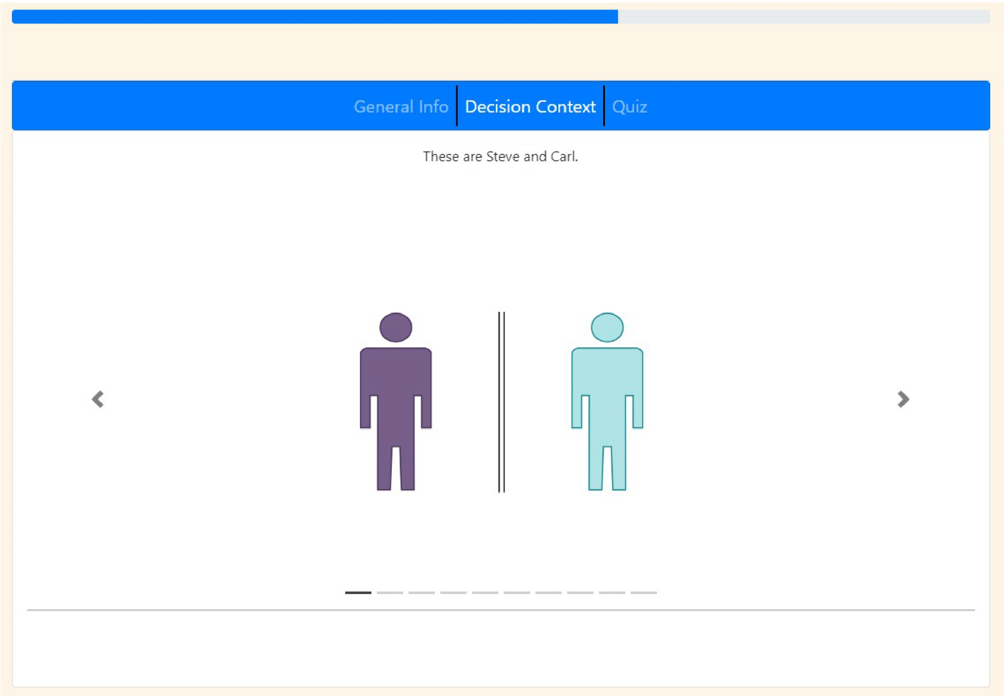
## Your Considerations

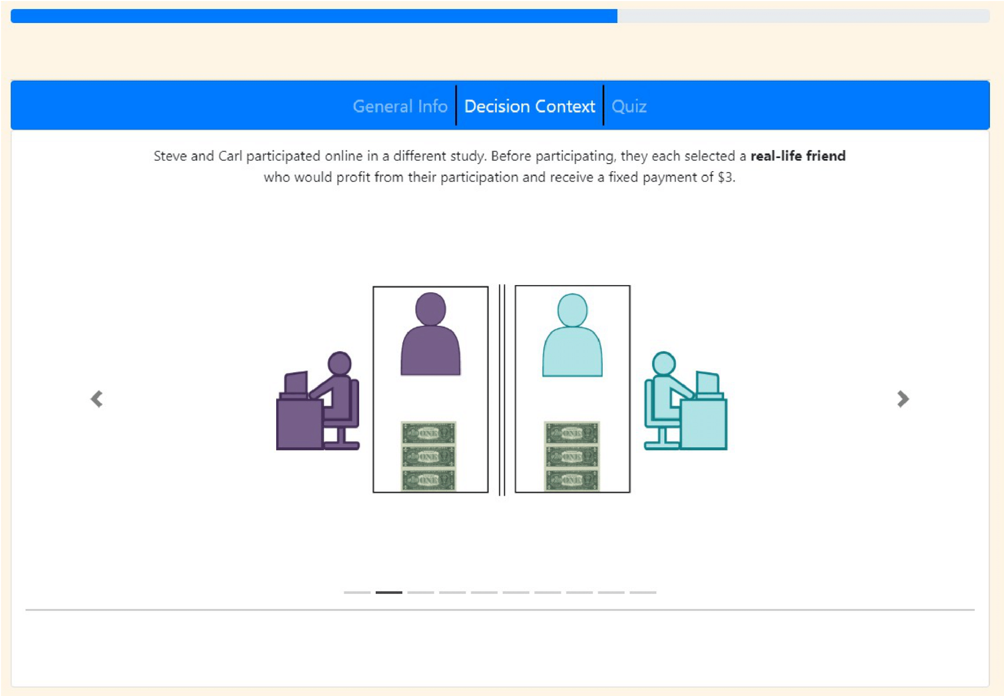
To help us understand why you consider the final payments you implemented in the past six decisions appropriate, please briefly describe the reasoning behind your decisions.

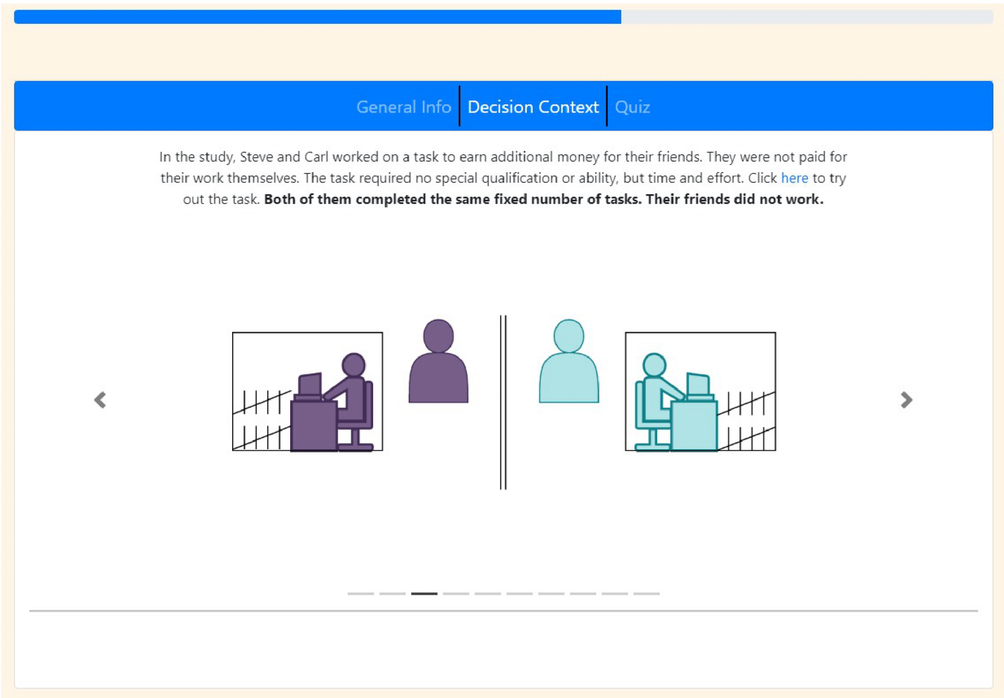
Next

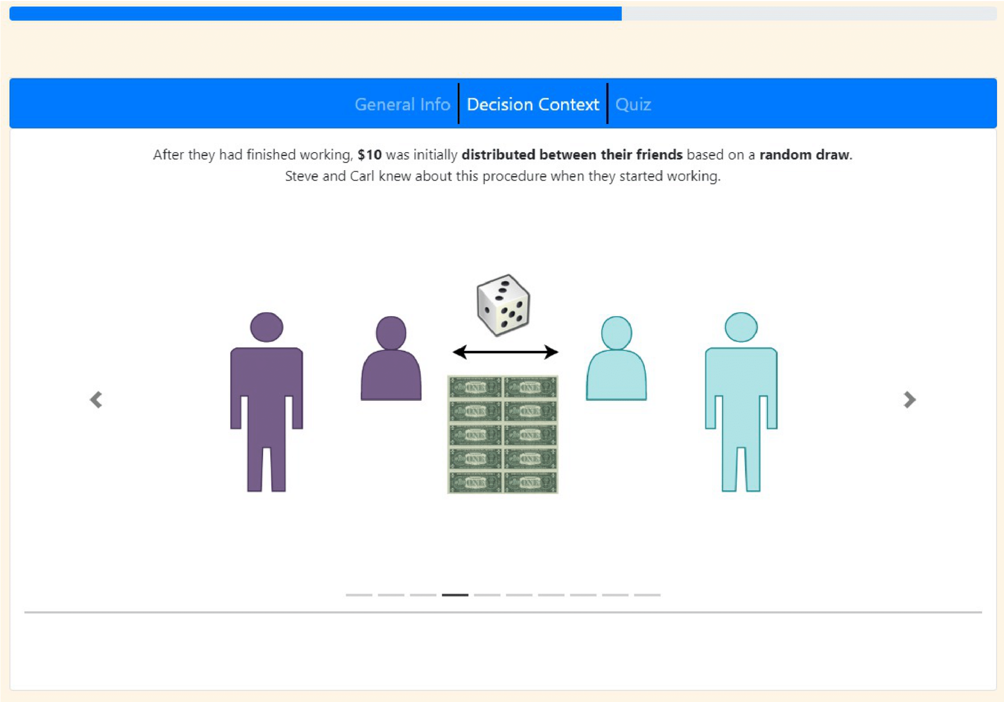
## Block 4

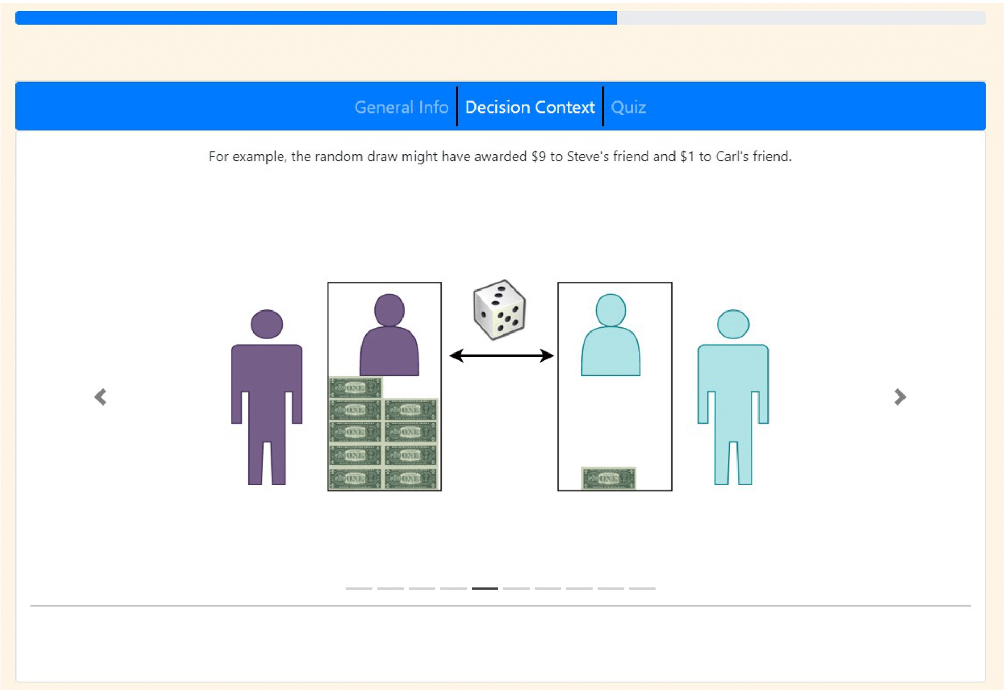

General Info
Decision Context
Quiz
Again, please refer to the slideshow in the "Decision Context" tab to learn what happened in **version four** of the other study.
Even though the general structure does not change, please carefully look at all slides before you proceed to understand how this version differs from what you have seen before. A short quiz will test your understanding.
Next

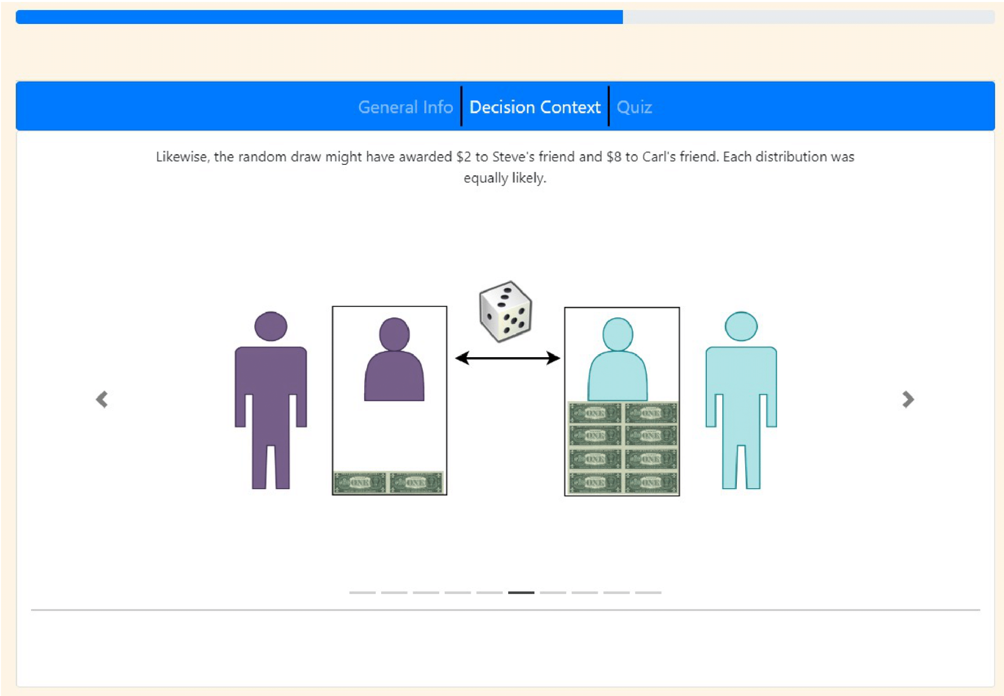

General Info
Decision Context
Quiz
These are Steve and Carl.

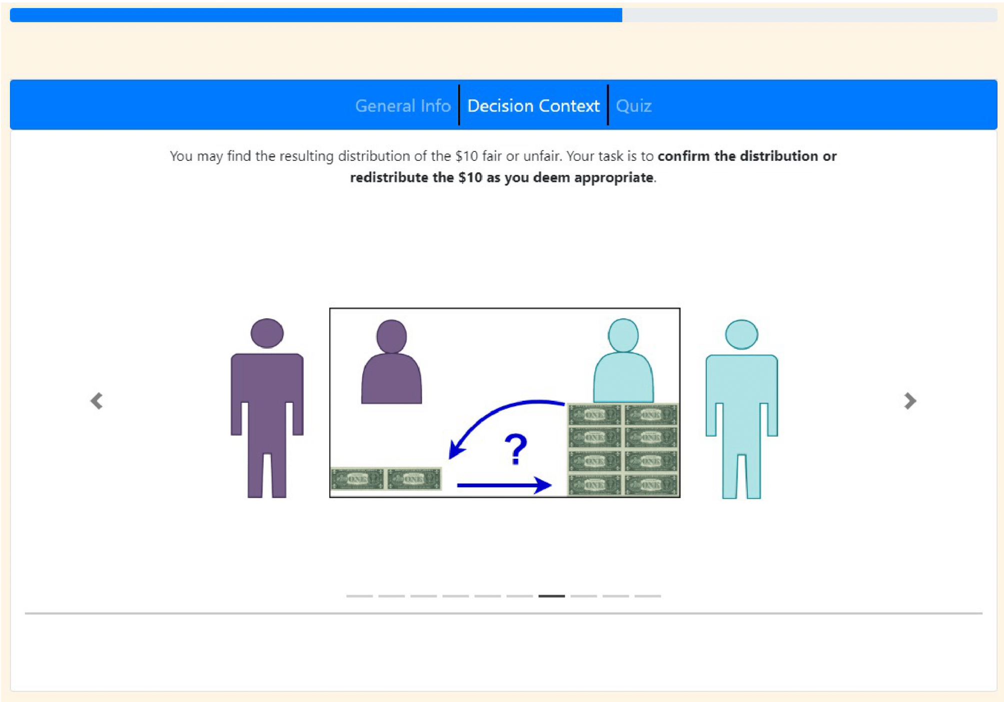

General Info
Decision Context
Quiz
Steve and Carl participated online in a different study. Before participating, they each selected a **real-life friend** who would profit from their participation and receive a fixed payment of $3.

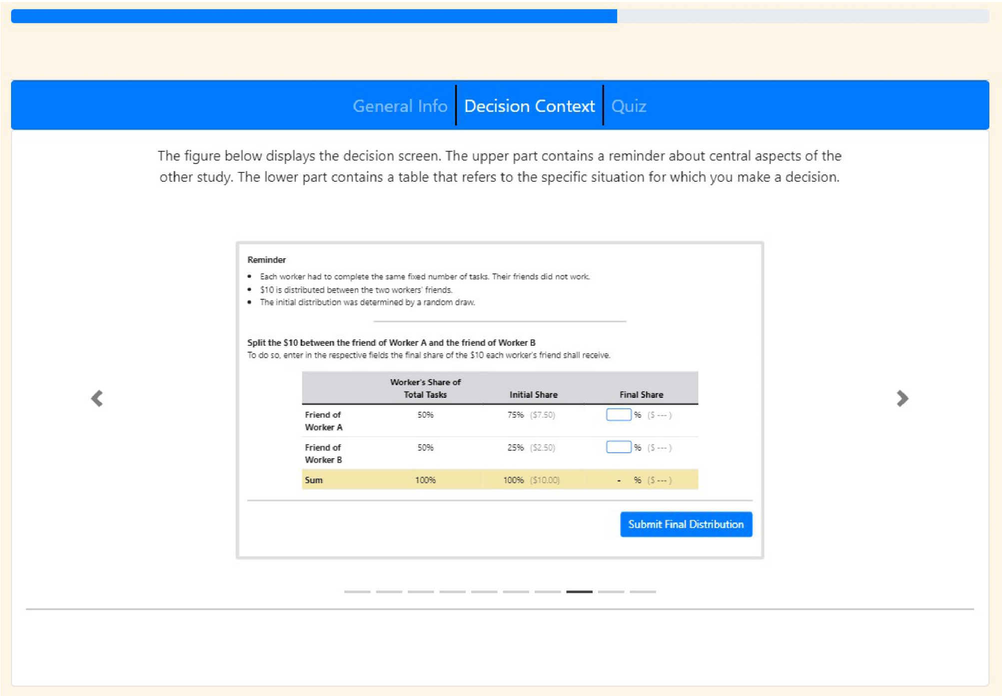

General Info
Decision Context
Quiz
In the study, Steve and Carl worked on a task to earn additional money for their friends. They were not paid for their work themselves. The task required no special qualification or ability, but time and effort. Click here to try out the task. **Both of them completed the same fixed number of tasks. Their friends did not work.**

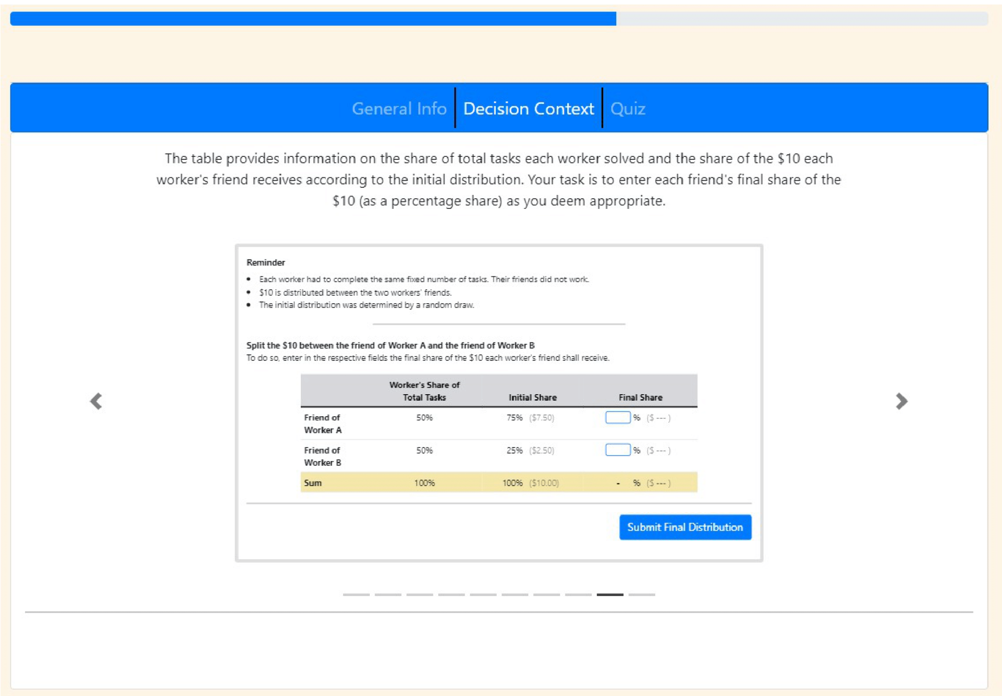

General Info
Decision Context
Quiz
After they had finished working, $10 was initially distributed between their friends based on a random draw.
Steve and Carl knew about this procedure when they started working.

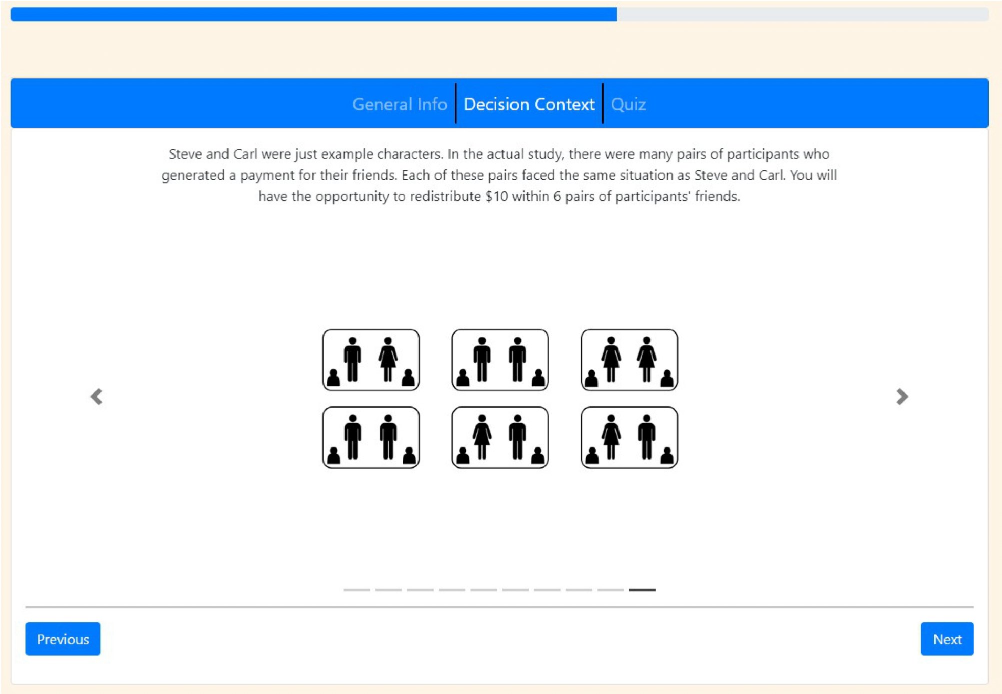

General Info
Decision Context
Quiz
For example, the random draw might have awarded $9 to Steve's friend and $1 to Carl's friend.

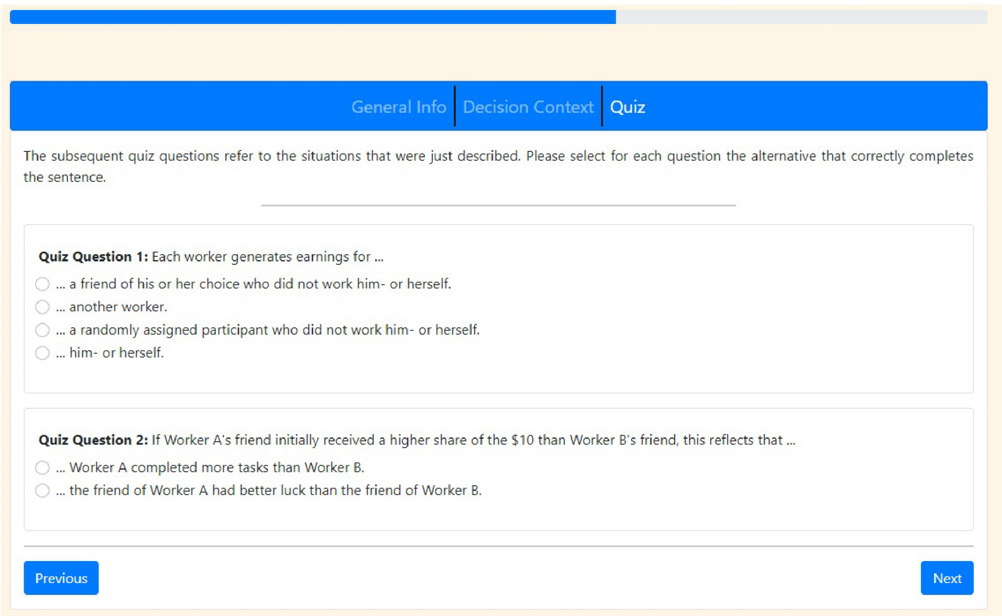

General Info
Decision Context
Quiz
Likewise, the random draw might have awarded $2 to Steve's friend and $8 to Carl's friend. Each distribution was equally likely.


General Info
Decision Context
Quiz
You may find the resulting distribution of the $10 fair or unfair. Your task is to **confirm the distribution or redistribute the $10 as you deem appropriate**.
?


General Info
Decision Context
Quiz
The figure below displays the decision screen. The upper part contains a reminder about central aspects of the other study. The lower part contains a table that refers to the specific situation for which you make a decision.
Reminder
Each worker had to complete the same fixed number of tasks. Their friends did not work.
$10 is distributed between the two workers' friends.
The initial distribution was determined by a random draw.
Split the $10 between the friend of Worker A and the friend of Worker B
To do so, enter in the respective fields the final share of the $10 each worker's friend shall receive.
Worker's Share of Total Tasks
Initial Share
Final Share
Friend of Worker A
50%
75% ($7.50)
% ($ ---)
Friend of Worker B
50%
25% ($2.50)
% ($ ---)
Sum
100%
100% ($10.00)
- % ($ ---)
Submit Final Distribution


General Info
Decision Context
Quiz
The table provides information on the share of total tasks each worker solved and the share of the $10 each worker's friend receives according to the initial distribution. Your task is to enter each friend's final share of the $10 (as a percentage share) as you deem appropriate.
Reminder
Each worker had to complete the same fixed number of tasks. Their friends did not work.
$10 is distributed between the two workers' friends.
The initial distribution was determined by a random draw.
Split the $10 between the friend of Worker A and the friend of Worker B
To do so, enter in the respective fields the final share of the $10 each worker's friend shall receive.
Worker's Share of Total Tasks
Initial Share
Final Share
Friend of Worker A
50%
75% ($7.50)
% ($ ---)
Friend of Worker B
50%
25% ($2.50)
% ($ ---)
Sum
100%
100% ($10.00)
- % ($ ---)
Submit Final Distribution


General Info
Decision Context
Quiz
Steve and Carl were just example characters. In the actual study, there were many pairs of participants who generated a payment for their friends. Each of these pairs faced the same situation as Steve and Carl. You will have the opportunity to redistribute $10 within 6 pairs of participants' friends.
Previous
Next


General Info
Decision Context
Quiz
The subsequent quiz questions refer to the situations that were just described. Please select for each question the alternative that correctly completes the sentence.
Quiz Question 1: Each worker generates earnings for ...
... a friend of his or her choice who did not work him- or herself.
... another worker.
... a randomly assigned participant who did not work him- or herself.
... him- or herself.
Quiz Question 2: If Worker A's friend initially received a higher share of the $10 than Worker B's friend, this reflects that ...
... Worker A completed more tasks than Worker B.
... the friend of Worker A had better luck than the friend of Worker B.
Previous
Next

## Proceed to Decisions

On the following six pages, you will make your redistribution decisions for six different pairs of friends.

Before you proceed, please take a short moment to deliberately think about which distribution you find appropriate in this kind of situation.

Proceed to Decisions

## Situation 1

**Reminder**

- Each worker had to complete the same fixed number of tasks. Their friends did not work.
- $10 is distributed between the two workers' friends.
- The initial distribution was determined by a random draw.

**Split the $10 between the friend of Worker A and the friend of Worker B**
To do so, enter in the respective fields the final share of the $10 each worker's friend shall receive.

| | Worker's Share of Total Tasks | Initial Share | Final Share |
|---|---|---|---|
| **Friend of Worker A** | 50% | 0% ($0.00) | % ($ --- ) |
| **Friend of Worker B** | 50% | 100% ($10.00) | % ($ --- ) |
| **Sum** | 100% | 100% ($10.00) | - % ($ --- ) |

Submit Final Distribution

## Your Considerations

To help us understand why you consider the final payments you implemented in the past six decisions appropriate, please briefly describe the reasoning behind your decisions.

Next

*The following pages were shown to all subjects in the same order as in this document.*

## Make a Guess

To conclude this section, we have two guessing questions for you.

**Context**
Unrelated to their work on the task, we asked workers in the other study to make another decision.

- We asked those workers who could earn money for themselves how they would distribute $10 between themselves and the worker they were matched to if they could freely decide.
- Similarly, we asked those workers who could earn money for their friends how they would distribute $10 between their own friend and the friend of the worker they were matched to.

As we had announced to the workers before, we will randomly select one of them and implement his or her choice.

**Your Task & Reward for Correct Guesses**
Your task is to guess for both groups of workers how they distributed the $10 on average. For each correct guess, you are rewarded with a bonus of $0.20. A guess will count as correct if it is at most $0.20 away from the corresponding true average distribution.

**Guessing Question 1**
Please use the slider below to make your guess about the group of workers who earned money for themselves.

On average, how much of the $10 did these workers keep for themselves, and how much did they give to the other worker?

$x for themselves | $x for the other worker

**Guessing Question 2:**
Please use the slider below to make your guess about the group of workers who earned money for their friends.

On average, how much of the $10 did these workers give to their own friend, and how much did they give to the friend of the other worker?

$x for their own friend | $x for the friend of the other worker

Previous | Next

## Instructions

You have reached the final part of this study!

In this part, we are mostly asking for **your personal view** – in these cases, there are no right or wrong answers.

Please click on the "Next" button to proceed.

Next

## Your Beliefs and Assessment

Please complete the statements below by selecting the option which corresponds to your view most closely.

If one person receives more than another due to **having better luck**, I find that ...

clearly unfair | unfair | rather unfair | rather fair | fair | clearly fair

If one person receives more than another due to **exerting higher effort**, I find that ...

clearly unfair | unfair | rather unfair | rather fair | fair | clearly fair

If children born to affluent parents are more likely to be affluent themselves later in life compared to children born to less well-off parents, I find that ...

clearly unfair | unfair | rather unfair | rather fair | fair | clearly fair

How much do you agree or disagree with the following statement? "Last year all citizens of the USA received exactly the same income."

strongly disagree | disagree | agree | strongly agree

Next

## Your Policy Views

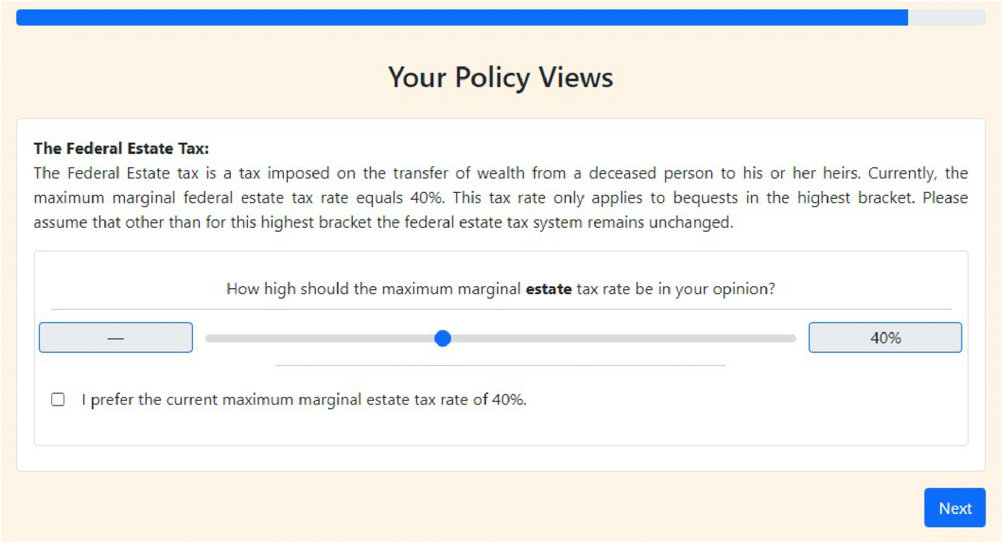
**The Federal Income Tax**
Currently, the maximum marginal federal income tax rate is 37%. This tax rate only applies to incomes in the highest tax bracket. Please assume that other than for this highest bracket the federal income tax system remains unchanged.

How high should the maximum marginal **income** tax rate be in your opinion?

— | 37%

☐ I prefer the current maximum marginal income tax rate of 37%.

Next


Your Policy Views

**The Federal Estate Tax:**
The Federal Estate tax is a tax imposed on the transfer of wealth from a deceased person to his or her heirs. Currently, the maximum marginal federal estate tax rate equals 40%. This tax rate only applies to bequests in the highest bracket. Please assume that other than for this highest bracket the federal estate tax system remains unchanged.

How high should the maximum marginal **estate** tax rate be in your opinion?

— 40%

☐ I prefer the current maximum marginal estate tax rate of 40%.

Next

## Your Policy Views

For the questions on this page, please complete the sentences by selecting the option that most closely corresponds to your view.

The federal-state unemployment insurance (UI) system uses tax money to provide assistance to people who have lost their jobs and are eligible for benefits by temporarily replacing part of their wages.

The unemployment insurance (UI) system should be ...

significantly reduced | reduced | moderately reduced | neither reduced nor extended | moderately extended | extended | significantly extended

The government uses tax money to run programs such as the Social Security program and the Supplemental Security Income disability program which provide assistance to people with disabilities.

Government funding for programs that provide assistance to disabled people should be ...

significantly reduced | reduced | moderately reduced | neither reduced nor extended | moderately extended | extended | significantly extended

The government uses tax money to finance institutions and runs programs to – among other things – provide assistance to children from less well-off families; examples include public schools, colleges and universities, tuition waivers, and health coverage programs such as Medicaid and the Children's Health Insurance Program (CHIP).

Government funding for these institutions and programs should be ...

significantly reduced | reduced | moderately reduced | neither reduced nor extended | moderately extended | extended | significantly extended

Next

## About Yourself

Now we would like to ask you a few things about yourself.

Please select your state of residence.

--------

Please select your educational attainment.

- No High School degree
- High school diploma equivalent
- Bachelor's or Associate's degree
- Master's degree or higher

How much was your pre-tax (gross) household income between January and December 2021? In case you are not sure, please provide your best estimate.

- Below $34,000
- Between $34,000 and $68,000
- Between $68,000 and $125,000
- More than $125,000

Please estimate your household's net worth (value of all assets - sum of all liabilities) and indicate in which of the below categories your estimate falls.

- Less than $13,000.
- More than $13,000 but less than $124,000.
- More than $124,000 but less than $410,000.
- More than $410,000.

If you had to use one of these five commonly-used terms to describe your social class, which one would it be?

- Lower Class or Poor
- Working Class
- Middle Class
- Upper-middle Class
- Upper Class

On economic policy matters, where do you see yourself on the liberal/conservative spectrum?

- Very liberal
- Liberal
- Rather liberal
- Moderate
- Rather conservative
- Conservative
- Very conservative

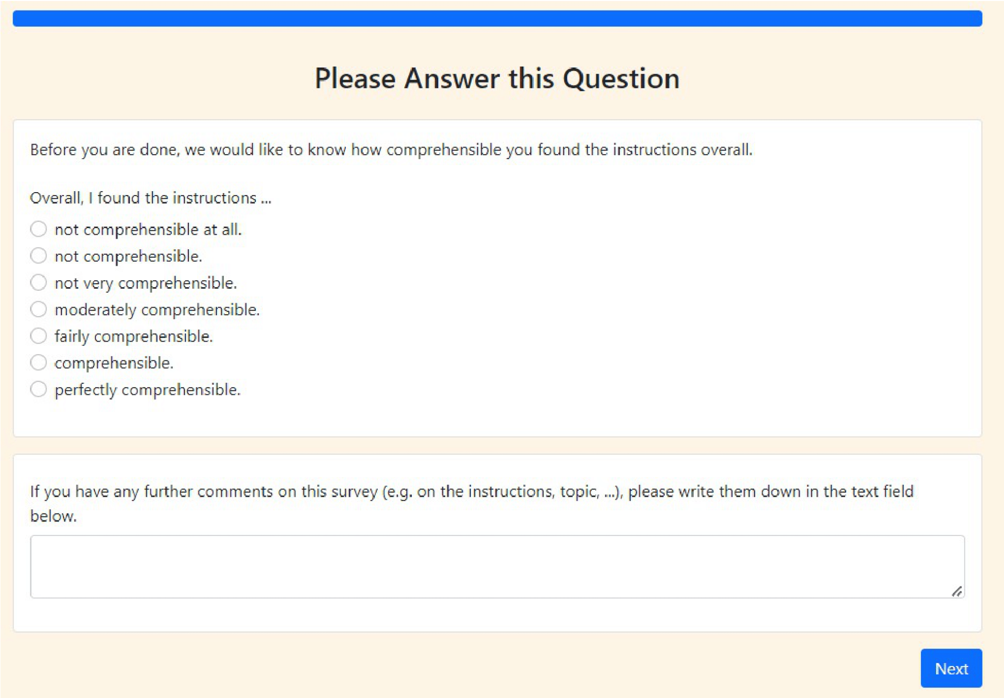
Generally speaking, where do you see yourself on the Republican/Democrat spectrum?

- Clearly Republican
- Republican
- Rather Republican
- Neither Republican nor Democrat
- Rather Democrat
- Democrat
- Clearly Democrat

Which of the options below best describes how regularly you vote?

- I do not vote in elections
- Rarely
- Some elections
- Approximately every other election
- I may have missed a few
- Almost every election
- Every election without exception

Next



## Please Answer this Question

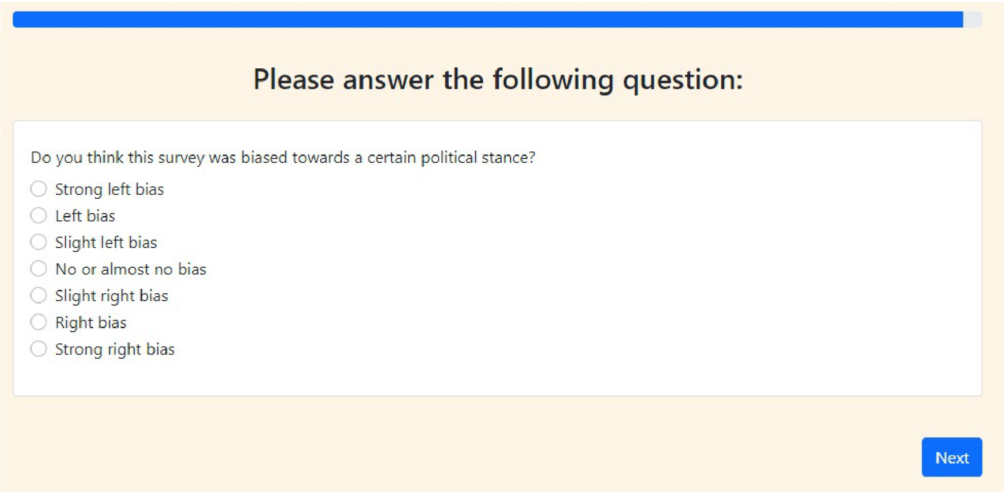
Before you are done, we would like to know how comprehensible you found the instructions overall.

Overall, I found the instructions ...

- not comprehensible at all.
- not comprehensible.
- not very comprehensible.
- moderately comprehensible.
- fairly comprehensible.
- comprehensible.
- perfectly comprehensible.

If you have any further comments on this survey (e.g. on the instructions, topic, ...), please write them down in the text field below.

Next



## Please answer the following question:

Do you think this survey was biased towards a certain political stance?

- Strong left bias
- Left bias
- Slight left bias
- No or almost no bias
- Slight right bias
- Right bias
- Strong right bias

Next

*After spectators clicked the "Next" button on the last page, they were redirected to the Prolific platform.*

## References

**Alesina, Alberto, and George-Marios Angeletos.** 2005. "Fairness and Redistribution." *American Economic Review* 95 (4): 960–80. https://doi.org/10.1257/0002828054825655. [10]

**Alesina, Alberto, and Paola Giuliano.** 2011. "Preferences for Redistribution." In *Handbook of Social Economics,* 1: 93–131. North-Holland: Elsevier B.V. https://doi.org/10.1016/B978-0-444-53187-2.00004-8. [10]

**Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso.** 2018. "Intergenerational Mobility and Preferences for Redistribution." *American Economic Review* 108 (2): 521–54. https://doi.org/10.1257/aer.20162015. [10, 43]

**Almås, Ingvild, Alexander W. Cappelen, Erik Ø. Sørensen, and Bertil Tungodden.** 2022. "Global Evidence on the Selfish Rich Inequality Hypothesis." *Proceedings of the National Academy of Sciences* 119 (3). https://doi.org/10.1073/pnas.2109690119. [19]

**Almås, Ingvild, Alexander W. Cappelen, and Bertil Tungodden.** 2020. "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" *Journal of Political Economy* 128 (5): 1753–88. https://doi.org/10.1086/705551. [7, 11, 13, 16]

**Andre, Peter.** 2022. "Shallow Meritocracy." [10, 43]

**Bénabou, Roland.** 2000. "Meritocracy, Redistribution and the Size of the Pie." Chap. 12 in *Meritocracy and Economic Inequality,* edited by Kenneth J. Arrow, Samuel Bowles, and Steven N. Durlauf, 317–39. Princeton University Press. [10]

**Bernhard, Helen, Urs Fischbacher, and Ernst Fehr.** 2006. "Parochial Altruism in Humans." *Nature* 442: 912–15. https://doi.org/10.1038/nature04981. [41]

**Björklund, Anders, Jesper Roine, and Daniel Waldenström.** 2012. "Intergenerational Top Income Mobility in Sweden: Capitalist Dynasties in the Land of Equal Opportunity?" *Journal of Public Economics* 96 (5-6): 474–84. https://doi.org/10.1016/j.jpubeco.2012.02.003. [10]

**Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch.** 2014. "Hroot: Hamburg Registration and Organization Online Tool." *European Economic Review* 71: 117–20. https://doi.org/10.1016/j.euroecorev.2014.07.003. [16]

**Bowles, Samuel, and Herbert Gintis.** 2002. "The Inheritance of Inequality." *Journal of Economic Perspectives* 16 (3): 3–30. https://doi.org/10.1257/089533002760278686. [10]

**Cappelen, Alexander W., Benjamin Enke, and Bertil Tungodden.** 2022. "Moral Universalism: Global Evidence." [41, 42]

**Cappelen, Alexander W., Astri Drange Hole, Erik Ø. Sorensen, and Bertil Tungodden.** 2007. "The Pluralism of Fairness Ideals: An Experimental Approach." *American Economic Review* 97 (3): 818–27. https://doi.org/10.1257/aer.97.3.818. [10]

**Cappelen, Alexander W., James Konow, Erik Ø. Sørensen, and Bertil Tungodden.** 2013. "Just Luck: An Experimental Study of Risk-Taking and Fairness." *American Economic Review* 103 (4): 1398–413. https://doi.org/10.1257/aer.103.4.1398. [7, 10, 11, 13]

**Cappelen, Alexander W., John List, Anya Samek, and Bertil Tungodden.** 2020. "The Effect of Early-Childhood Education on Social Preferences." *Journal of Political Economy* 128 (7): 2739–58. https://doi.org/10.1086/706858. [8, 10]

**Cappelen, Alexander W., Johanna Mollerstrom, Bjørn-Atle Reme, and Bertil Tungodden.** 2022. "A Meritocratic Origin of Egalitarian Behaviour." *Economic Journal* 132 (646): 2101–17. https://doi.org/10.1093/ej/ueac008. [10, 19, 43]

**Cappelen, Alexander W., Erik Ø. Sørensen, and Bertil Tungodden.** 2010. "Responsibility for What? Fairness and Individual Responsibility." *European Economic Review* 54 (3): 429–41. https://doi.org/10.1016/j.euroecorev.2009.08.005. [10]

**Cappelen, Alexander W., and Bertil Tungodden.** 2017. "Fairness and the Proportionality Principle." *Social Choice and Welfare* 49: 709–19. https://doi.org/10.1007/s00355-016-1016-6. [10, 19, 43]

**Cassar, Lea, and Arnd H. Klein.** 2019. "A Matter of Perspective: How Failure Shapes Distributive Preferences." *Management Science* 65 (11): 5050–64. https://doi.org/10.1287/mnsc.2018.3185. [37, 43]

**Chen, Daniel L., Martin Schonger, and Chris Wickens.** 2016. "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments." *Journal of Behavioral and Experimental Finance* 9: 88–97. https://doi.org/10.1016/j.jbef.2015.12.001. [16]

**Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz.** 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106 (4): 855–902. https://doi.org/10.1257/aer.20150572. [10]

**Cohen, Rachelle, Amnon Maltz, and Yuval Ofek-Shanny.** 2022. "Social Preferences over Taste-Based Inequality." [10, 11, 43]

**Deffains, Bruno, Romain Espinosa, and Christian Thöni.** 2016. "Political Self-Serving Bias and Redistribution." *Journal of Public Economics* 134: 67–74. https://doi.org/10.1016/j.jpubeco.2016.01.002. [37, 43]

**Eisenkopf, Gerald, Urs Fischbacher, and Franziska Föllmi-Heusi.** 2013. "Unequal Opportunities and Distributive Justice." *Journal of Economic Behavior and Organization* 93: 51–61. https://doi.org/10.1016/j.jebo.2013.07.011. [43]

**Gächter, Simon, Chris Starmer, and Fabio Tufano.** 2015. "Measuring the Closeness of Relationships: A Comprehensive Evaluation of the 'Inclusion of the Other in the Self' Scale." *PLoS ONE* 10 (6). https://doi.org/10.1371/journal.pone.0129478. [14]

**Gärtner, Manja, Johanna Mollerstrom, and David Seim.** 2017. "Individual Risk Preferences and the Demand for Redistribution." *Journal of Public Economics* 153: 49–55. https://doi.org/10.1016/j.jpubeco.2017.06.009. [10]

**Gill, David, and Victoria Prowse.** 2012. "A Structural Analysis of Disappointment Aversion in a Real Effort Competition." *American Economic Review* 102 (1): 469–503. https://doi.org/10.1257/aer.102.1.469. [14]

**Konow, James.** 2000. "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." *American Economic Review* 90 (4): 1072–92. [7, 13, 37, 43]

**Konow, James.** 2001. "Fair and Square: The Four Sides of Distributive Justice." *Journal of Economic Behavior and Organization* 46 (2): 137–64. https://doi.org/10.1016/S0167-2681(01)00194-9. [42]

**Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk.** 2020. "The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy* 128 (2): 434–67. https://doi.org/10.1086/704386. [10]

**Kube, Sebastian, and Christian Traxler.** 2011. "The Interaction of Legal and Social Norm Enforcement." *Journal of Public Economic Theory* 13 (5): 639–60. https://doi.org/10.1111/j.1467-9779.2011.01515.x. [16]

**Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva.** 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." *American Economic Review* 105 (4): 1478–508. https://doi.org/10.1257/aer.20130360. [10]

**Mollerstrom, Johanna, Bjørn-Atle Reme, and Erik Ø. Sørensen.** 2015. "Luck, Choice and Responsibility - An Experimental Study of Fairness Views." *Journal of Public Economics* 131: 33–40. https://doi.org/10.1016/j.jpubeco.2015.08.010. [10]

**Obama, Barack.** 2006. *The Audacity of Hope: Thoughts on Reclaiming the American Dream.* New York: Crown. [6]

**Palan, Stefan, and Christian Schitter.** 2018. "Prolific.Ac - A Subject Pool for Online Experiments." *Journal of Behavioral and Experimental Finance* 17: 22–27. https://doi.org/10.1016/j.jbef.2017.12.004. [17]

**Peer, Eyal, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer.** 2022. "Data Quality of Platforms and Panels for Online Behavioral Research." *Behavior Research Methods,* no. 54, 1643–62. https://doi.org/10.3758/s13428-021-01694-3. [17]

**Piketty, Thomas, and Emmanuel Saez.** 2013. "A Theory of Optimal Inheritance Taxation." *Econometrica* 81 (5): 1851–86. https://doi.org/10.3982/ecta10712. [10]

**Posner, Richard A.** 1972. *Economic Analysis of Law.* Boston: Little, Brown & Co. [6]

**Rodriguez-Lara, Ismael, and Luis Moreno-Garrido.** 2012. "Self-Interest and Fairness: Self-Serving Choices of Justice Principles." *Experimental Economics* 15: 158–75. https://doi.org/10.1007/s10683-011-9295-3. [37, 43]

**Roth, Christopher, and Johannes Wohlfart.** 2018. "Experienced Inequality and Preferences for Redistribution." *Journal of Public Economics* 167: 251–62. https://doi.org/10.1016/j.jpubeco.2018.09.012. [10]

**Schaube, Sebastian, and Louis Strang.** 2022. "(Not) Everyone Can Be a Winner – The Role of Payoff Interdependence for Redistribution." [10, 16, 19]

**Schwaiger, Rene, Jürgen Huber, Michael Kirchler, Daniel Kleinlercher, and Utz Weitzel.** 2022. "Unequal Opportunities, Social Groups, and Redistribution: Evidence from Germany." *Journal of Behavioral and Experimental Economics* 100: 101911. https://doi.org/10.1016/j.socec.2022.101911. [43]

**Stantcheva, Stefanie.** 2021. "Understanding Tax Policy: How Do People Reason?" *Quarterly Journal of Economics* 136 (4): 2309–69. https://doi.org/10.1093/qje/qjab033. [7, 10, 40]

**Stantcheva, Stefanie.** 2022. "How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible." [18]

**Valero, Vanessa.** 2022. "Redistribution and Beliefs about the Source of Income Inequality." *Experimental Economics* 25 (3): 876–901. https://doi.org/10.1007/s10683-021-09733-8. [37, 43]

# Chapter 2

# Prosociality predicts individual behavior and collective outcomes in the COVID-19 pandemic*

*Joint with Ximeng Fang, Chui-Yee Ho, Zihua Chen, and Lorenz Götte*

## 2.1 Introduction

To curb the COVID-19 pandemic, individuals have to engage in costly preventive behaviors such as reducing social contacts, wearing face masks, or using contact tracing apps. However, the benefits from a lower rate of transmission accrue to society at large and thus constitute a public good. This results in a social dilemma, where "the maximization of short-term self-interest yields outcomes leaving all participants worse off than feasible alternatives." (Ostrom, 1998, p.1). In this sense, the pandemic is comparable to other collective action problems such as civic engagement or the fight against climate change.

Which factors determine the success of groups or societies in overcoming collective action problems has been a long-standing question in the social sciences. One plausible determinant is the extent to which individual members are prosocial, i.e., how willing they are to behave in a way that primarily benefits other people or society at large. Prosocial individuals may help their groups in achieving more beneficial outcomes in the face of social dilemmas, both by contributing

* We would like to thank Peter Andre, Felix Chopra, Thomas Dohmen, Luca Henkel, Sven Heuser, Sebastian Kube, Anna Schulze-Tilling, as well as participants of the IAME Applied Micro Coffee for helpful comments and suggestions already at early stages of this project. Financial support by the Deutsche Forschungsgemeinschaft (DFG) through CRC TR 224 (Project B07) and Germany's Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged.

This paper is published as Fang, X., Freyer, T., Ho, C. Y., Chen, Z., & Goette, L., 2022. "Prosociality predicts individual behavior and collective outcomes in the COVID-19 pandemic." *Social Science & Medicine*, 308, 115192. https://doi.org/10.1016/j.socscimed.2022.115192.

more to a common cause themselves and by increasing cooperation rates among other members — for example through establishing and enforcing corresponding social norms (Fehr and Gächter, 2002; Fehr and Fischbacher, 2003; Fischbacher and Gächter, 2010; Albrecht, Kube, and Traxler, 2018; Fehr and Schurtenberger, 2018). Previous studies have documented associations between (pro-)social preferences and, amongst others, pro-environmental behavior (Andre et al., 2021; Fuhrmann-Riebel, D'Exelle, and Verschoor, 2021; Lades, Laffan, and Weber, 2021), donation and volunteering decisions (Falk et al., 2018), redistributive voting (Epper, Fehr, and Senn, 2020), as well as labor market outcomes (Dohmen et al., 2008; Burks, Carpenter, and Goette, 2009; Kosse and Tincani, 2020). However, combining data of both individual- and group-level behavior and outcomes under collective action problems in real-world contexts remains challenging.

In this paper, we examine the relationship between prosociality and individual behavior as well as collective health outcomes in the context of the COVID-19 pandemic. When fighting the pandemic, governments and public health experts have recurringly appealed to people's altruistic motivations to protect others from getting infected by embracing voluntary behavioral changes. More prosocial individuals may be more likely to respond to (and propagate) such norms and appeals, and they may generally be more inclined to internalize the health externalities that their behavior imposes on others. Consistent with this, studies have found that more prosocial individuals tend to follow social distancing and hygiene guidelines more stringently (van Hulsen, Rohde, and van Exel, 2020; Campos-Mercade et al., 2021; Müller and Rau, 2021). One implication is that regions with higher average levels of prosociality in the population might be more successful in slowing the spread of the virus. This is also proposed theoretically in recent susceptible-infected-recovered (SIR) models with endogenous behavior (Alfaro et al., 2021b; Farboodi, Jarosch, and Shimer, 2021; Quaas et al., 2021). Indeed, some empirical studies provide evidence that proxies for social (or civic) capital are related to mobility flows and COVID-19 incidence rates at the subnational level (Bartscher et al., 2020; Alfaro et al., 2021a; Barrios et al., 2021; Durante, Guiso, and Gulino, 2021; Makridis and Wu, 2021), but they do not combine regional-level associations with individual-level data.

We study the role of prosociality in the COVID-19 pandemic by employing data from a representative online survey in Germany ($n = 5,843$) that we conducted during the second coronavirus wave, between mid-November and mid-December 2020. This period was characterized by steeply increasing incidence rates and a relatively lenient "lockdown light". To measure individuals' public health behavior (PHB) during that time, we included a series of questions about the extent to which they engage in physical distancing, mask-wearing, precautionary hygiene measures, self-quarantining, etc., which we then combine into a single index variable of PHB by means of a factor analysis. Although imperfect, self-reported PHB measures such as ours have been shown to be good indicators of actual behavior in the pandemic (Jensen, 2020; Gollwitzer et al., 2021). We further use experimentally-validated survey measures by Falk et al. (2016) to elicit different components of individuals' prosocial preferences and beliefs — altruism, trust, positive reciprocity, and indirect (negative) reciprocity — and collapse them into single summary measure of "prosociality".

Our data confirms that prosociality is strongly positively related to compliance with recommended social distancing and hygiene measures. Due to the large sample size, we can further aggregate our survey measures to regional-level averages across NUTS-2 regions in Germany and link them to official statistical data on COVID-19 incidence and deaths reported by the Robert-Koch-Institut (RKI), the federal government agency and research institute responsible for disease control and prevention in Germany. Our focus on within-country variation has the advantage that policy mandates and regulations in response to the pandemic remain largely similar. We find that the individual-level relation between prosociality and PHB translates into better health outcomes at the regional level — the spread of Sars-CoV-2 is slower in regions where average prosociality in the population is high. This relationship is mediated by compliance with public health measures, which supports our suggested pathway of prosociality leading to greater PH compliance, which in turn leads to lower incidence rates.

## 2.2 Theoretical Predictions

The rates of social contact and disease transmission are key parameters in epidemiological models, namely the susceptible-infected-recovered (SIR) model and its various modifications (Kermack and McKendrick, 1927; Keeling and Rohani, 2011), but they are typically determined exogenously and do not respond to voluntary behavioral adaptation by individuals in a pandemic.

Canonical SIR models can be extended by endogenizing behavioral responses of forward-looking agents who face a trade-off between utility from social contacts and disutility from increased risk of getting infected (e.g., Bauch and Earn, 2004; Fenichel et al., 2011; Jones, Philippon, and Venkateswaran, 2021). To protect themselves, individuals may choose to engage in preventive health behaviors even in the absence of government restrictions. However, individuals' actions also impose health externalities on others, and social costs of infections can exceed private costs significantly — e.g. for young and healthy individuals in the COVID-19 pandemic. Hence, behavioral adaption due to purely self-interested motives (i.e., avoiding to get infected) only flattens the infection trajectory to a limited extent.

Recent theoretical studies have explicitly incorporated prosocial motives in SIR models with endogenous behavior (Alfaro et al., 2021b; Farboodi, Jarosch, and Shimer, 2021; Quaas et al., 2021). Agents in these models are not only concerned about their own health, but also about other people's health. Thus, they partially

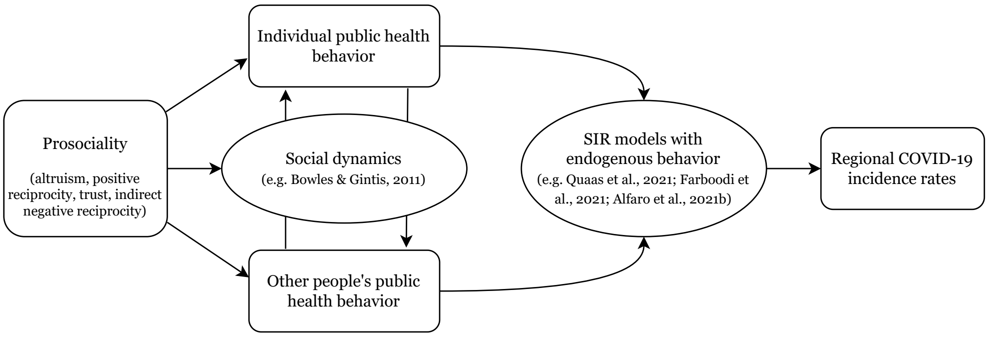



**Figure 2.1.** Framework

internalize the health risks that their own behavior imposes on susceptible individuals around them. This is particularly relevant for people who are uncertain about whether they are susceptible or infectious (e.g., due to asymptomatic cases and limited testing capacities), which applies to the majority of the population during our study period, since most people in Germany had not experienced a COVID-19 infection yet. To prevent that they unknowingly spread the virus, prosocial agents endogenously engage in lower levels of (risky) social activity.

While prosocial engagement in social distancing follows from an assumption on exogenously given preferences in these models, it can also be derived more explicitly from theories of human behavior that take a stance on where preferences to behave prosocially come from (e.g., Batson and Powell, 2003). For example, as an anonymous referee pointed out to us, a link between individuals' prosociality and their public health behavior can be explained by different variants of consistency theory (Festinger, 1957; Heider, 1958; Abelson et al., 1968). Specifically, individuals who hold strong prosocial values and attitudes may experience cognitive dissonance if they do not adjust their behavior in the pandemic accordingly.

In this empirical study, we consider several distinct components of prosociality that all reflect a positive disposition towards others: altruism, positive reciprocity, trust, and indirect (negative) reciprocity. Altruism constitutes a direct concern for others' well-being and links most closely to the above-mentioned models. Positive reciprocity is the tendency to return favors, which can facilitate norms of conditional cooperation (Bowles and Gintis, 2011). Trust is a composite trait reflecting preferences as well as beliefs about whether other people in general hold good intentions; higher generalized trust may encourage individuals to behave more prosocially towards friends and strangers alike. Indirect negative reciprocity describes the willingness to punish those who treat others unfairly and act detrimentally to the group. In the context of the pandemic, this could for example entail confronting others who disregard rules or norms regarding mask wearing and social distancing. This sort of third-party punishment can deter norm violation and free-riding and is therefore considered to be prosocial (Fehr and Gächter, 2002; Albrecht, Kube, and Traxler, 2018). In summary, as illustrated in Figure 2.1, individuals' prosocial attitudes can positively affect compliance with health measures both directly, out of concern for not (unintentionally) infecting others, as well as indirectly, through the social dynamics of cooperation and norm adoption. Thus, our first prediction is that more prosocial individuals are more likely to engage in preventive health measures in the pandemic.

Through the lens of a SIR model with endogenous behavior, increased compliance due to higher prosociality leads to a lower rate of disease transmission and thus fewer infections in the population, all else equal. In a dynamic setting, this positive effect is dampened, as lower incidence rates will reduce perceived infection risks and thus subsequent readjustment towards more social interactions. However, it can be shown that higher prosociality will still lead to a flatter infection curve in equilibrium (Alfaro et al., 2021b; Farboodi, Jarosch, and Shimer, 2021; Quaas et al., 2021). Thus, our second prediction is that infection rates will tend to be lower in regions with more prosocial individuals.

There are many other determinants of health behavior that are not considered in Figure 2.1. Importantly, the models highlight that behavior should adapt strongly to the perceived threat of COVID-19, which can vary based on the contemporaneous regional incidence rates and based on heterogeneity in expected health/mortality risks, e.g. due to age. Furthermore, time and risk preferences also play a role, as more patient individuals place a higher weight on future risks of infection (relative to immediate utility from social interactions) and more risk averse individuals shy away from uncertain consequences of a potential infection. Indeed, previous empirical studies have found positive associations of patience and risk aversion with better health behaviors and outcomes both in the COVID-19 pandemic (e.g., Chan et al., 2020; Alfaro et al., 2021a) and in other health-related domains such as smoking or obesity (e.g., Khwaja, Sloan, and Salm, 2006; Burks et al., 2012; Sutter et al., 2013; de Oliveira et al., 2016).

## 2.3 Data and Measurements

### 2.3.1 Survey Data

We partnered with the market research firm Dynata to recruit a target sample of 6000 German participants and conducted our web-based survey between November 11 to December 17, 2020. Participants were invited via email and sampled using demographic quotas on age, gender, and state, to achieve national-level representativeness of the population aged 18 to 65. Our final analysis sample consists of 5,843 responses that fulfilled the quality criteria for inclusion in the analysis: a minimum response duration, passing an attention check, no inconsistencies in demographic information, and no excessive straightlining.

To measure health behavior in the pandemic, we obtain responses (on a 7-point Likert scale) to ten questions about subjects' social distancing, hygiene behavior, etc. These questions were selected based on public health guidelines in Germany at that time. Using responses to these questions, we then construct an index by factor analysis. This index is our main measure of compliance to PHB. The eigenvalue of the first factor is 4.47 (0.25 for the second factor), which points towards a single underlying factor driving adherence to different PH measures. The Cronbach's $\alpha$ is 0.87, indicating that the different aspects of PHB are strongly interrelated.

We elicited subjects' time, risk, and social preferences using experimentally validated measures that have been employed in a large-scale representative global survey (Falk et al., 2016; Falk et al., 2018). Although the validation was conducted in a German student sample, it is plausible that the measures remain informative in our context, as language and culture are constant and there is no evidence that insights from student experiments fundamentally misrepresent behavior in the general population (Exadaktylos, Espín, and Branas-Garza, 2013; Falk, Meier, and Zehnder, 2013). To construct an individual-level measure of prosociality, we follow Falk et al. (2018) and Kosse and Tincani (2020) and combine several facets of social preferences and beliefs — altruism, trust, positive reciprocity, and indirect (negative) reciprocity — into one index variable by extracting their first principal component (eigenvalue $= 1.789$). This component places positive weight on all input variables and is thus congruent with the common notion of prosociality. We deviate from previous studies by also including indirect negative reciprocity, which reflects altruistic punishment and is positively correlated with our measure of altruism ($\rho = 0.257$, see Appendix Table 2.A.1).

We further collected information on demographic characteristics, education, income, political attitudes, beliefs and attitudes towards the COVID-19 pandemic, news consumption, conspiracy mentality, and Big Five personality factors. We construct the Big Five personality traits of openness, conscientiousness, neuroticism, agreeableness, and extraversion using the 15-item BFI-S scale by Gerlitz and Schupp (2005). See Appendix 2.B for a detailed description of all survey questions and variables.

### 2.3.2 Regional-Level Aggregation

For regional-level analyses, we aggregate our survey measures at the administrative NUTS-2 region level in Germany (38 regions; visit https://ec.europa.eu/eurostat/web/nuts/background for information on the NUTS classification system) by calculating the average of all respondents who currently live in that region. The sample size per region ranges from 46 to 427 (mean 154, median 124). We use sampling weights from a raking procedure (Battaglia, Hoaglin, and Frankel, 2009) to improve regional representativeness by age and gender (age above/below 40

$\times$ gender) as well as the share of adults with a college degree. To validate the regional representativeness of our sample, we compare vote shares of the main political parties in the 2019 election with the implied vote shares in our survey based on self-reported party preferences (Appendix, Table 2.A.7). The regional correlations are extremely high — $\rho$ between 0.76 and 0.86 — for all parties except for the FDP, the German liberal party ($\rho = 0.29$).

We further obtain information on the official daily number of confirmed COVID-19 cases and deaths at the county-level (NUTS-3 region) reported by the Robert-Koch-Institut (RKI), the federal government agency and research institute responsible for disease control and prevention in Germany. We use data obtained from infas360 to construct a local policy stringency index by summing up a total of 23 indicator variables for whether local mandates in a certain category (e.g. curfew, school closure) were in place. We normalize this index to range between 0 (no restriction) and 100 (full restriction). Finally, we collect a host of demographic information and socio-economic indicators for each county in Germany from the joint database of the statistical offices of the German states. See Appendix 2.C for detailed descriptions of regional-level data.

## 2.4 Individual-Level Prosociality and Public Health Behavior

We begin by establishing a robust positive relationship between prosociality and PHB at the individual level using data from our representative online sample. To do so, we regress the PHB variable on our measures of prosociality, time and risk preferences, and a number of controls, using ordinary least squares (OLS). The statistical model is

$$PHB_{ic} = \alpha + \beta_1 \cdot Prosocial_i + \beta_2 \cdot Patience_i + \beta_3 \cdot RiskT_i + \boldsymbol{\gamma}' \boldsymbol{x}_{ic} + \varepsilon_{ic}\,, \qquad (2.1)$$

where $PHB_{ic}$ is the public health behavior factor for individual $i$ (living in county $c$) and $Prosocial_i$ is his or her level of prosociality. $Patience_i$ and $RiskT_i$ denote her level of patience and risk-taking, respectively, which we include as these are generally correlated with prosociality (Falk et al., 2016) and may also have an influence on individual's willingness to engage in preventive health measures. $\boldsymbol{x}_{ic}$ is a vector of control variables that differ by specifications. Standard errors are always clustered at the county level.

Table 2.1 presents the regression estimates from the baseline specification in equation 2.1 without additional control variables. Column 1 shows that prosociality strongly predicts individual behavior in the pandemic, with a one SD increase in prosociality being associated with a one third SD increase in PHB ($p < 0.001$). Additionally, we find that more patient and less risk-tolerant individuals are also more likely to adhere to social distancing and hygiene measures. These results are consistent with our theoretical predictions from Section 2.2.

**Table 2.1.** Individual-Level Association between Preferences and PHB

| | Public Health Behavior (PHB) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Prosociality | 0.3356*** | 0.3059*** | 0.3071*** | 0.2182*** | 0.1611*** |
| | (0.0162) | (0.0165) | (0.0167) | (0.0173) | (0.0144) |
| Patience | 0.1983*** | 0.1969*** | 0.1921*** | 0.1689*** | 0.0809*** |
| | (0.0150) | (0.0151) | (0.0150) | (0.0149) | (0.0126) |
| Risk-taking | -0.2095*** | -0.1710*** | -0.1725*** | -0.1715*** | -0.0785*** |
| | (0.0141) | (0.0144) | (0.0143) | (0.0138) | (0.0107) |
| Socio-demographic controls | *No* | *Yes* | *Yes* | *Yes* | *Yes* |
| NUTS-2 region FEs | *No* | *No* | *Yes* | *Yes* | *Yes* |
| Big 5 personality traits | *No* | *No* | *No* | *Yes* | *Yes* |
| COVID-19 perceptions | *No* | *No* | *No* | *No* | *Yes* |
| Observations | 5843 | 5660 | 5660 | 5660 | 5660 |
| Clusters (counties) | 397 | 396 | 396 | 396 | 396 |
| $R^2$ | 0.209 | 0.234 | 0.242 | 0.298 | 0.495 |

*Notes:* In the interest of brevity, we report only the coefficients on economic preference variables here; Appendix Table 2.A.2 reports estimates on other variables included in each specification. Socio-demographic controls include age and age-squared, gender, education, income, employment status, household size, number of children, and an indicator for having children below age 16. COVID-19 perceptions include general attitudes towards the pandemic, infection experiences, and worrying about oneself, family members, and others being infected. Standard errors (in parentheses) are clustered at the county level. See Appendix Tables 2.A.3 and 2.A.4 for detailed results using individual elements of prosociality or PHB. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

People who are more prosocial also tend to differ with regard to other characteristics that may be associated with differential costs and benefits of adhering to recommended PHBs. For example, infection risk and disease severity vary with demographic factors such as age or gender, whereas economic factors such as occupation, income, or household situation could determine the costs of complying with certain preventive measures. Regional differences in current and past infection rates could further influence individual behavior, e.g., if regions hit more severely have stricter policy measures in place, or have developed stricter norms in enforcing such measures. In general, all these factors tend to be correlated with prosociality and could thus act as confounders (Falk et al., 2018). However, columns 2 and 3 of Table 2.1 show that the estimated coefficient for prosociality remains stable and highly statistically significant when controlling for demographic and socio-economic characteristics as well as region fixed effects.

Apart from economic preferences, certain psychological personality traits such as agreeableness and openness from the Big Five inventory have also been linked with stronger adherence to PH measures in the COVID-19 pandemic (Nikolov et al., 2020; Zettler et al., 2022) and are also correlated with prosociality to some degree (see e.g. Appendix Table 2.A.6). However, as the estimates in column 4 of Table 2.1 show, differences in Big Five personality traits do not drive the association between prosociality and PHB. This squares with the general observation that personality traits and economic preferences seem to be partially distinct concepts (Becker et al., 2012; Jagelka, 2020), and both retain explanatory value for individual behavior in the pandemic (see Appendix Table 2.A.2).

Finally, we also investigate to which degree the role of prosociality can be explained by individuals' perceptions and attitudes regarding the COVID-19 pandemic (Table 2.1 column 5). However, even controlling for these factors leaves a strong association between prosociality and PHB intact.

## 2.5 Regional-Level Prosociality and Collective Health Outcomes

In the next step, we examine how regional variation in prosociality across Germany relates to public health outcomes during the COVID-19 pandemic. For this purpose, we construct regional averages of our prosociality and PHB measures by aggregating individual survey responses at NUTS-2 level ("Regierungsbezirk") as described in Section 2.3.

### 2.5.1 Descriptive Overview

We document substantial variation in our measure of prosociality within Germany, as illustrated by the map in Figure 2.2a. Average prosociality ranges from $-0.37$ to 0.42 across NUTS-2 regions, thus spanning about 80% of an individual-level standard deviation. These regional differences are statistically significant ($p < 0.05$) and explain about 50% additional variation in individual-level prosociality compared to other socio-demographic variables alone (Appendix Table 2.A.8). Moreover, regional prosociality patterns are related to commonly used proxies for social (or civic) capital: higher average prosociality is associated with higher voter turnout in the 2019 EU election ($\varrho = 0.3098$, $p = 0.0169$) and larger density of civic associations in 2008 ($\varrho = 0.1394$, $p = 0.0657$), see Appendix Table 2.A.9. Thus, our measure seems to capture stable and meaningful variation.

Figure 2.2b shows that average prosociality is closely linked with average PHB in the pandemic at the regional level. In fact, the regional-level correlation ($\varrho = 0.5795$, $p < 0.001$) is substantially stronger than what would have been predicted solely based on the unconditional individual-level correlation ($\varrho = 0.3503$, $p < 0.001$), suggesting that prosocial individuals may also raise general health compliance indirectly through social influence and normative channels.

Figure 2.2c plots the evolution of COVID-19 cases per 100,000 population in Germany over the course of the pandemic, split by regions with above-median and below-median prosociality. Incidence rates in high-prosociality regions dropped

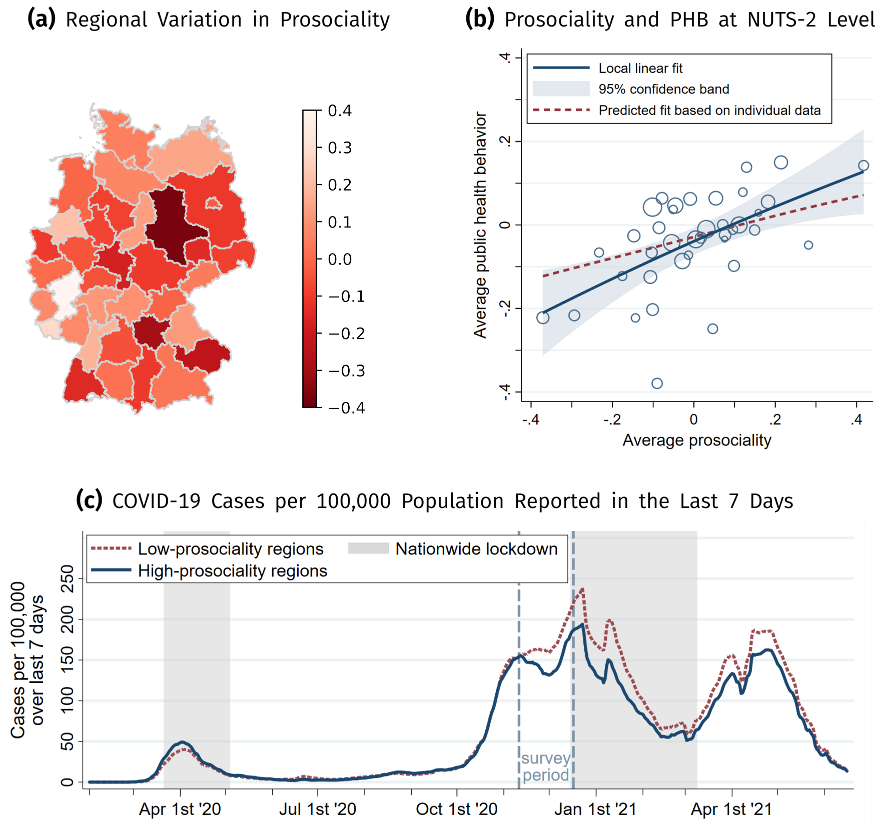



**Figure 2.2.** Prosociality, Public Health Behavior, and COVID-19 Incidence Rates

*Notes:* Panel (a): Map of the 38 NUTS-2 regions in Germany, with color intensity indicating average level of prosociality based on our survey measures. The unit is individual-level SDs. Panel (b): Relation between average prosociality and average PHB on NUTS-2 level, both expressed in terms of individual-level SDs. The solid fitted line is constructed from an unweighted local linear regression (Gaussian kernel, bandwidth = 0.3) of average PHB on average prosociality at NUTS-2 region level (N = 38). The dashed line shows the association between average prosociality and the average fitted values from an individual-level regression of PHB on prosociality and prosociality-squared. Bubbles indicate NUTS-2 regions and are proportional to population size. Panel (c): Official number of COVID-19 cases reported by RKI between Feb 1, 2020, and Jun 15, 2021. Grey shaded areas indicate time periods of strict nationwide lockdowns in Germany (as of March 8, 2021, restrictions were tied to the regional incidence rate, although the lockdown formally remained in place).

persistently below those in low-prosociality regions starting from around Nov 2020, in the period of the so-called “lockdown light”, which was in place at the beginning of the second wave in Germany and had the goal of reducing social contacts while avoiding a complete economic standstill. At the height of the second wave, high-prosociality regions experienced around 15-25% lower incidence rates, and 20-30% fewer COVID-19 deaths (see Appendix Figure 2.A.2, which also shows differential mobility patterns during the second wave). These descriptive observations hint at a meaningful role of prosociality in determining how well a region can slow the spread of the virus and protect vulnerable groups. However, regions with different levels of prosociality also differ by other characteristics such as population density and socio-economic factors. Therefore, we will now move on to our formal statistical analyses.

### 2.5.2 Association between Prosociality and COVID-19 Incidence Rates

Our main outcome variable is the weekly COVID-19 incidence rate, i.e. the confirmed number of new cases per 100,000 population within 7 days, as reported by the RKI for each county in Germany. We additionally take the logarithm of the incidence rate to capture the exponential nature of infectious disease dynamics. Results for COVID-19 deaths are reported in the Appendix and in general very similar. As a first step in examining the relation between regional incidence rates and prosociality, we use OLS to estimate the following statistical model:

$$\log(cases_{crt}) = \alpha_t + \beta_1 \cdot \overline{Prosocial}_r + \beta_2 \cdot \overline{Patience}_r + \beta_3 \cdot \overline{RiskT}_r + \boldsymbol{\gamma}_t' \boldsymbol{x}_c + \varepsilon_{crt}\,, \tag{2.2}$$

where $log(cases_{crt})$ is the log COVID-19 incidence rate in county $c$ (NUTS-3 level) and week $t$. Our main regressor of interest is $\overline{Prosocial}_r$, which is the average prosociality in NUTS-2 region $r$. $\overline{Patience}_r$ and $\overline{RiskT}_i$ denote the average level of patience and risk-taking, respectively. For ease of interpretation, we standardize these three preference measures to mean 0 and standard deviation 1 across regions. $\boldsymbol{x}_c$ is a vector of pre-pandemic county characteristics, which we interact with week dummies to allow the coefficient vector $\boldsymbol{\gamma}_t$ to change over time. To account for the highly dynamic nature of the pandemic, all specifications include week fixed effects $\alpha_t$. We focus our analysis on the two-month period from Nov 16 to Jan 17, around the peak of the second wave in Germany, because this is when our survey measures are most applicable. Note that we include an additional month of data from the end our survey onwards, as the effects of changes in behavior or policies will only manifest themselves with a delay, which is exacerbated by reporting lags by local health authorities during Christmas and New Year. Statistical inference is robust to clustering at the NUTS-2 region level. Due to the relatively low number of clusters (38), we report confidence intervals based on a wild cluster bootstrap-t procedure (Cameron, Gelbach, and Miller, 2008; Roodman et al., 2019).

Table 2.2 presents the baseline results, which indicate a robust association between regional incidence rates and prosociality. The estimated coefficient in column 1 shows that, without controlling for any other county characteristics, a one SD higher prosociality is associated with a 13% lower weekly incidence rate in the time period we study. This effect is both statistically significant ($p < 0.001$) and quantitatively sizeable, corresponding to about 8% of the region-week SD in incidence rates(see Appendix table 2.A.16). This association remains robust to including regional-level time and risk preferences as regressors (column 2), although

**Table 2.2.** Weekly Incidence at the Time of the Survey

| | $y_{c,t} = \log(cases_{c,t})$ in county $c$ and week $t$ | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Prosociality | -0.1391 *** | -0.1270 * | -0.1241 ** | -0.1189 ** | 0.0183 |
| | [-0.283, -0.061] | [-0.303, 0.010] | [-0.296, -0.021] | [-0.246, -0.033] | [-0.088, 0.106] |
| Patience | – | -0.0286 | 0.0024 | -0.0054 | 0.0602 |
| | | [-0.211, 0.133] | [-0.117, 0.181] | [-0.111, 0.129] | [-0.019, 0.188] |
| Risk-taking | – | 0.0106 | -0.0377 | -0.0454 | -0.0814 * |
| | | [-0.107, 0.126] | [-0.154, 0.092] | [-0.137, 0.072] | [-0.149, 0.005] |
| Public health behavior | – | – | – | – | -0.2996 *** |
| | | | | | [-0.443, -0.158] |
| Wave 1 severity | *No* | *No* | *No* | *Yes* | *Yes* |
| County controls × Week | *No* | *No* | *Yes* | *Yes* | *Yes* |
| Week fixed effects | *Yes* | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 3609 | 3609 | 3609 | 3609 | 3609 |
| Spatial units (counties) | 401 | 401 | 401 | 401 | 401 |
| Clusters (NUTS-2 regions) | 38 | 38 | 38 | 38 | 38 |
| $R^2$ | 0.116 | 0.118 | 0.357 | 0.415 | 0.481 |

*Notes:* Bootstrapped 95%-confidence-intervals in brackets (clustered at NUTS-2 level), obtained using wild bootstrapping with Rademacher-weights and 9,999 simulations. The outcome variable is the log weekly incidence rate by county, ranging from Nov 16, 2020, until Jan 17, 2021 (9 weeks). County controls include 18 variables: log population density, log GDP per capita, log average income per capita, share of college graduates, employment share, share of non-German residents, share of workers in the service sector, share of population below age 18, share of population age 65 or above, and border county dummies for each neighboring country of Germany. Controls for wave 1 severity include the log of aggregate case numbers, its square, and case fatality rate in the time period from the first confirmed infection until May 17th, 2020. See Appendix Table 2.A.10 for results with the individual elements of prosociality. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

its precision decreases due to the covariates being correlated with each other. The estimated coefficients for patience and risk-taking are small and insignificant.

Importantly, we verify whether the association between prosociality and COVID-19 incidence rates is robust to controlling for other demographic and socio-economic county characteristics that could influence the regional spread of the virus. In column 3, we therefore add pre-pandemic county characteristics ($\boldsymbol{x}_c$) and allow their effect to vary by week. The vector of county controls consists of log population density, log GDP per capita, log average income per capita, share of college graduates, employment share, share of workers in the service sector, share of non-German residents, share of population below age 18, share of population age 65 or above, and border county dummies for each neighboring country of Germany. Another potential concern is that regional differences in severity of the pandemic experienced during the first wave may have had an impact on the level of prosociality, but simultaneously also on other factors like general attitudes or local government preparedness. To flexibly account for this, we further add control variables for counties' first wave (February-May) infection outcomes in another specification.

After including this rich set of control variables in columns 3-4 of Table 2.2, the explanatory power of the regression increases drastically by a factor of more than three. Crucially, the coefficient for prosociality remains nearly unchanged, with a one SD increase being associated with $11-12\%$ lower weekly incidence rates ($p < 0.05$).

Why is the incidence rate lower in regions with higher prosociality? Our theoretical considerations suggest that more prosocial individuals should be more willing to comply with recommended or mandatory social distancing and hygiene measures, which is confirmed empirically by our individual-level results. The models discussed in Section 2.2 would then predict that stricter engagement in preventive health behaviors leads to a lower contact and transmission rate, and thus eventually to a lower COVID-19 incidence rate in high-prosociality regions. To test this mediating role of behavior, we include our measure of average PHB as additional regressor in column 5 of Table 2.2 (Baron and Kenny, 1986). Upon doing so, the coefficient size for prosociality is reduced by 85% to almost zero, whereas we observe a remarkably strong relation between self-reported PHB and incidence rates: a one SD increase in PHB is associated with a 26% decrease in the weekly number of cases per 100,000 population. This is consistent with the hypothesis that the effect of prosociality is mediated by differences in PHB across regions. Interestingly, risk-taking has a weakly significant negative effect conditional on PHB, which could potentially be explained with a higher willingness to experiment with new strategies or to adopt new technologies.

Although we have controlled for a host of demographic and socio-economic county characteristics, there could still be other, unobserved factors that lead to generally lower levels of infections in a county, while also being positively correlated with prosociality and PHB. To circumvent this issue, we test whether regions with higher prosociality also exhibit lower growth rates of new cases, as this partials out any time-invariant differences across counties that can affect absolute levels of infection rates in the pandemic. We approximate growth rates by the weekly change in log incidence rates $\Delta \log(cases_{crt}) = \log(cases_{c,t}) - \log(cases_{c,t-1})$ in county $c$ and week $t$ and estimate the following statistical model:

$$\begin{aligned}\Delta \log(cases_{crt}) \;=\; & \alpha_t + \beta_1 \cdot \overline{Prosocial}_r + \beta_2 \cdot \overline{Patience}_r + \beta_3 \cdot \overline{RiskTak}_r \\ & + \boldsymbol{\gamma}_t' \boldsymbol{x}_c + \boldsymbol{\delta}' \boldsymbol{w}_c + \varepsilon_{crt}\,,\end{aligned} \tag{2.3}$$

where everything is defined as in equation 2.2. We include the full set of previously used control variables in all specifications, including the vector of controls for wave 1 severity $\boldsymbol{w}_c$.

Although high- and low-prosociality regions start from roughly similar levels of incidence at the beginning of the second wave (see Figure 2.2c), differences in

**Table 2.3.** Weekly Growth Rate of Confirmed Cases at the Time of the Survey

| | $y_{c,t} = \log(cases_{c,t}) - \log(cases_{c,t-1})$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Prosociality | -0.0091 ** | -0.0097 | -0.0218 *** | -0.0072 |
| | [-0.018, -0.001] | [-0.022, 0.002] | [-0.037, -0.011] | [-0.025, 0.008] |
| Patience | -0.0012 | -0.0015 | -0.0012 | 0.0062 |
| | [-0.014, 0.007] | [-0.015, 0.009] | [-0.011, 0.014] | [-0.008, 0.026] |
| Risk-taking | 0.0002 | 0.0003 | -0.0044 | -0.0092 |
| | [-0.012, 0.013] | [-0.012, 0.012] | [-0.016, 0.010] | [-0.026, 0.007] |
| Public health behavior | – | 0.0012 | – | -0.0340 ** |
| | | [-0.021, 0.022] | | [-0.066, -0.006] |
| $\log(cases_{c,t-2})$ | – | – | -0.1081 *** | -0.1209 *** |
| | | | [-0.126, -0.093] | [-0.146, -0.096] |
| Policy stringency$_{c,t-2}$ | – | – | -0.2403 | -0.2050 |
| | | | [-0.857, 0.289] | [-0.765, 0.228] |
| Wave 1 severity | *Yes* | *Yes* | *Yes* | *Yes* |
| County controls × Week | *Yes* | *Yes* | *Yes* | *Yes* |
| Week fixed effects | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 3609 | 3609 | 3609 | 3609 |
| Spatial units (counties) | 401 | 401 | 401 | 401 |
| Clusters (NUTS-2 regions) | 38 | 38 | 38 | 38 |
| $R^2$ | 0.293 | 0.293 | 0.315 | 0.317 |

*Notes:* Bootstrapped 95%-confidence-intervals in brackets (clustered at NUTS-2 level), obtained using wild bootstrapping with Rademacher-weights and 9,999 simulations. The outcome variable is the change in log weekly incidence rate in a county, ranging from Nov 16th, 2020 until Jan 17th, 2021 (9 weeks). All control variables are defined as in Table 2.2. See Appendix Table 2.A.11 for results with the individual elements of prosociality. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

the growth rate would gradually drive incidence levels apart over time, eventually resulting in large cumulative differences. Indeed, our baseline specification in Table 2.3 shows that, in the time period we study, the growth rate of new cases was about 1%p lower in regions with a one SD higher prosociality ($p < 0.05$). We find no evidence for mediation through PHB in column 2 yet.

However, the estimated effects of prosociality and social distancing might be attenuated due to dynamic interactions between incidence rates, behavior, and policy responses that push towards regional convergence. For one, the share of susceptibles in the population is naturally higher in regions with fewer past infections, although this effect may have been negligible at that stage of the pandemic. Moreover, SIR models with endogenous behavior predict that in regions with lower incidence rates, people may endogenously reengage in more social contacts in response to reduced infection risks. Local governments could also feel encouraged to partially lift curtailment measures. Thus, more prosocial regions could become the victims of their own success. For this reason, we further add the 2-week lagged incidence rate $\log(cases_{c,t-2})$ as well as a 2-week lagged local policy stringency index (see Section 2.3.2) as covariates in equation 2.3. After including these lagged variables, the coefficient size for prosociality more than doubles, implying a 2%*p* lower weekly growth rate per SD increase ($p < 0.01$) — this corresponds to about 3% of a region-week SD in incidence growth rates (see Appendix 2.A.17). This is a sizeable effect given that small differences in growth rates accumulate to large absolute differences over time. In column 4, prosociality becomes insignificant after adding average PHB, further supporting the hypothesis that better compliance with social distancing and hygiene measures mediates the effect of higher prosociality on collective health outcomes during the pandemic.

Finally, we check whether our results are influenced by comparisons between West Germany and East Germany, as previous studies document that historical institutional differences between these two regions before the German reunification still have a persistent effect on preferences, norms, and outcomes (Torgler, 2002; Alesina and Fuchs-Schündeln, 2007; Brosig-Koch et al., 2011; Becker, Mergele, and Woessmann, 2020). Therefore, we rerun our analyses adding an East-Germany dummy as control variable, and further interacting it with our measure of average prosociality (Appendix Tables 2.A.13-2.A.15). The results show that the estimated coefficients for prosociality remain robust, and that there is no evidence for a differential association between higher prosociality and lower COVID-19 incidence rates in East and West Germany, although the low number of regional units in the East precludes any conclusive statement.

## 2.6 Discussion

How well a group of individuals succeeds in achieving desirable collective outcomes in the face of social dilemma depends, amongst other things, on how willingly individual members engage in actions that incur personal costs but that benefit the group as a whole. We have provided suggestive evidence that, in the context of the COVID-19 pandemic, more prosocial individuals are significantly more willing to engage in public health behaviors (e.g. physical distancing and mask-wearing) aimed at slowing the spread of the virus. We further presented evidence that, in turn, regions in Germany with higher average prosociality in the population also tend to experience a lower incidence of COVID-19 cases and deaths. The estimated (conditional) correlations are quantitatively sizeable: a 1 SD higher average prosociality in a region is associated with around 11% lower COVID-19 incidence rates and 2%*p* lower incidence growth rates.

### 2.6.1 Role of the Study Context

The interpretation of our results needs to take into account the broader context in which our study is embedded, as the role of prosociality may be moderated, among others, by the stage of the pandemic, the regional severity of the outbreak, and the stringency of government-mandated restrictions and policy measures. Our survey was conducted in the late fall of 2020, before the peak of the second wave in Germany, during the so-called lockdown light. In contrast, most related studies examining determinants of PHB were conducted in the first wave of the pandemic, when more fear and uncertainty was revolving around the disease and the spread of the virus (Harper et al., 2020). Thus, we confirm previous results on the importance of prosociality (Campos-Mercade et al., 2021; Müller and Rau, 2021) also for later stages of the pandemic, when people had become more accustomed to and more weary of the situation (Petherick et al., 2021). In Table 2.A.18 of the Appendix, we compare predictors of regional incidence rates in the first and the second COVID-19 wave in Germany. We observe that the same set of demographic and socio-economic county characteristics (e.g. population density, employment share) has much higher explanatory value in the first wave ($R^2 = 0.497$) than in the second wave ($R^2 = 0.265$), possibly because behavioral responses in the population were more homogeneous early on in the pandemic.

The quickly rising case numbers at the time period of our survey might have further driven attitudes and behavioral responses apart for people in different regions and with different individual characteristics, as protecting those vulnerable to the disease becomes especially relevant when the risk of infection and transmission is high. In contrast, private gatherings may not be considered irresponsible acts of selfishness in periods of low incidence such as the summer of 2020 in Germany. Another potentially amplifying factor for the role of prosociality in our context may be that the lockdown light in Germany left plenty of wiggle room in the extent of social distancing behavior within the limits of what was allowed, thereby putting considerable weight on voluntary reduction of social contacts. Although voluntary adaptions and government-mandated restrictions can be partly substitutable (Alfaro et al., 2021b), prosociality may affect health behaviors and outcomes even under more stringent lockdown regimes, as perfect monitoring and enforcement of compliance are infeasible, and drastic government measures can also influence public perceptions of severity and social norms (Casoria, Galeotti, and Villeval, 2021; Galbiati et al., 2021).

### 2.6.2 Potential Endogeneity Concerns

Finally, a natural question in our context is to which extent the conditional correlations we find in our empirical analyses can be interpreted as causal. There are several potential concerns against such a causal interpretation. First, our sample may not be regionally representative due to self-selection into completing the survey. While such selection effects are hard to rule out, they could only explain our results if systematically more prosocial individuals respond to our survey in regions with lower incidence rates, which seems implausible. Second, one might worry that our measures of prosociality and economic preferences are themselves affected by the COVID-19 pandemic (Bauer et al., 2016; Branas-Garza et al., 2020; Cappelen et al., 2021; Frondel, Osberghaus, and Sommer, 2021; Shachat, Walker, and Wei, 2021). If any influence on individuals' survey responses reflects true changes in preferences and attitudes, our measures remain internally valid for the time period around which we conducted the survey. On the other hand, we might overestimate the role of prosociality if respondents' answers to broadly framed questions overreflected their behavior during the pandemic, e.g. due to availability bias (Tversky and Kahneman, 1973). We cannot directly investigate this issue with our cross-sectional survey data, but note that regional prosociality in our data correlates with pre-pandemic outcomes such as election turnout, and that our results are robust to controlling for first-wave severity of the pandemic. Moreover, Campos-Mercade et al. (2021) provide evidence that individual health behavior during the pandemic is predicted by prosociality measured before the COVID-19 outbreak, which is consistent with the notion that individual's (social) preferences are fairly stable in general (Volk, Thöni, and Ruigrok, 2012; Carlsson, Johansson-Stenman, and Nam, 2014). A third concern is reverse causality, because regional incidence rates may also influence PHB and its relation to prosociality. However, this would presumably lead to an underestimation of the true effect since lower incidence rates allow residents and policymakers to become more lenient in their responses. Consistent with this convergence effect, we have shown in Table 2.3 that the estimated association between average prosociality and weekly incidence growth rate doubles in magnitude when controlling for lagged incidence levels.

The fourth and arguably most important concern is omitted variable bias. At the individual level, it seems unlikely that the relation between prosociality and PHB is entirely driven by some unobserved factor, as we control for a host of demographic and socio-economic characteristics, and further confirm robustness to including personality factors and political attitudes as regressors. At the regional level, we control for a variety of relevant county characteristics. However, it is difficult to rule out all potentially confounding factors, e.g., the stringency of local implementation and enforcement of containment measures, contact tracing efficiency, etc., which may themselves be a function of prosociality in the population. Most notably, the distribution of (pro-)social preferences, values, norms, and beliefs is inherently endogenous to social, cultural, political, and institutional factors. Because these factors are imperfectly observable and the underlying causal relationships highly complex and interdependent, our empirical investigation must inevitably remain correlational.

### 2.6.3 Concluding Remarks

Our paper is inspired by several previous studies that measure individual and geographical variation of (pro-)social behavior and preferences in order to advance our understanding of how collective societal outcomes may be shaped by the prevalent values, norms, and preferences in the population, and vice versa, how individual dispositions may vary due to ecological, cultural, or socio-economic factors (Henrich et al., 2006; Nettle, Colléony, and Cockerill, 2011; Falk et al., 2018; Cohn et al., 2019; Barsbai, Lukas, and Pondorfer, 2021; Caicedo, Dohmen, and Pondorfer, 2021). Recent experimental evidence further highlights the malleability of prosociality by documenting the importance of socialization and role models (Kosse et al., 2020). Cultivating prosocial values and norms within a society may strengthen its capacity to face challenges such as pandemics or global warming that require widespread cooperation and collective action.

## Appendix 2.A Additional Results and Robustness Checks

### 2.A.1 Supplementary Figures

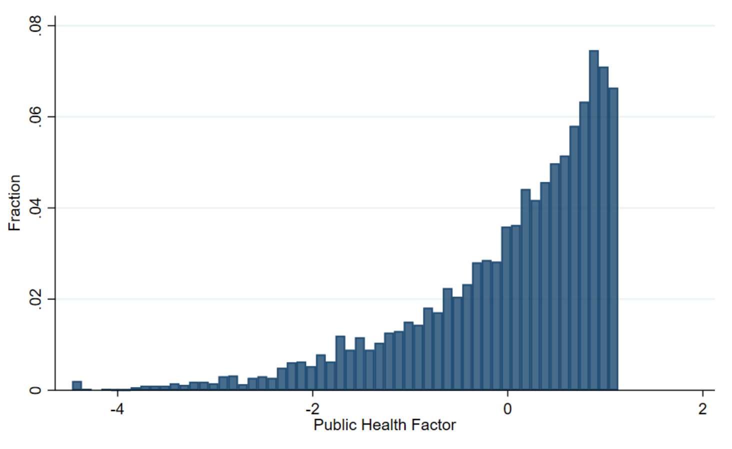
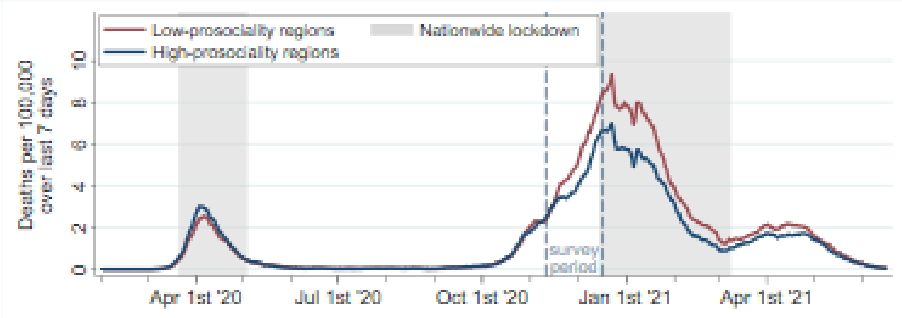



**Figure 2.A.1.** Histogram of PHB Values

*Notes:* Histogram of public health behavior, using width of 0.1.

**(a)** COVID-19 Deaths per 100.000 Population in 7 Days (by Date of Infection)




**(b)** Change in Mobility [%] Compared to 2019 (7-Day Moving Average)

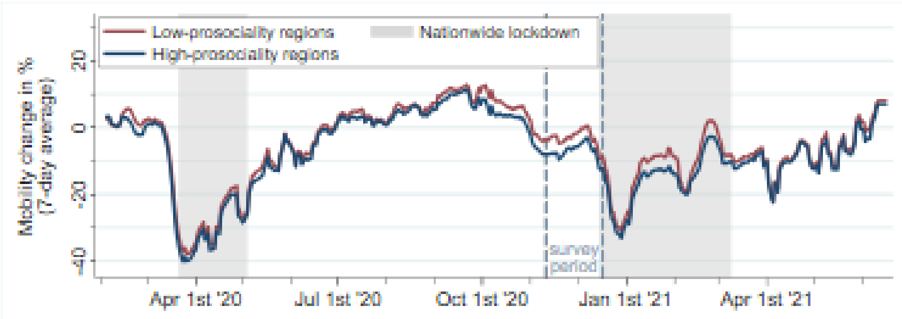
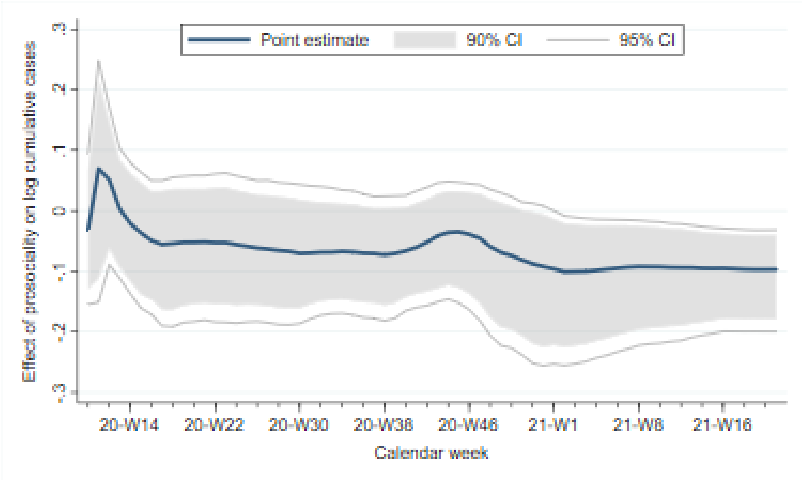



**Figure 2.A.2.** The COVID-19 Pandemic in Germany

*Notes:* The time labels in Figure 2.A.2a refer to the day the coronavirus infection of the deceased person was first reported to the RKI, not the day of death. Grey shaded areas indicate time periods of strict nationwide lockdowns in Germany (as of March 8, 2021, restrictions were tied to the regional incidence rate, although the lockdown formally remained in place).

**(a)** Log Cumulative Cases per Population




**(b)** Log Cumulative Deaths per Population

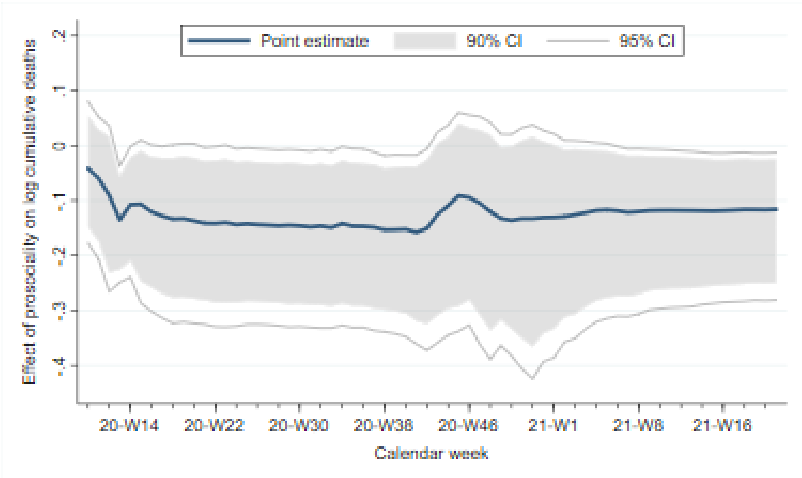



**Figure 2.A.3.** Estimated Effect of Prosociality on Cumulative Cases and Deaths

*Notes:* Confidence-intervals are obtained using the wild bootstrap (9,999 simulations) with clustering on NUTS-2 region level and Rademacher-weights. The time labels in Panel (b) refer to the day the coronavirus infection of the deceased person was first reported to the RKI, not the day of death.

### 2.A.2 Supplementary Tables

**Table 2.A.1.** Correlation Matrix of Prosociality Components

| | Altruism | Positive reciprocity | Trust | Indirect neg. reciprocity |
|---|---|---|---|---|
| Altruism | 1 | | | |
| Positive reciprocity | 0.3344 | 1 | | |
| Trust | 0.2591 | 0.1503 | 1 | |
| Indirect neg. reciprocity | 0.2574 | 0.1705 | 0.1488 | 1 |
| Observations | | 5949 | | |

*Notes:* Pearson correlation coefficients of altruism, positive reciprocity, trust, and indirect (negative) reciprocity across individual survey respondents.

**Table 2.A.2.** Individual-Level Association between Preferences and PHB

| | Public Health Behavior (PHB) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Prosociality | 0.3356*** | 0.3059*** | 0.3115*** | 0.2216*** | 0.1625*** |
| | (0.0162) | (0.0165) | (0.0168) | (0.0173) | (0.0148) |
| Patience | 0.1983*** | 0.1969*** | 0.1858*** | 0.1633*** | 0.0777*** |
| | (0.0150) | (0.0151) | (0.0155) | (0.0155) | (0.0131) |
| Risk-taking | -0.2095*** | -0.1710*** | -0.1722*** | -0.1683*** | -0.0790*** |
| | (0.0141) | (0.0144) | (0.0148) | (0.0141) | (0.0110) |
| Negative reciprocity (Direct) | -0.1231*** | -0.1078*** | -0.1075*** | -0.0662*** | -0.0184 |
| | (0.0141) | (0.0145) | (0.0151) | (0.0156) | (0.0127) |
| Female | | 0.1546*** | 0.1542*** | 0.0895*** | 0.0800*** |
| | | (0.0267) | (0.0269) | (0.0266) | (0.0225) |
| Age | | 0.0146* | 0.0141* | 0.0084 | 0.0127* |
| | | (0.0083) | (0.0085) | (0.0081) | (0.0070) |
| Age$^2$ | | -0.0000 | -0.0000 | 0.0000 | -0.0001 |
| | | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Big 5: Openness | | | | 0.0578*** | 0.0423*** |
| | | | | (0.0135) | (0.0116) |
| Big 5: Conscientiousness | | | | 0.1596*** | 0.1577*** |
| | | | | (0.0157) | (0.0129) |
| Big 5: Extraversion | | | | 0.0192 | 0.0070 |
| | | | | (0.0135) | (0.0114) |
| Big 5: Agreeableness | | | | 0.1186*** | 0.1055*** |
| | | | | (0.0162) | (0.0137) |
| Big 5: Neuroticism | | | | 0.0418*** | -0.0121 |
| | | | | (0.0136) | (0.0116) |
| Affected by pandemic | | | | | 0.0252** |
| | | | | | (0.0121) |
| Take pandemic seriously | | | | | 0.2974*** |
| | | | | | (0.0157) |
| Worry: Self | | | | | 0.0211** |
| | | | | | (0.0084) |
| Worry: Family & Friends | | | | | 0.0761*** |
| | | | | | (0.0107) |
| Worry: Others | | | | | 0.0557*** |
| | | | | | (0.0101) |
| Socio-demographic factors | *No* | *Yes* | *Yes* | *Yes* | *Yes* |
| County FEs | *No* | *No* | *Yes* | *Yes* | *Yes* |
| Big Five | *No* | *No* | *No* | *Yes* | *Yes* |
| COVID-19 Perceptions | *No* | *No* | *No* | *No* | *Yes* |
| Observations | 5843 | 5660 | 5653 | 5653 | 5653 |
| Clusters | 397 | 396 | 389 | 389 | 389 |
| $R^2$ | 0.209 | 0.234 | 0.293 | 0.345 | 0.529 |

*Notes:* This table estimates the same specifications as in Table 2.1, but reports additional estimates that might be of interest to the reader. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Table 2.A.3.** Individual-Level Association between Individual Preferences and PHB

| | Public Health Behavior (PHB) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Altruism | 0.1547*** | 0.1545*** | 0.1492*** | 0.1141*** | 0.0598*** |
| | (0.0146) | (0.0148) | (0.0149) | (0.0150) | (0.0138) |
| Positive reciprocity | 0.2383*** | 0.2048*** | 0.2125*** | 0.1342*** | 0.1361*** |
| | (0.0170) | (0.0170) | (0.0174) | (0.0171) | (0.0149) |
| Negative reciprocity (Indirect) | 0.0218 | 0.0253* | 0.0258* | 0.0150 | 0.0011 |
| | (0.0148) | (0.0152) | (0.0152) | (0.0147) | (0.0133) |
| Trust | 0.0708*** | 0.0582*** | 0.0663*** | 0.0609*** | 0.0512*** |
| | (0.0130) | (0.0131) | (0.0131) | (0.0132) | (0.0113) |
| Patience | 0.1807*** | 0.1813*** | 0.1690*** | 0.1561*** | 0.0675*** |
| | (0.0156) | (0.0157) | (0.0159) | (0.0157) | (0.0130) |
| Risk-taking | -0.1949*** | -0.1632*** | -0.1648*** | -0.1662*** | -0.0772*** |
| | (0.0140) | (0.0144) | (0.0148) | (0.0142) | (0.0110) |
| Negative reciprocity (Direct) | -0.0741*** | -0.0665*** | -0.0654*** | -0.0341* | 0.0086 |
| | (0.0179) | (0.0179) | (0.0184) | (0.0187) | (0.0160) |
| Socio-demographic factors | — | Yes | Yes | Yes | Yes |
| County FEs | — | — | Yes | Yes | Yes |
| Big Five | — | — | — | Yes | Yes |
| COVID-19 Perceptions | — | — | — | — | Yes |
| Observations | 5843 | 5660 | 5653 | 5653 | 5653 |
| $R^2$ | 0.223 | 0.243 | 0.302 | 0.348 | 0.533 |
| Clusters | 397 | 396 | 389 | 389 | 389 |

*Notes:* This table estimates the same specifications as in Table 2.1, but with the individual social preferences of altruism, trust, positive reciprocity and indirect negative reciprocity as independent variables instead of prosociality. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Table 2.A.4.** Individual-Level Association between Preferences and Individual PHB Survey Items

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| Prosociality | 0.2928*** | 0.3807*** | 0.2893*** | 0.3562*** | 0.3749*** | 0.4040*** | 0.3317*** | 0.3647*** | 0.2626*** | 0.3738*** |
| | (0.0224) | (0.0268) | (0.0236) | (0.0240) | (0.0366) | (0.0301) | (0.0285) | (0.0241) | (0.0295) | (0.0244) |
| Patience | 0.2112*** | 0.2257*** | 0.2339*** | 0.1562*** | 0.2061*** | 0.1473*** | 0.2299*** | 0.1531*** | 0.2824*** | 0.1608*** |
| | (0.0223) | (0.0247) | (0.0254) | (0.0217) | (0.0353) | (0.0275) | (0.0279) | (0.0245) | (0.0277) | (0.0202) |
| Risk-taking | -0.1870*** | -0.2419*** | -0.1546*** | -0.1429*** | -0.1728*** | -0.1473*** | -0.1993*** | -0.1189*** | -0.3499*** | -0.1206*** |
| | (0.0208) | (0.0236) | (0.0252) | (0.0217) | (0.0350) | (0.0244) | (0.0244) | (0.0243) | (0.0248) | (0.0179) |
| Negative reciprocity (Direct) | -0.0835*** | -0.1463*** | -0.0436* | -0.1407*** | -0.0766** | -0.1922*** | -0.1085*** | -0.1322*** | -0.0569** | -0.1762*** |
| | (0.0235) | (0.0258) | (0.0260) | (0.0224) | (0.0389) | (0.0269) | (0.0228) | (0.0248) | (0.0278) | (0.0212) |
| Observations | 5653 | 5653 | 5653 | 5653 | 5653 | 5653 | 5653 | 5653 | 5653 | 5653 |
| $R^2$ | 0.206 | 0.215 | 0.199 | 0.204 | 0.184 | 0.179 | 0.198 | 0.197 | 0.198 | 0.233 |
| Clusters | 389 | 389 | 389 | 389 | 389 | 389 | 389 | 389 | 389 | 389 |

*Notes:* This table estimates the specification (3) of Table 2.1, but using individual survey items of the PHB index as dependent variables. The columns are defined as follows: 1) Social distancing of 1.5 meters 2) Self-quarantining in the case of risky contact 3) Keeping oneself informed about the pandemic 4) Washing and disinfecting hands 5) Willingness to get vaccinated 6) Sneezing and coughing into elbow 7) Wearing mask 8) Ventilating when indoors 9) Avoiding social contacts 10) Informing others if infected. Each survey item is measured on a 7-point scale, with 1 indicating “Do not agree” and 7 indicating “Agree completely”. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.5.** Economic Preferences, Personality Traits and COVID-19 Perceptions

| | (1) Pandemic serious | (2) Worry: Family & Friends | (3) Worry: Others |
|---|---|---|---|
| Prosociality | 0.1059*** | 0.1875*** | 0.2682*** |
| | (0.0192) | (0.0362) | (0.0329) |
| Patience | 0.1838*** | 0.2446*** | 0.1765*** |
| | (0.0168) | (0.0310) | (0.0313) |
| Risk-taking | -0.1962*** | -0.2275*** | -0.1954*** |
| | (0.0181) | (0.0349) | (0.0292) |
| Negative reciprocity (Direct) | -0.1429*** | -0.0558 | -0.0932*** |
| | (0.0177) | (0.0349) | (0.0332) |
| Big 5: Openness | -0.0016 | 0.0814*** | 0.0981*** |
| | (0.0164) | (0.0309) | (0.0279) |
| Big 5: Conscientiousness | -0.0116 | 0.0311 | -0.0237 |
| | (0.0190) | (0.0303) | (0.0294) |
| Big 5: Extraversion | -0.0204 | 0.1032*** | 0.1500*** |
| | (0.0164) | (0.0325) | (0.0269) |
| Big 5: Agreeableness | 0.0042 | 0.1136*** | 0.0518 |
| | (0.0180) | (0.0304) | (0.0318) |
| Big 5: Neuroticism | -0.0144 | 0.3725*** | 0.3351*** |
| | (0.0148) | (0.0299) | (0.0295) |
| Observations | 5653 | 5653 | 5653 |
| $R^2$ | 0.190 | 0.200 | 0.192 |
| Clusters | 389 | 389 | 389 |

*Notes:* Pandemic serious is a factor comprised of two survey items measuring (on a 5-point scale) how much the respondent disagrees with the statements that the media takes the pandemic too seriously, and that government measures are too strict. Worry: Family & Friends and Worry: Others measure (on a 7-point scale) how much the respondent worries about their family and friends, and others around them, respectively. All specifications include socio-demographic controls and county FEs. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.6.** Correlation Matrix of Prosociality and BFI Personality Traits

| | Prosocial-ity | Agreeable-ness | Conscient-iousness | Extravers-ion | Neurotic-ism | Openness |
|---|---|---|---|---|---|---|
| Prosociality | 1.0000 | | | | | |
| Agreeableness | 0.3070*** | 1.0000 | | | | |
| Conscientiousness | 0.2446*** | 0.4353*** | 1.0000 | | | |
| Extraversion | 0.2451*** | 0.2347*** | 0.3021*** | 1.0000 | | |
| Neuroticism | -0.0314* | -0.0209 | -0.1554*** | -0.2268*** | 1.0000 | |
| Openness | 0.2777*** | 0.2399*** | 0.2638*** | 0.4142*** | -0.0060 | 1.0000 |

*Notes:* Pearson correlation coefficients of prosociality, agreeableness, conscientiousness, extraversion, neuroticism, and openness across individual survey respondents. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.7.** Regional Correlations of Vote Shares for the Major Political Parties

| | *Regional correlation with 2019 election outcome* | | | | | |
|---|---|---|---|---|---|---|
| | CDU/CSU | SPD | Grüne | FDP | Die Linke | AfD |
| Survey vote shares | 0.808*** | 0.854*** | 0.757*** | 0.290* | 0.861*** | 0.784*** |
| 2017 election outcomes | 0.904*** | 0.923*** | 0.844*** | 0.763*** | 0.980*** | 0.970*** |
| Observations | 38 | 38 | 38 | 38 | 38 | 38 |
| Overall 2019 vote share [%] | 22.6 | 15.8 | 20.5 | 5.4 | 5.5 | 11.0 |

*Notes:* The first row shows the Pearson's correlation coefficients of 2019 election vote shares with the implied vote shares from our survey on NUTS-2 region level. For comparison, the second rows shows the correlation of 2019 election outcomes with 2017 election outcomes. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.8.** Variation of Prosociality across NUTS-2 Regions in Germany

| | (1) *No controls* | | (2) *With controls* | | (3) *Only controls* |
|---|---|---|---|---|---|
| | Coefficient | SE | Coefficient | SE | |
| Brandenburg | -0.059 | (0.102) | -0.096 | (0.103) | – |
| Bremen | 0.140 | (0.124) | 0.109 | (0.117) | – |
| Direktionsbezirk Chemnitz | 0.056 | (0.137) | 0.039 | (0.136) | – |
| Direktionsbezirk Dresden | -0.046 | (0.113) | -0.089 | (0.115) | – |
| Direktionsbezirk Leipzig | -0.086 | (0.131) | -0.105 | (0.131) | – |
| Hamburg | 0.095 | (0.120) | 0.079 | (0.120) | – |
| Mecklenburg-Vorpommern | 0.133 | (0.118) | 0.121 | (0.121) | – |
| Reg.-Bez. Arnsberg | -0.013 | (0.093) | -0.011 | (0.093) | – |
| Reg.-Bez. Darmstadt | 0.104 | (0.087) | 0.068 | (0.088) | – |
| Reg.-Bez. Detmold | -0.037 | (0.122) | -0.021 | (0.123) | – |
| Reg.-Bez. Düsseldorf | -0.055 | (0.083) | -0.070 | (0.084) | – |
| Reg.-Bez. Freiburg | -0.089 | (0.106) | -0.087 | (0.107) | – |
| Reg.-Bez. Gießen | -0.017 | (0.178) | -0.097 | (0.193) | – |
| Reg.-Bez. Karlsruhe | 0.157 | (0.100) | 0.137 | (0.101) | – |
| Reg.-Bez. Kassel | -0.110 | (0.147) | -0.171 | (0.147) | – |
| Reg.-Bez. Köln | 0.025 | (0.087) | -0.003 | (0.088) | – |
| Reg.-Bez. Mittelfranken | -0.199 | (0.100) | -0.239 | (0.099) | – |
| Reg.-Bez. Münster | 0.181 | (0.101) | 0.171 | (0.103) | – |
| Reg.-Bez. Niederbayern | -0.153 | (0.137) | -0.177 | (0.142) | – |
| Reg.-Bez. Oberbayern | 0.044 | (0.089) | 0.023 | (0.090) | – |
| Reg.-Bez. Oberfranken | 0.011 | (0.114) | -0.019 | (0.117) | – |
| Reg.-Bez. Oberpfalz | 0.111 | (0.124) | 0.100 | (0.128) | – |
| Reg.-Bez. Schwaben | -0.056 | (0.116) | -0.080 | (0.117) | – |
| Reg.-Bez. Stuttgart | -0.020 | (0.089) | -0.048 | (0.090) | – |
| Reg.-Bez. Tübingen | 0.034 | (0.111) | 0.002 | (0.113) | – |
| Reg.-Bez. Unterfranken | 0.093 | (0.128) | 0.077 | (0.131) | – |
| Saarland | 0.232 | (0.140) | 0.245 | (0.144) | – |
| Sachsen-Anhalt | -0.256 | (0.106) | -0.278 | (0.105) | – |
| Schleswig-Holstein | 0.061 | (0.095) | 0.011 | (0.097) | – |
| Statistische Region Braunschweig | 0.118 | (0.117) | 0.096 | (0.120) | – |
| Statistische Region Hannover | -0.043 | (0.098) | -0.016 | (0.099) | – |
| Statistische Region Lüneburg | 0.074 | (0.117) | 0.025 | (0.119) | – |
| Statistische Region Weser-Ems | 0.014 | (0.104) | -0.012 | (0.105) | – |
| Thüringen | -0.055 | (0.101) | -0.074 | (0.104) | – |
| früher: Reg.-Bez. Koblenz | 0.333 | (0.186) | 0.299 | (0.203) | – |
| früher: Reg.-Bez. Rheinhessen-Pfalz | 0.078 | (0.104) | 0.057 | (0.106) | – |
| früher: Reg.-Bez. Trier | 0.078 | (0.168) | 0.053 | (0.169) | – |
| Constant | -0.021 | (0.064) | 0.023 | (0.205) | – |
| Socio-demographic controls | – | | *yes* | | *yes* |
| Observations | 5843 | | 5660 | | 5660 |
| F-statistic (NUTS-2 dummies) | 1.426 | | 1.480 | | – |
| p-value (NUTS-2 dummies) | .0455 | | .0307 | | – |
| $R^2$ | 0.011 | | 0.034 | | 0.023 |

*Notes:* The baseline region is Berlin. Socio-demographic controls include age and age-squared, gender, education, income, employment status, household size, number of children, and an indicator for having children below age 16. Robust standard errors in parentheses.

**Table 2.A.9.** Prosociality and Measures of Social Capital

| | Turnout in 2019 election [%] | | | Civic associations per 100k pop. in 2008 | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Prosociality | 1.52 ** | 1.57 ** | 1.51 *** | 14.63 * | 10.79 * | 10.92 |
| | [0.37, 2.51] | [0.36, 2.93] | [0.56, 2.55] | [-1.06, 23.85] | [-1.91, 19.16] | [-7.49, 24.83] |
| Patience | – | -0.46 | -0.26 | – | 12.62 * | 12.78 |
| | | [-1.93, 0.58] | [-1.40, 0.79] | | [-1.57, 30.41] | [-12.39, 40.19] |
| Risk-taking | – | 0.74 | 0.36 | – | -10.75 | -16.82 ** |
| | | [-0.50, 1.74] | [-1.15, 1.75] | | [-24.60, 3.87] | [-30.38, -1.64] |
| County controls | *No* | *No* | *Yes* | *No* | *No* | *Yes* |
| Population mean | 61.37 | 61.37 | 61.37 | 280.82 | 280.82 | 280.82 |
| Observations | 401 | 401 | 401 | 401 | 401 | 401 |
| Clusters | 38 | 38 | 38 | 38 | 38 | 38 |
| $R^2$ | 0.096 | 0.117 | 0.542 | 0.019 | 0.035 | 0.415 |

*Notes:* Bootstrapped 95%-confidence-intervals in brackets (clustered at NUTS-2 level), obtained using wild bootstrapping with Rademacher-weights and 9,999 simulations. Control variables include log GDP per capita, log average income per capita, share of college graduates, share of non-German residents, share of population below age 18, share of population age 65 or above, and indicators for the degree of urbanization. Under civic associations, we include (non-profit) organizations focused on social and economic welfare, political asssociations, and interest groups, following a classification by Franzen and Botzen (2011). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.10.** Weekly Incidence at the Time of the Survey

| | $y_{c,t} = \log(cases_{c,t})$ in county $c$ and week $t$ | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Altruism | -0.1041 * | -0.1042 * | -0.0745 | -0.0742 | 0.0229 |
| | [-0.227, 0.011] | [-0.238, 0.018] | [-0.211, 0.047] | [-0.188, 0.024] | [-0.122, 0.163] |
| Trust | 0.1055 | 0.1095 | 0.0655 | 0.0493 | 0.0649 |
| | [-0.048, 0.293] | [-0.065, 0.304] | [-0.079, 0.217] | [-0.072, 0.182] | [-0.023, 0.177] |
| Positive Reciprocity | -0.0640 | -0.0657 | -0.0640 | -0.0527 | -0.0296 |
| | [-0.179, 0.079] | [-0.178, 0.086] | [-0.209, 0.089] | [-0.183, 0.071] | [-0.133, 0.067] |
| Negative Reciprocity (ind.) | -0.1018 | -0.1017 | -0.1032 | -0.0910 * | -0.0428 |
| | [-0.239, 0.028] | [-0.258, 0.056] | [-0.240, 0.041] | [-0.198, 0.020] | [-0.135, 0.055] |
| Patience | – | 0.0051 | 0.0323 | 0.0215 | 0.0718 * |
| | | [-0.168, 0.175] | [-0.102, 0.210] | [-0.081, 0.158] | [-0.021, 0.230] |
| Risk-taking | – | -0.0157 | -0.0484 | -0.0550 | -0.0837 ** |
| | | [-0.092, 0.066] | [-0.139, 0.058] | [-0.128, 0.044] | [-0.158, -0.020] |
| Public health behavior | – | – | – | – | -0.2878 *** |
| | | | | | [-0.431, -0.144] |
| Wave 1 severity | *No* | *No* | *No* | *Yes* | *Yes* |
| County controls × Week | *No* | *No* | *Yes* | *Yes* | *Yes* |
| Week fixed effects | *Yes* | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 3609 | 3609 | 3609 | 3609 | 3609 |
| Spatial units (counties) | 401 | 401 | 401 | 401 | 401 |
| Clusters | 38 | 38 | 38 | 38 | 38 |
| $R^2$ | 0.170 | 0.171 | 0.385 | 0.433 | 0.492 |

*Notes:* This table estimates the same specifications as in Table 2.2, but with the individual social preferences of altruism, trust, positive reciprocity and indirect negative reciprocity as independent variables instead of prosociality. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Table 2.A.11.** Weekly Growth Rate of Confirmed Cases at the Time of the Survey

| | $y_{c,t} = \log(cases_{c,t}) - \log(cases_{c,t-1})$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Altruism | -0.0028 | -0.0025 | -0.0119 | -0.0010 |
| | [-0.015, 0.009] | [-0.020, 0.016] | [-0.031, 0.005] | [-0.027, 0.028] |
| Trust | -0.0037 | -0.0037 | 0.0037 | 0.0062 |
| | [-0.016, 0.012] | [-0.017, 0.014] | [-0.010, 0.023] | [-0.010, 0.026] |
| Positive Reciprocity | -0.0073 | -0.0072 | -0.0114 | -0.0092 |
| | [-0.022, 0.006] | [-0.021, 0.006] | [-0.031, 0.007] | [-0.026, 0.006] |
| Negative Reciprocity (ind.) | 0.0016 | 0.0017 | -0.0116 | -0.0066 |
| | [-0.010, 0.015] | [-0.010, 0.016] | [-0.030, 0.006] | [-0.024, 0.011] |
| Patience | -0.0029 | -0.0028 | 0.0013 | 0.0074 |
| | [-0.013, 0.006] | [-0.014, 0.007] | [-0.015, 0.022] | [-0.010, 0.032] |
| Risk-taking | 0.0005 | 0.0004 | -0.0055 | -0.0097 |
| | [-0.012, 0.014] | [-0.014, 0.014] | [-0.020, 0.009] | [-0.028, 0.007] |
| Public health behavior | – | -0.0010 | – | -0.0341 ** |
| | | [-0.027, 0.023] | | [-0.069, -0.005] |
| $\log(cases_{c,t-2})$ | – | – | -0.1112 *** | -0.1234 *** |
| | | | [-0.129, -0.096] | [-0.148, -0.101] |
| Policy stringency$_{c,t-2}$ | – | – | -0.2759 | -0.2260 |
| | | | [-1.009, 0.298] | [-0.901, 0.250] |
| Wave 1 severity | *Yes* | *Yes* | *Yes* | *Yes* |
| County controls × Week | *Yes* | *Yes* | *Yes* | *Yes* |
| Week fixed effects | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 3609 | 3609 | 3609 | 3609 |
| $R^2$ | 0.294 | 0.294 | 0.316 | 0.318 |

*Notes:* This table estimates the same specifications as in Table 2.3, but with the individual social preferences of altruism, trust, positive reciprocity and indirect negative reciprocity as independent variables instead of prosociality. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.12.** Effect of Preferences and Behavior on Weekly Deaths

| | $y = \log deaths_t$ | | | $y = \log deaths_t - \log deaths_{t-1}$ | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Prosociality | -0.1272 * | -0.1241 * | 0.0488 | -0.0134 * | -0.0051 |
| | [-0.315, 0.009] | [-0.288, 0.007] | [-0.089, 0.176] | [-0.033, 0.000] | [-0.035, 0.019] |
| Patience | -0.0095 | -0.0163 | 0.0678 | -0.0010 | 0.0032 |
| | [-0.174, 0.207] | [-0.180, 0.180] | [-0.051, 0.222] | [-0.015, 0.020] | [-0.013, 0.024] |
| Risk-taking | -0.0271 | -0.0307 | -0.0852 | -0.0147 | -0.0181 |
| | [-0.139, 0.110] | [-0.134, 0.107] | [-0.196, 0.022] | [-0.048, 0.013] | [-0.053, 0.016] |
| Public health behavior | – | – | -0.3851 *** | – | -0.0197 |
| | | | [-0.520, -0.240] | | [-0.056, 0.022] |
| $\log cases_{t-2}$ | – | – | – | -0.1476 *** | -0.1549 *** |
| | | | | [-0.195, -0.103] | [-0.211, -0.101] |
| Policy measures$_{t-2}$ | – | – | – | -0.2032 | -0.1806 |
| | | | | [-1.079, 0.321] | [-0.948, 0.299] |
| Wave 1 severity | | *Yes* | *Yes* | *Yes* | *Yes* |
| County controls × Week | *Yes* | *Yes* | *Yes* | *Yes* | *Yes* |
| Week fixed effects | *Yes* | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 3395 | 3395 | 3395 | 3213 | 3213 |
| Spatial units (counties) | 401 | 401 | 401 | 401 | 401 |
| Clusters | 38 | 38 | 38 | 38 | 38 |
| $R^2$ | 0.249 | 0.257 | 0.299 | 0.090 | 0.090 |

*Notes:* Bootstrapped 95%-confidence-intervals in brackets (clustered at NUTS-2 level), obtained using wild bootstrapping with Rademacher-weights and 9,999 simulations. The outcome variable is the (change in the) log of weekly deaths per 100000 population in a county, ranging from Nov 11th 2020 until Jan 17th 2021. Controls for wave 1 severity include the log of aggregate case numbers, its square, and case fatality rate in the time period from the first confirmed case until May 17th, 2020. County controls include log population density, log GDP per capita, log average income per capita, share of college graduates, employment share, share of non-German residents, share of workers in the service sector, share of population below age 18, share of population age 65 or above, and border country dummies for each neighboring country of Germany. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Table 2.A.13.** Individual-Level Association with PHB — East and West Germany

| | Public Health Behavior (PHB) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Prosociality | 0.3269*** | 0.2971*** | 0.2974*** | 0.2073*** | 0.1529*** |
| | (0.0177) | (0.0180) | (0.0181) | (0.0184) | (0.0155) |
| Prosociality × East Germany | 0.0573 | 0.0569 | 0.0589 | 0.0647* | 0.0485 |
| | (0.0392) | (0.0385) | (0.0391) | (0.0375) | (0.0302) |
| Patience | 0.1939*** | 0.1924*** | 0.1913*** | 0.1680*** | 0.0804*** |
| | (0.0149) | (0.0149) | (0.0149) | (0.0149) | (0.0126) |
| Risk-taking | -0.2100*** | -0.1710*** | -0.1718*** | -0.1708*** | -0.0780*** |
| | (0.0140) | (0.0143) | (0.0143) | (0.0138) | (0.0108) |
| Negative reciprocity (Direct) | -0.1228*** | -0.1075*** | -0.1070*** | -0.0671*** | -0.0178 |
| | (0.0141) | (0.0145) | (0.0145) | (0.0150) | (0.0124) |
| Socio-demographic factors | *No* | *Yes* | *Yes* | *Yes* | *Yes* |
| East Germany dummy | *Yes* | *Yes* | *Yes* | *Yes* | *Yes* |
| NUTS-2 region FEs | *No* | *No* | *Yes* | *Yes* | *Yes* |
| Big 5 personality traits | *No* | *No* | *No* | *Yes* | *Yes* |
| COVID-19 perceptions | *No* | *No* | *No* | *No* | *Yes* |
| Observations | 5843 | 5660 | 5660 | 5660 | 5660 |
| Clusters | 397 | 396 | 396 | 396 | 396 |
| $R^2$ | 0.213 | 0.239 | 0.243 | 0.299 | 0.495 |

*Notes:* Socio-demographic controls include age and age-squared, gender, education, income, employment status, household size, number of children, and an indicator for having children below age 16. COVID-19 perceptions include general attitudes towards the pandemic, infection experiences, and worrying about oneself, family members, and others being infected. SEs (in parentheses) are clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.14.** Weekly Incidence at the Time of the Survey — East and West Germany

| | $y_{c,t} = \log(cases_{c,t})$ in county $c$ and week $t$ | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Prosociality | -0.1108 ** | -0.1286 ** | -0.0943 ** | -0.0927 ** | 0.0021 |
| | [-0.246, -0.024] | [-0.331, -0.016] | [-0.238, -0.003] | [-0.200, -0.024] | [-0.116, 0.091] |
| Prosociality × East Germany | 0.0312 | 0.0375 | -0.0178 | 0.0005 | -0.0056 |
| | [-1.577, 0.992] | [-1.422, 1.100] | [-1.885, 0.793] | [-1.557, 0.821] | [-1.352, 0.787] |
| Patience | – | 0.0256 | 0.0447 | 0.0386 | 0.0741 ** |
| | | [-0.092, 0.208] | [-0.052, 0.206] | [-0.052, 0.162] | [0.007, 0.194] |
| Risk-taking | – | 0.0386 | -0.0299 | -0.0377 | -0.0661 |
| | | [-0.110, 0.188] | [-0.144, 0.095] | [-0.128, 0.082] | [-0.120, 0.038] |
| Public health behavior | – | – | – | – | -0.2194 *** |
| | | | | | [-0.354, -0.071] |
| Wave 1 severity | *No* | *No* | *No* | *Yes* | *Yes* |
| County controls × Week | *No* | *No* | *Yes* | *Yes* | *Yes* |
| East Germany × Week | *Yes* | *Yes* | *Yes* | *Yes* | *Yes* |
| Week fixed effects | *Yes* | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 3609 | 3609 | 3609 | 3609 | 3609 |
| Spatial units (counties) | 401 | 401 | 401 | 401 | 401 |
| Clusters (NUTS-2 regions) | 38 | 38 | 38 | 38 | 38 |
| $R^2$ | 0.189 | 0.194 | 0.424 | 0.483 | 0.513 |

*Notes:* Bootstrapped 95%-confidence-intervals in brackets (clustered at NUTS-2 level), obtained using wild bootstrapping with Rademacher-weights and 9,999 simulations. The time period of analysis ranges from Nov 16, 2020, until Jan 17, 2021. County controls include log population density, log GDP per capita, log average income per capita, share of college graduates, employment share, share of non-German residents, share of workers in the service sector, share of population below age 18, share of population age 65 or above, and border country dummies for each neighboring country of Germany. Controls for wave 1 severity include the log of aggregate case numbers, its square, and case fatality rate in the time period from the first confirmed case until May 17th, 2020. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.15.** Weekly Growth Rate of Confirmed Cases — East and West Germany

| | $y_{c,t} = \log(cases_{c,t}) - \log(cases_{c,t-1})$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Prosociality | -0.0053 | -0.0102 * | -0.0163 *** | -0.0086 |
| | [-0.016, 0.004] | [-0.024, 0.001] | [-0.034, -0.006] | [-0.027, 0.004] |
| Prosociality × East Germany | -0.0074 | -0.0071 | -0.0068 | -0.0072 |
| | [-0.064, 0.009] | [-0.068, 0.016] | [-0.220, 0.129] | [-0.210, 0.123] |
| Patience | 0.0015 | -0.0003 | 0.0058 | 0.0090 * |
| | [-0.012, 0.010] | [-0.014, 0.008] | [-0.003, 0.019] | [-0.002, 0.025] |
| Risk-taking | 0.0009 | 0.0024 | -0.0038 | -0.0065 |
| | [-0.010, 0.013] | [-0.008, 0.012] | [-0.013, 0.011] | [-0.018, 0.008] |
| Public health behavior | – | 0.0113 | – | -0.0186 |
| | | [-0.010, 0.030] | | [-0.043, 0.005] |
| $\log(cases_{c,t-2})$ | – | – | -0.1232 *** | -0.1285 *** |
| | | | [-0.147, -0.103] | [-0.155, -0.105] |
| Policy stringency$_{c,t-2}$) | – | – | -0.1672 | -0.1548 |
| | | | [-0.632, 0.181] | [-0.593, 0.153] |
| Wave 1 severity | *Yes* | *Yes* | *Yes* | *Yes* |
| County controls × Week | *Yes* | *Yes* | *Yes* | *Yes* |
| East Germany × Week | *Yes* | *Yes* | *Yes* | *Yes* |
| Week fixed effects | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 3609 | 3609 | 3609 | 3609 |
| Spatial units (counties) | 401 | 401 | 401 | 401 |
| Clusters (NUTS-2 regions) | 38 | 38 | 38 | 38 |
| $R^2$ | 0.302 | 0.302 | 0.327 | 0.328 |

*Notes:* Bootstrapped 95%-confidence-intervals in brackets (clustered at NUTS-2 level), obtained using wild bootstrapping with Rademacher-weights and 9,999 simulations. The outcome variable is the change in the log of weekly cases per capita in a county, ranging from Nov 16th 2020 until Jan 17th 2021. County controls include log population density, log GDP per capita, log average income per capita, share of college graduates, employment share, share of non-German residents, share of workers in the service sector, share of population below age 18, share of population age 65 or above, and border country dummies for each neighboring country of Germany. Controls for wave 1 severity include the log of aggregate case numbers, its square, and case fatality rate in the time period from the first confirmed case until May 17th, 2020. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.16.** Weekly Incidence at the Time of the Survey — Standardized

| | $y_{c,t} = \log(cases_{c,t})$ in county $c$ and week $t$ | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Prosociality | -0.0820 *** | -0.0749 * | -0.0732 ** | -0.0701 ** | 0.0108 |
| | [-0.167, -0.036] | [-0.178, 0.006] | [-0.174, -0.013] | [-0.145, -0.019] | [-0.052, 0.063] |
| Patience | – | -0.0168 | 0.0014 | -0.0032 | 0.0355 |
| | | [-0.125, 0.079] | [-0.069, 0.106] | [-0.065, 0.076] | [-0.011, 0.111] |
| Risk taking | – | 0.0062 | -0.0222 | -0.0268 | -0.0480 * |
| | | [-0.063, 0.074] | [-0.091, 0.054] | [-0.081, 0.042] | [-0.088, 0.003] |
| Public health behavior | – | – | – | – | -0.1767 *** |
| | | | | | [-0.262, -0.093] |
| Wave 1 severity | *No* | *No* | *No* | *Yes* | *Yes* |
| County controls × Week | *No* | *No* | *Yes* | *Yes* | *Yes* |
| Week fixed effects | *Yes* | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 3609 | 3609 | 3609 | 3609 | 3609 |
| Spatial units (counties) | 401 | 401 | 401 | 401 | 401 |
| Clusters (NUTS-2 regions) | 38 | 38 | 38 | 38 | 38 |
| $R^2$ | 0.116 | 0.118 | 0.357 | 0.415 | 0.481 |

*Notes:* This table estimates the same specifications as Table 2.2, but with the dependent variable standardized. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Table 2.A.17.** Weekly Growth Rate of Confirmed Cases — Standardized

| | $y_{c,t} = \log(cases_{c,t}) - \log(cases_{c,t-1})$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Prosociality | -0.0133 ** | -0.0141 | -0.0318 *** | -0.0105 |
| | [-0.027, -0.002] | [-0.033, 0.003] | [-0.055, -0.016] | [-0.037, 0.011] |
| Patience | -0.0018 | -0.0021 | -0.0017 | 0.0090 |
| | [-0.020, 0.010] | [-0.022, 0.013] | [-0.016, 0.020] | [-0.012, 0.038] |
| Risk taking | 0.0002 | 0.0004 | -0.0064 | -0.0134 |
| | [-0.017, 0.019] | [-0.018, 0.018] | [-0.023, 0.014] | [-0.038, 0.011] |
| Public health behavior | – | 0.0018 | – | -0.0496 ** |
| | | [-0.031, 0.032] | – | [-0.096, -0.009] |
| $\log(cases_{c,t-2})$ | – | – | -0.1578 *** | -0.1765 *** |
| | | | [-0.183, -0.135] | [-0.214, -0.141] |
| Policy stringency$_{c,t-2}$ | – | – | -0.3508 | -0.2993 |
| | | | [-1.251, 0.422] | [-1.117, 0.332] |
| Wave 1 severity | *Yes* | *Yes* | *Yes* | *Yes* |
| County controls × Week | *Yes* | *Yes* | *Yes* | *Yes* |
| Week fixed effects | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 3609 | 3609 | 3609 | 3609 |
| Spatial units (counties) | 401 | 401 | 401 | 401 |
| Clusters (NUTS-2 regions) | 38 | 38 | 38 | 38 |
| $R^2$ | 0.293 | 0.293 | 0.315 | 0.317 |

*Notes:* This table estimates the same specifications as Table 2.3, but with the dependent variable standardized. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Table 2.A.18.** Overall Number of Confirmed Cases in First and Second Wave

| | $y_i$ = *log overall confirmed cases per 100000 population in county i* | | | |
|---|---|---|---|---|
| | "first wave" | | "second wave" | |
| | (1) | (2) | (3) | (4) |
| Prosociality | – | -0.0546 | – | -0.0913 ** |
| | | [-0.186, 0.053] | | [-0.231, -0.011] |
| Patience | – | 0.0113 | – | 0.0025 |
| | | [-0.110, 0.182] | | [-0.092, 0.146] |
| Risk-taking | – | 0.0938 | – | -0.0238 |
| | | [-0.017, 0.212] | | [-0.124, 0.089] |
| log population density | 0.4055 ** | 0.4142 ** | 0.0634 | 0.0847 |
| | [0.045, 0.738] | [0.047, 0.757] | [-0.192, 0.347] | [-0.151, 0.341] |
| Employed / population | 3.5720 *** | 3.6969 *** | 1.5709 * | 1.4675 * |
| | [2.072, 5.091] | [2.150, 5.276] | [-0.056, 3.428] | [-0.156, 3.458] |
| Share of jobs in service sector | -3.1460 *** | -3.0334 *** | -1.4531 * | -1.4196 * |
| | [-4.694, -1.551] | [-4.559, -1.429] | [-3.077, 0.086] | [-3.052, 0.078] |
| Further county characteristics | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 401 | 401 | 401 | 401 |
| Clusters | 38 | 38 | 38 | 38 |
| $R^2$ | 0.497 | 0.509 | 0.265 | 0.323 |

*Notes:* Bootstrapped 95%-confidence-intervals in brackets (clustered at NUTS-2 level), obtained using wild bootstrapping with Rademacher-weights and 9,999 simulations. The "first wave" is defined as the time period until May 17th, 2020; the "second wave" is defined as time period between Sep 28th 2020 and Feb 28th 2021. Further regressors include log GDP per capita, log average income per capita, share of college graduates, share of non-German residents, share of population below age 18, share of population age 65 or above, and border country dummies for each neighboring country of Germany. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Table 2.A.19.** Aggregate Number of Deaths in First and Second Wave

| | $y_i$ = *log COVID-19 deaths per 100000 population in county i* | | | |
|---|---|---|---|---|
| | "First wave" | | "Second wave" | |
| | (1) | (2) | (3) | (4) |
| Prosociality | – | -0.1835 ** | – | -0.1157 * |
| | | [-0.373, -0.043] | | [-0.312, 0.003] |
| Patience | – | 0.0571 | – | -0.0345 |
| | | [-0.106, 0.261] | | [-0.157, 0.185] |
| Risk-taking | – | 0.2022 *** | – | -0.0254 |
| | | [0.066, 0.376] | | [-0.144, 0.101] |
| log population density | 0.2898 | 0.3214 | 0.0433 | 0.0686 |
| | [-0.175, 0.786] | [-0.147, 0.789] | [-0.256, 0.378] | [-0.200, 0.353] |
| Employed / population | 5.1239 *** | 5.4715 *** | 1.1927 | 0.9480 |
| | [2.655, 7.635] | [2.984, 7.960] | [-0.743, 3.316] | [-1.224, 3.444] |
| Share of jobs in service sector | -4.1070 *** | -3.8792 *** | -1.1468 | -1.0467 |
| | [-6.428, -1.782] | [-6.176, -1.563] | [-3.186, 0.791] | [-3.141, 0.904] |
| Further county controls | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 381 | 381 | 401 | 401 |
| Clusters | 38 | 38 | 38 | 38 |
| $R^2$ | 0.288 | 0.322 | 0.272 | 0.321 |

*Notes:* Bootstrapped 95%-confidence-intervals in brackets (clustered at NUTS-2 level), obtained using wild bootstrapping with Rademacher-weights and 9,999 simulations. The "first wave" is defined as the time period until May 17th, 2020; the "second wave" is defined as time period between Sep 28th 2020 and Feb 28th 2021. Further controls include log average income per capita, share of college graduates, share of non-German residents, share of population below age 18, share of population age 65 or above, and border country dummies for each neighboring country of Germany. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Table 2.A.20.** Aggregate Number of Cases and Deaths in Third Wave

| | *"Third wave": starting from March 1st, 2021* | | | |
|---|---|---|---|---|
| | log cumulative cases | | log cumulative deaths | |
| | (1) | (2) | (3) | (4) |
| Prosociality | – | -0.1020 *** | – | -0.0947 ** |
| | | [-0.186, -0.049] | | [-0.240, -0.004] |
| Patience | – | 0.0106 | – | -0.0257 |
| | | [-0.064, 0.121] | | [-0.134, 0.163] |
| Risk-taking | – | 0.0220 | – | -0.0097 |
| | | [-0.036, 0.106] | | [-0.110, 0.113] |
| log population density | 0.0973 | 0.1196 | 0.0773 | 0.0977 |
| | [-0.072, 0.258] | [-0.030, 0.266] | [-0.259, 0.389] | [-0.230, 0.410] |
| log GDP per capita | -0.0826 ** | -0.1105 ** | -0.4362 *** | -0.4030 *** |
| | [0.489, 9.790] | [0.237, 9.521] | [3.679, 12.716] | [3.037, 12.165] |
| Employed / population | 2.3504 *** | 2.3284 *** | 2.0686 | 1.8922 |
| | [0.731, 4.251] | [0.676, 4.329] | [-0.514, 4.643] | [-0.962, 4.837] |
| Share of jobs in service sector | -2.3614 *** | -2.2861 *** | -2.3779 * | -2.2885 * |
| | [-4.179, -0.668] | [-4.092, -0.604] | [-4.840, 0.099] | [-4.825, 0.192] |
| Population share age 65 or above | 5.3880 | 5.0146 | 11.4636 | 10.9036 |
| | [-0.412, 0.301] | [-0.415, 0.236] | [-0.516, 0.549] | [-0.517, 0.579] |
| Further county controls | *Yes* | *Yes* | *Yes* | *Yes* |
| Observations | 401 | 401 | 401 | 401 |
| Clusters | 38 | 38 | 38 | 38 |
| $R^2$ | 0.305 | 0.365 | 0.319 | 0.346 |

*Notes:* Bootstrapped 95%-confidence-intervals in brackets (clustered at NUTS-2 level), obtained using wild bootstrapping with Rademacher-weights and 9,999 simulations. Dependent variables are log cumulative cases (deaths) per 100000 population. The time period of analysis goes until July 8, 2021. Further controls include log average income per capita, share of college graduates, share of non-German residents, share of population below age 18, share of population age 65 or above, and border country dummies for each neighboring country of Germany. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

## Appendix 2.B Survey Questions and Data

In this section, we describe all the survey questions that respondents were asked to complete as part of the questionnaire (subsections 2.B.1- 2.B.6), including those that we use to construct major dependent or independent variables for the main paper, i.e. pandemic-related behavior and prosocial preferences. We translated all questions into English for this Section. For the complete original questionnaire in German language, see Section 2.D. In subsections 2.B.7, we describe our sample recruiting and data cleaning procedures, and in subsection 2.B.8, we describe how we construct our individual-level variables based on the survey items.

### 2.B.1 Public Health Behavior

To what extent do the following statements apply to your own behavior? *Please rate on a scale from 1 to 7. The value 1 means: does not apply at all. The value 7 means: fully applies.*

- I keep a distance of at least 1.5 meters from other people.
- I will socially isolate myself if I have had contact with an infected person.
- I always keep up to date on news about the pandemic.
- I wash and disinfect my hands regularly.
- I am going to get vaccinated against the coronavirus when a vaccine becomes available.
- I cough and sneeze into the crook of my elbow.
- I wear a face mask in public.
- I ventilate regularly when several people are using a room.
- I avoid social contacts as much as possible.
- I will inform other people if I am infected with the coronavirus.

### 2.B.2 Questions from the Peference Survey Module

To elicit time, risk, and social preferences, we included some questions from experimentally-validated preference survey module by Falk et al. (2016) and Falk et al. (2018) in our questionnaire. All qualitative questions were rated on an 11-point Likert scale from 0 to 10, where the value of 0 indicates complete disagreement or unwillingness, and the value 10 indicates complete agreement or willingness.

Altruism was elicited using one qualitative question and a quantitative decision involving a hypothetical donation. Positive reciprocity, indirect negative reciprocity, and trust were elicited using one qualitative question each. Direct negative reciprocity was elicited using two qualitative questions, and patience and risk taking were elicited using one qualitative item each.

#### Altruism, Reciprocity, and Trust

How willing are you to give to good causes without expecting anything in return.

not willing at all □—□—□—□—□—□—□—□—□—□ very willing to do it

Imagine the following situation:

> Today you unexpectedly received 1,000 euros. How much of this amount would you donate to a good cause? ____________

If someone does me a favor, I am willing to return it.

> does not describe me at all □—□—□—□—□—□—□—□—□—□ describes me perfectly

How willing are you to punish someone who treats you unfairly, even if there may be costs for you?

> not at all willing to do it □—□—□—□—□—□—□—□—□—□ very willing to do it

If I am treated very unfairly, I will take revenge at the first occasion, even if there is a cost to do so.

> does not describe me at all □—□—□—□—□—□—□—□—□—□ describes me perfectly

How willing are you to punish someone who treats others unfairly, even if there may be costs for you?

> not at all willing to do it □—□—□—□—□—□—□—□—□—□ very willing to do it

I assume that people have only the best intentions.

> does not describe me at all □—□—□—□—□—□—□—□—□—□ describes me perfectly

**Risk and time preferences**

In general, how willing are you to take risks?

> completely unwilling to take risks □—□—□—□—□—□—□—□—□—□ very willing to take risks

How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?

> completely unwilling to do so □—□—□—□—□—□—□—□—□—□ very willing to do so

**Control questions**

I am good at math.

does not describe me at all ☐—☐—☐—☐—☐—☐—☐—☐—☐—☐—☐ describes me perfectly

I tend to put off tasks even when I know it would be better to do them now.

does not describe me at all ☐—☐—☐—☐—☐—☐—☐—☐—☐—☐—☐ describes me perfectly

### 2.B.3 Demographic and socio-economic questions

Please enter your year of birth. ____________

Please select your gender.
- ☐ Female
- ☐ Male
- ☐ Others

Which state do you live in? ____________

Please enter your zip code? ____________

How long have you been living at your current place of residence? ____________

What is your highest educational qualification?
- ☐ No degree
- ☐ Elementary/secondary school certificate (GDR: 8th grade)
- ☐ Secondary school leaving certificate (GDR: 10th grade)
- ☐ Fachhochschulreife (qualification from a technical college)
- ☐ Abitur/university entrance qualification
- ☐ Fachhochschule (formerly: engineering school, teacher training, GDR: engineer and technical college degree)
- ☐ University, college degree
- ☐ Doctorate
- ☐ Other educational qualification

How many people live in your household (ie unit living and working together)? ____________

How many children do you have? ____________

Which of the following best describes your current employment status?
- ☐ full-time employed
- ☐ part-time employed
- ☐ self-employed
- ☐ in educational/vocational training
- ☐ non-employed

What is approximately your net monthly household income in Euro?
- ☐ under €900
- ☐ €900 to under €1,300
- ☐ €1,300 to less than €1,500
- ☐ €1,500 to less than €2,000
- ☐ €2,000 to less than €2,600
- ☐ €2,600 to less than €3,200
- ☐ €3,200 to less than €4,500
- ☐ €4,500 to less than €6,000
- ☐ €6,000 and more
- ☐ not specified

### 2.B.4 Big Five personality index

How well does each of the following statements describe you as a person?
*Please answer as honestly and spontaneously as possible on a scale from 1 to 7. The value 1 means: Not at all applicable. The value 7 means: Completely applies.*

I am someone who...

- ... works thoroughly.
- ... is communicative, talkative.
- ... is sometimes rude to others.
- ... is original, brings in new ideas.
- ... worries a lot.
- ... can forgive.
- ... is rather lazy.
- ... is outgoing and sociable.
- ... values artistic experiences.
- ... gets nervous easily.
- ... gets tasks done effectively and efficiently.

- ... is reserved.
- ... treats others with respect and kindness.
- ... has a vivid imagination.
- ... is relaxed, can handle stress well.

### 2.B.5 Other pandemic-related questions

How much do you agree with the following statements? *Please rate on a scale from 1 to 5. The value 1 means: completely disagree. The value 5 means: completely agree.*

- The pandemic has a negative effect on my financial situation.
- The pandemic has a negative effect on my personal life.
- The government's measures against the pandemic are way too strict.
- Overall, Germany has managed the pandemic well so far.
- The media takes COVID-19 way too seriously.

Have you contracted COVID-19 before?

- ☐ Yes
- ☐ No
- ☐ prefer not to say

Do you personally know someone who has contracted COVID-19?

- ☐ Yes
- ☐ No
- ☐ prefer not to say

Do you personally know someone who has died from Covid-19?

- ☐ Yes
- ☐ No
- ☐ prefer not to say

When was the last time you had the flu? ____________

The 7-day incidence rate indicates the number of new COVID-19 cases (i.e., people who tested positive for the coronavirus) per 100,000 inhabitants within the past seven days. It is considered an important indicator for assessing the current pandemic situation.

Please estimate the 7-day incidence rate in your city (note: as of December 17, the value for all of Germany is 179). ____________

*Please rate on a scale from 1 to 7. The value 1 means: not worried at all. The value 7 means: extremely worried.*
How worried are you about ...

- ... contracting COVID-19 yourself.
- ... friends or relatives contracting COVID-19.
- ... other people in general contracting COVID-19.

How high do you rate the risk of contracting COVID-19 within the next 3 months?

Very low □—□—□—□—□—□—□ Very high

Have you installed the Corona-Warn-App on your current mobile phone?

□ Yes

□ No

Which type of smartphone do you use most of the time?

□ Android smartphone (e.g. Samsung, Huawei, ...)

□ iPhone (Apple)

□ other smartphone (e.g. Windows-Phone, Blackberry, ...)

□ I don't use a smartphone

If not: How likely is it that you would install the Corona-Warn-App within the next few weeks?

Very unlikely □—□—□—□—□—□—□ Very likely

If yes: How likely is it that you would report your infection status to the Corona-Warn-App in case you would be tested positiv?

Very unlikely □—□—□—□—□—□—□ Very likely

How much do you agree with the following statements about the Corona-Warn-App?
*Please answer on a scale from 1 to 7. The value 1 means: I do not agree at all. The value 7 means: I completely agree.*

The Corona-Warn-App ...

- ... helps to slow down the spread of the coronavirus in Germany.
- ... helps to slow down the spread of the coronavirus in my city.
- ... is of no real use to me personally.
- ... is a good way to trace infection chains.
- ... is not used by enough people yet.

**Donation option:** The following scenario has a 25% probability of actually being implemented. So you should think carefully about what you want to do. It may involve real money.

You have 1 Euro at your disposal. You are free to decide how much of this you donate and what share you keep for yourself. Your donation will be used for an online advertising campaign on social media, which encourages more people (including in your region) to use the Corona-Warn-App. Past data has shown that 50 cents of advertising expenditure correspond to 1 additional Corona-Warn-App installation on average. You will get to keep the rest of the amount that you don't donate.

At the end of the survey, a random number generator determines whether this donation and the additional remuneration will actually be paid out. Please move the slider to decide on your allocation:

**How much would you like to donate?** Your donation is ______ Euro.

### 2.B.6 News consumption, political attitudes, and values

Where do you inform yourself about the news?
People use different news sources to learn about what is happening around them and in the world. For each of the following sources, please indicate how often you use them:

| | Daily | Weekly | Monthly | Less than monthly | Never |
|---|---|---|---|---|---|
| **Newspaper** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **TV** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Radio** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **News sites on the internet** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Mobile phone (WhatsApp, Telegram, etc.)** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Social media (Facebook, Twitter, etc.)** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Conversations with friends, colleagues and acquaintances** | ☐ | ☐ | ☐ | ☐ | ☐ |

About how much time do you spend on social media (e.g. Facebook, Instagram)?: ____________

If there were general elections tomorrow, which party would you vote for?

- ☐ CDU/CSU
- ☐ Buendnis'90/Die Gruenen
- ☐ SPD
- ☐ AfD
- ☐ Die Linke
- ☐ FDP
- ☐ Other

Would you actually vote?

- ☐ Yes
- ☐ No
- ☐ Undecided

How satisfied are you with how the political system in Germany works today?
*Please rate on a scale of 0 to 10, where 0 is "not at all satisfied" and 10 is "completely satisfied".*

not at all satisfied ☐—☐—☐—☐—☐—☐—☐—☐—☐—☐—☐ completely satisfied

How much do you agree with the following statements?

| | totally agree not to | | | | totally agree and quite to |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| **There are many very important things happening in the world which the public is never informed about.** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Government agencies monitor all citizens.** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **There are secret powers that control the world.** | ☐ | ☐ | ☐ | ☐ | ☐ |

People have different views about themselves and how strong they feel connected to their environment and the rest of the world.
How strongly do you feel connected to...

| | Not at all connected | A little connected | Somewhat connected | Quite connected | Very connected |
|---|---|---|---|---|---|
| **The town or city you live in** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **The region you live in** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Germany** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Europe** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **The whole world** | ☐ | ☐ | ☐ | ☐ | ☐ |

### 2.B.7 Data cleaning

The survey was administered to a sample of individuals in Germany through the market research company Dynata. Participants between 18 and 65 years old were recruited via email-invitation, with quotas on age, gender, and state to achieve national-level representativeness along these dimensions for the relevant age group of our sample. The questionnaire was web-based could be completed on-line on PC, laptop, tablet, or smartphone. It consisted of 20 pages in total and the median response time was about 13 minutes. A total 7,052 individuals responded to our survey, and 6,826 respondents completed every survey question on preferences and public health behavior. In accordance with Dynata policy, we used several different criteria to check response quality and to exclude bad responses: speeding (i.e. unreasonable quick response time), inconsistencies or conflicting answers, excessive straightlining (e.g. always ticking the same box in Likert scales), and an attention check question.

To check for speeding, we recorded the duration spent on answering questions on each page of the survey, as well as for completing the entire survey. We immediately excluded all responses where the survey-taker spent less than 2 seconds per question on average on at least 3 pages. Then, we flagged responses as potentially bad if the survey was completed in less than one-third of the median completion time. With regard to inconsistencies, we flagged responses as potentially bad if they would imply that the respondent became a parent at the age of 12 or younger, that the respondent lived at the current place of residence since before he or she was born, or if the zip code did not match the state of residence. With regard to straightlining, we flagged responses as potentially bad if they included at least 2 modules of Likert-scale-type sequences (e.g. preferences survey module, public health behavior) in which always the same value was selected. Finally, we flagged responses as potentially bad if the survey-taker failed an attention check question at the beginning of the survey. The attention check consisted of an absurd question ("[...] How interested are you in learning about the impact of traffic noise on the singing bird population in German cities?") for which the description prescribed a particular response in order to "demonstrate that you answer this survey carefully". We excluded all responses which were flagged as potentially bad in at least 2 out of 4 criteria. In total, 992 responses (i.e. below 15%) were removed for our analyses, thus giving us our main sample size of 5,843. In some analyses that include socioeconomic variables as controls, an additional 183 responses drop out due to missing information about education or income.

### 2.B.8 Variable Construction

#### Public Health Behavior

To construct the factor variable on public health behavior (PHB), we assume that compliance to public health behavior is driven by one underlying factor, and conduct factor analysis on the ten survey items on PHB (see Section 2.B.1). The results of our factor analysis support this notion. From Figure 2.B.1, we see that the eigenvalue on the first factor is 4.47, whereas those on the remaining factors are below 1. Table 2.B.1, which shows the factor loadings on each survey item, indicates that all survey items are highly correlated with the underlying factor. Furthermore, Cronbach's alpha is 0.87, indicating that all the PHB items are highly interrelated.

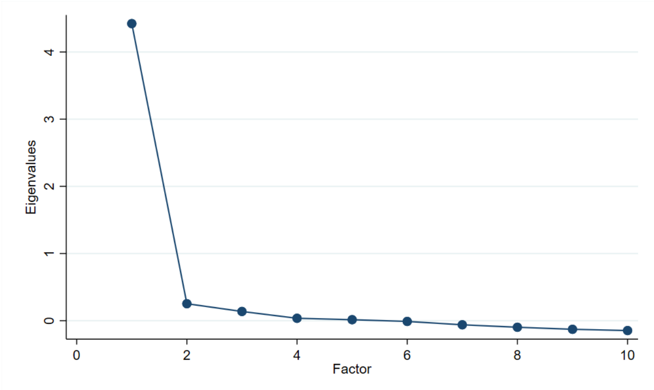



**Figure 2.B.1.** Scree Plot PHB

*Notes:* Eigenvalues on factors obtained from a factor analysis of PHB survey items.

**Table 2.B.1.** Factor Loadings PHB

| | Factor loadings |
|---|---|
| Social distancing 1.5 meters | 0.769 |
| Self-quarantine if risky contact | 0.746 |
| Keep informed about pandemic | 0.618 |
| Wash and disinfect hands | 0.686 |
| Get vaccinated when vaccine available | 0.434 |
| Sneeze and cough in elbow | 0.589 |
| Wear mask | 0.690 |
| Regular ventilation when indoors | 0.707 |
| Avoid social contacts | 0.713 |
| Would inform others if infected | 0.633 |

*Notes:* Factor loadings on survey items used to construct PHB.

### Prosociality

We construct the prosociality variable via principal component analysis on the five (standardized) survey items for altruism, positive reciprocity, trust, and indirect negative reciprocity. See Section 2.B.2 for the wording and scale of the questions. Note that we do not include the two questions on direct negative reciprocity ("If I am treated very unfairly, I will take revenge at the first occasion, even if there is a cost to do so.", and "How willing are you to punish someone who treats you unfairly, even if there may be costs for you?") as these do not square with our notion of prosociality. From Table 2.B.2, we see that the first principal component for prosociality explains approximately 36% of the total variance. The subsequent components explain 20%, 17%, 17%, and 10% of the variance respectively, which suggest that there are several distinct aspects to social preferences. Though these components could also explain adherence to PHB, this is not the aim of our study. Rather, our analysis is guided by theoretical considerations— We are interested in how a particular aspect of social preferences, i.e. prosociality, predicts adherence to PHB. In this regard, we see from Table 2.B.3 that the first principal component assigns weights to the underlying variables that are congruent with our notion of prosociality: 0.2 and 0.6 for the two altruism survey items, 0.49 for positive reciprocity, 0.4 for trust, and 0.4 for indirect negative reciprocity.

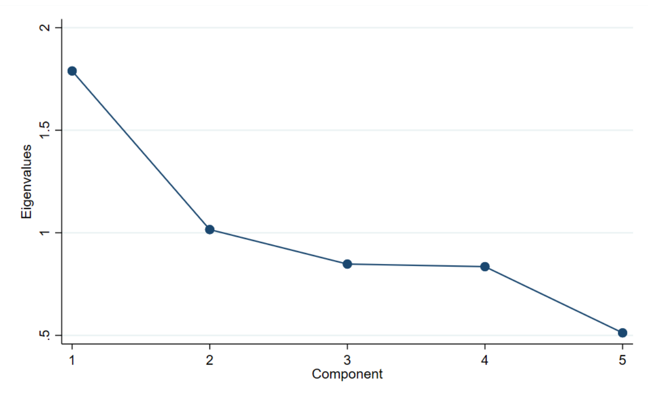



**Figure 2.B.2.** Scree Plot Prosociality

*Notes:* Eigenvalues on components obtained from a principal component analysis of prosocial preference survey items.

**Table 2.B.2.** Eigenvalues and Proportion of Total Variance, Prosocial Preferences Components

| | Eigenvalues | Proportion |
|---|---|---|
| Component 1 | 1.789 | 0.358 |
| Component 2 | 1.016 | 0.203 |
| Component 3 | 0.848 | 0.170 |
| Component 4 | 0.835 | 0.167 |
| Component 5 | 0.512 | 0.102 |

*Notes:* Eigenvalues and proportion of total variance on components of principal component analysis on standardized prosocial preferences survey items.

**Table 2.B.3.** Weights on Prosociality Survey Items, Prosocial Preferences Components

| | Comp 1 | Comp 2 | Comp 3 | Comp 4 | Comp 5 |
|---|---|---|---|---|---|
| Willingness to give for a good cause | 0.602 | -0.0638 | 0.00970 | -0.269 | -0.749 |
| Donation amount out of 1000 Euro | 0.257 | 0.835 | 0.129 | -0.381 | 0.274 |
| Postive reciprocity | 0.485 | -0.514 | -0.0608 | -0.403 | 0.578 |
| Negative reciprocity (indirect) | 0.415 | -0.00562 | 0.679 | 0.592 | 0.130 |
| General trust towards people | 0.405 | 0.187 | -0.720 | 0.519 | 0.113 |

*Notes:* Weights on prosociality survey items for each component obtained by principal component analysis of prosocial preferences survey items.

### Other Variables

**Table 2.B.4.** Overview of All Individual-Level Control Variables Used in the Paper

| Variable | Question(s) | Value formatting |
|---|---|---|
| Age | Please state your year of birth. | Age in years and age$^2$ |
| Gender | Please select your gender. | Categorical |
| County | Please enter your zip code | Categorical |
| Education | What is your highest educational attainment? | 4 categories: No degree, Secondary degree, Abitur, University degree |
| Income | What is approximately your net monthly household income in Euro? | 10 categories: <900, 900-1300, 1300-1500, 1500-2000, 2000-2600, 1.6k-3.2k, 3200-4500, 4500-6000, >6000, prefer not to say |
| Employment status | Which of the following best describes your current employment status? | 5 categories: full-time employed, part-time employed, self-employed, educational/vocational training, non-employed |
| Household size | How many people live in your household? | 5 categories: 1, 2, 3, 4, 5 or more |
| Number of children | How many children do you have? | 4 categories: none, 1, 2, 3 or more |
| Children below 16 | What is the age of your youngest child? | Indicator for age $\leq$ 16 years |
| Pandemic skeptical | 1) The government's measures against the pandemic are way too strict.<br>2) The media takes COVID-19 way too seriously. | Mean of two 5-point Likert scales (standardized) |
| Pandemic affected | 1) The pandemic has a negative effect on my financial situation.<br>1) The pandemic has a negative effect on my personal life. | Mean of two 5-point Likert scales (standardized) |
| Worry self | How worried are you about contracting COVID-19 yourself? | 7-point Likert scales |
| Worry family | How worried are you about friends or family contracting COVID-19? | 7-point Likert scales |
| Worry others | How worried are you about people in general contracting COVID-19? | 7-point Likert scales |
| Infected | Have you contracted COVID-19 before? | Categorical: yes, no, prefer not to say |
| Know infected | Do you personally know someone who has contracted COVID-19? | Categorical: yes, no, prefer not to say |
| Know died | Do you personally know someone who has died from COVID-19? | Categorical: yes, no, prefer not to say |
| Patience | How willing are you to give up sth that is beneficial for you today in order to benefit more from that in the future? | 11-point Likert scale |
| Risk taking | In general, how willing are you to take risks? | 11-point Likert scale |
| Big Five personality | 15 item BFI-S (Gerlitz and Schupp, 2005) | 5 standardized variables: openness, conscientiousness, extraversion, agreeableness, neuroticism |

**Table 2.B.5.** Survey Items Used to Construct Big Five Personality Factors

| Personality Trait | Definitions (Becker et al., 2012, p.466) | Survey items |
|---|---|---|
| Openness | Individual differences in the tendency to be open to new aesthetic, cultural, and intellectual experiences | ... is original, brings in new ideas.<br>... values artistic experiences.<br>... has a vivid imagination. |
| Conscientiousness | The tendency to be organized, responsible, and hardworking; located at one end of a dimension of individual differences (conscientiousness versus lack of direction) | ... works thoroughly.<br>... is rather lazy.<br>... tasks done effectively and efficiently. |
| Extraversion | An orientation of one's interests and energies toward the outerworld of people and things rather than the inner world of subjective experience; includes the qualities of being outgoing, gregarious, sociable, and openly expressive | ... is reserved.<br>... is communicative, talkative.<br>... can be outgoing, is sociable. |
| Agreeableness | The tendency to act in a cooperative, unselfish manner; located at one end of a dimension of individual differences (agreeableness versus disagreeableness) | ... sometimes being rude to others.<br>... can forgive.<br>... treat others with respect and kindness. |
| Neuroticism | A chronic level of emotional instability and proneness to psychological distress | ... often worries.<br>... is relaxed, can handle stress well.<br>... gets nervous easily. |

*Notes:* We construct each Big Five personality factor by conducting factor analysis on the relevant survey items, and then standardizing the resultant factor variable. Definitions are taken from Becker et al. (2012, p.466).

## Appendix 2.C Data Sources for Regional Data

### 2.C.1 Aggregation of Survey Measures

As the sample for our online survey was recruited to be representative only at the national-level, we weight observations to improve representativeness of at the NUTS-2 level. Specifically, we obtain official data on age, gender, and education by region (see 2.C.3) and calculate sampling weights to match the regional population with regard to age-by-gender ($2 \times 2$ matrix of age above/below 40 with female/male) and the share of adults with a university degree. We do so using a simple stepwise raking procedure (Battaglia, Hoaglin, and Frankel, 2009), in which we first calculate initial weights so that our sample matches the population age-gender distribution, then readjust these weights to match the share of adults with a university degree, then readjust to match age-gender again, and so on, until the weights converge. Using the final sampling weights, we then calculate the NUTS-2 region-level average of the PHB, prosociality, patience, and risk taking measures described in Section 2.B.

### 2.C.2 COVID-19 Incidences and Deaths

We obtained official data on the daily number of confirmed COVID-19 cases and deaths as reported by the Robert-Koch-Institut (RKI), the the federal government agency and research institute responsible for disease control and prevention in Germany. It can be publicly accessed via the Corona data hub (https://npgeo-corona-npgeo-de.hub.arcgis.com). The information is updated daily at the county level, although there can be delays in reporting by local health authorities, especially on weekends and on holidays. We therefore aggregate all numbers to the weekly level, with each week beginning on Monday and ending on Sunday. Furthermore, we adjust the number of cases and deaths by each county's population size to obtain the incidence rates, defined as number of confirmed cases/deaths per 100,000 population in a period of 7 days.

### 2.C.3 Demographic and Socio-Economic Information

We collect data on pre-pandemic county characteristics from the publicly accessible official database of the German federal statistical office and the state statistical offices (Regionaldatenbank, https://www.regionalstatistik.de/genesis/online). This includes information on population and demographics, education, economic indicators, employment statistics, etc. We complement this with data collected by infas360 in an effort to synthesize databases that can be relevant with regard to COVID-19 and make them available to researchers (Corona-Datenplattform, https://www.corona-datenplattform.de). In Table 2.C.1, we provide a complete list of all variables that we use in the paper, the data source, and from which year it is.

**Table 2.C.1.** Overview of All County-Level Control Variables Used in the Paper

| Variable | Year | Source(s) |
|---|---|---|
| Population density (settlement area only) | 2018 | Corona-Datenplattform |
| GDP per capita | 2017 | Regionaldatenbank |
| Average disposable income per capita | 2017 | Regionaldatenbank |
| Share of population with college degree | 2018 | Regionaldatenbank |
| Employment share | 2019 | Regionaldatenbank |
| Share of employees in service sector | 2017 | Corona-Datenplattform |
| Share of non-German residents | 2019 | Regionaldatenbank |
| Share of population below age 18 | 2019 | Regionaldatenbank |
| Share of population aged 65 or above | 2019 | Regionaldatenbank |
| Border country indicators | 2021 | Any map of choice |
| Local policy restrictions | 2021 | Corona-Datenplattform |
| 2019 EU parliament election turnout & vote shares | 2019 | Regionaldatenbank |
| 2017 general election turnout & vote shares | 2017 | Regionaldatenbank |
| Civic associations per 100,000 population | 2008 | Franzen and Botzen (2011) |

### 2.C.4 Local Policy Stringency

Finally, to evaluate the role of county-level stringency of policy restrictions aimed to combat the pandemic, we obtain data from the infas360 Corona-Datenplattform (https://www.corona-datenplattform.de/dataset/massnahmen_oberkategorien_kreise) which indicates for 23 categories of possible restrictions (e.g. curfew, school closure, ...) whether they were in place in a certain county at a particular point in time. To construct a local policy stringency index, we sum up all 23 indicator variables and then normalize this index to range between 0 and 100, where 0 means that not a single restriction was in place, and 100 means that every single restriction was mandated by the local government. The 23 categories entail restrictions regarding: private gatherings, public gatherings, secondary schools, primary schools, daycare centers, indoor public events, outdoor public events, cultural institutions (museums, theaters, ...), retail and wholesale, gastronomy, services and craft, nightclubs and bars, hotels, indoor sports, outdoor sports, domestic travel, international travel, mask wearing, workplace, curfews, public transport, physical distancing, testing.

## Appendix 2.D Full Original Questionnaire (in German)

# Einführung

COVID-19 (ugs. auch Corona) ist eine Infektionskrankheit, die von einem neu entdeckten Coronavirus ausgelöst wird. Die aktuelle Coronavirus-Pandemie hat sicher auch Ihr Leben stark verändert. Bei dieser Umfrage geht es um Ihre Einstellungen zur Corona-Pandemie sowie Ihre allgemeinen Einstellungen im Leben.
Die Umfragedauer beträgt in etwa 20 Minuten. Für die mobile Version empfehlen wir Ihnen das Smartphone im Querformat zu nutzen.

**Datenschutz- und Einwillungserklärung**
In der folgenden Umfrage werden wir Ihnen Fragen stellen zu möglicherweise sensiblen Themen bezüglich Ihrer Gesundheit und Ihren politischen Einstellungen. Ziel der Studie ist es, die Einstellungen von Menschen in Deutschland zur aktuellen Corona- Pandemie besser zu verstehen.
Die Studie wird duchgeführt von Forschern der Universität Bonn. Verantwortlicher für die Erhebung und Auswertung der Daten ist Ximeng Fang. Die erhobenen Daten werden auf sicheren Servern der Universität Bonn aufgezeichnet und ausschließlich für Forschungszwecke verwendet. Dabei sind keine Rückschlüsse auf Ihre Person möglich, Sie bleiben also jederzeit vollständig anonym.
Bitte geben Sie an, ob Sie mit der Erhebung und Auswertung Ihrer Daten aus dieser Umfrage einverstanden sind. Sie können Ihre Einwilligungserklärung jederzeit widerrufen.

**Sind Sie mit der Erhebung und Auswertung Ihrer Daten aus dieser Umfrage einverstanden?**

- ☐ Ich bin einverstanden und möchte an dieser Umfrage teilnehmen.
- ☐ Ich bin nicht einverstanden.

Manchmal lesen Umfrageteilnehmer nicht sorgfältig die Fragen und klicken sich einfach nur schnell durch. Das kann die Ergebnisse von wissenschaftlichen Studien verfälschen. Geben Sie darum bei der folgenden Frage bitte den Wert 2 als Antwort an, um zu zeigen, dass Sie diese Umfrage sorgfältig beantworten.

**In Anbetracht dieses Problems: Wie stark sind Sie interessiert an den Auswirkungen von Verkehrslärm auf die Vogelpopulation in deutschen Städten?**

Überhaupt nicht interessiert ☐—☐—☐—☐—☐—☐—☐ Sehr stark interessiert

# Allgemeine Angaben zu Ihrer Person

Vielen Dank, dass Sie sich die Zeit nehmen für unsere Umfrage! Wir möchten Ihnen zunächst einige allgemeine Fragen zu Ihrer Person stellen.

Bitte geben Sie ihr Geburtsjahr an. ____________

Bitte geben Sie Ihr Geschlecht an.

- ☐ Weiblich
- ☐ Männlich
- ☐ Divers

Wie viel Zeit verbringen Sie in etwa auf sozialen Medien (z.B. Facebook, Instagram)?: ____________

Was für ein Smartphone benutzen Sie im Alltag?

- ☐ Android-Smartphone (z.B. Samsung, Huawei, …)
- ☐ iPhone (Apple)
- ☐ anderes Smartphone (z.B. Windows-Phone, Blackberry, …)
- ☐ Ich benutze kein Smartphone

In welchem Bundesland leben Sie? ____________

Was ist Ihre Postleitzahl? ____________

Seit welchem Jahr leben Sie an Ihrem aktuellen Wohnort?: ____________

Menschen haben verschiedene Ansichten über sich selbst und wie stark Sie sich mit ihrem Umfeld und dem Rest der Welt verbunden fühlen.
Wenn Sie sich einmal diese Liste ansehen, **wie stark fühlen Sie sich verbunden mit…**

| | Überhaupt nicht verbunden | Nicht sehr verbunden | Ein wenig verbunden | Ziemlich verbunden | Sehr verbunden |
|---|---|---|---|---|---|
| **Dem Ort oder der Stadt, in der Sie leben** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Der Region, in der Sie leben** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Deutschland** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Europa** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Der ganzen Welt** | ☐ | ☐ | ☐ | ☐ | ☐ |

Wie sehr stimmen Sie den folgenden Aussagen zu?
*Bitte bewerten Sie auf einer Skala von 1 bis 5. Der Wert 1 bedeutet: Stimme überhaupt nicht zu. Der Wert 5 bedeutet: Stimme voll und ganz zu.*

| | stimme überhaupt nicht zu | | | | stimme voll und ganz zu |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| **Ich bin finanziell negativ betroffen von der Corona-Pandemie.** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Ich bin in meinem persönlichen Leben stark eingeschränkt durch die Pandemie.** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Ich finde die Regierungsmaßnahmen gegen Corona überzogen.** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Insgesamt betrachtet hat Deutschland die Corona-Krise bisher gut bewältigt.** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Die Medien nehmen das Coronavirus viel zu ernst.** | ☐ | ☐ | ☐ | ☐ | ☐ |

Haben Sie sich in der Vergangenheit mit dem Coronavirus infiziert?

☐ Ja

☐ Nein

☐ keine Angabe

Kennen Sie persönlich jemanden, der sich mit dem Coronavirus infiziert hat?

☐ Ja

☐ Nein

☐ weiß nicht

Kennen Sie persönlich jemanden, der an Covid-19 gestorben ist?

☐ Ja

☐ Nein

☐ weiß nicht

Wann sind Sie das letzte Mal an Grippe erkrankt? ____________

Nun etwas ganz anderes. Unsere alltäglichen Handlungen werden davon beeinflusst, welche Grundüberzeugungen wir haben. Darüber ist in der Wissenschaft wenig bekannt. In den folgenden Seiten zeigen wir Ihnen einige unterschiedliche Eigenschaften, die eine Person haben kann. Wahrscheinlich werden manche Eigenschaften auf Sie persönlich mehr zutreffen als andere.

Bei allen Fragen geht es darum, wie Sie sich tatsächlich einschätzen, und nicht darum, wie Sie gerne sein würden. Bitte antworten Sie deshalb so ehrlich und spontan wie möglich. Es gibt keine richtigen oder falschen Antworten.

Versuchen Sie im Allgemeinen, Risiken zu vermeiden, oder sind Sie im Allgemeinen ein risikobereiter Mensch? *Bitte schätzen Sie sich persönlich ein, auf einer Skala von 0 bis 10. Der Wert 0 bedeutet: Überhaupt nicht bereit, Risiken einzugehen. Der Wert 10 bedeutet: Sehr bereit, Risiken einzugehen.*

überhaupt riskobereit □—□—□—□—□—□—□—□—□—□—□ sehr risikobereit

**Wir fragen Sie nun nach Ihrer Bereitschaft sich in einer bestimmten Art zu verhalten.**
*Bitte verwenden Sie wieder eine Skala von 0 bis 10. Der Wert 0 bedeutet: Überhaupt nicht bereit es zu tun. Der Wert 10 bedeutet: Sehr bereit es zu tun.*

Wie sehr wären Sie bereit auf etwas zu verzichten, das für Sie heute Nutzen bringt, um dadurch in Zukunft mehr zu profitieren?

überhaupt nicht bereit es zu tun □—□—□—□—□—□—□—□—□—□—□ sehr bereit es zu tun

Wie sehr wären Sie bereit jemanden zu bestrafen, der Sie unfair behandelt, selbst wenn dies für Sie negative Konsequenzen haben würde?

überhaupt nicht bereit es zu tun □—□—□—□—□—□—□—□—□—□—□ sehr bereit es zu tun

Wie sehr wären Sie bereit jemanden zu bestrafen, der andere unfair behandelt, selbst wenn dies für Sie Kosten verursachen würde?

überhaupt nicht bereit es zu tun □—□—□—□—□—□—□—□—□—□—□ sehr bereit es zu tun

Wie sehr wären Sie bereit für einen guten Zweck zu geben, ohne etwas als Gegenleistung zu erwarten.

überhaupt nicht bereit es zu tun □—□—□—□—□—□—□—□—□—□—□ sehr bereit es zu tun

**Wie gut beschreibt jede der nachfolgenden Aussagen Sie als Person?**
*Bitte verwenden Sie erneut eine Skala von 0 bis 10. Der Wert 0 bedeutet: Beschreibt mich überhaupt nicht. Der Wert 10 bedeutet: Beschreibt mich perfekt.*

Wenn mir jemanden einen Gefallen tut, bin ich bereit ihn zu erwidern.

beschreibt mich überhaupt nicht □—□—□—□—□—□—□—□—□—□ beschreibt mich perfekt

Wenn ich sehr ungerecht behandelt werde, räche ich mich bei der ersten Gelegenheit, selbst wenn Kosten entstehen um das zu tun.

beschreibt mich überhaupt nicht □—□—□—□—□—□—□—□—□—□ beschreibt mich perfekt

Ich vermute, dass Leute nur die besten Absichten haben.

beschreibt mich überhaupt nicht □—□—□—□—□—□—□—□—□—□ beschreibt mich perfekt

Ich bin gut in Mathematik.

beschreibt mich überhaupt nicht □—□—□—□—□—□—□—□—□—□ beschreibt mich perfekt

Ich neige dazu Aufgaben zu verschieben, auch wenn ich weiß, dass es besser wäre sie gleich zu tun.

beschreibt mich überhaupt nicht □—□—□—□—□—□—□—□—□—□ beschreibt mich perfekt

Stellen Sie sich die folgende Situation vor:

Heute haben Sie unerwartet 1000 Euro erhalten. Wie viel von dem Geld würden Sie einem guten Zweck spenden? ____________

**Woher beziehen Sie Ihre Nachrichten?**
Menschen nutzen unterschiedliche Quellen, um zu erfahren, was um sie herum und in der Welt passiert. Geben Sie bitte für jede der folgenden Quellen an, wie oft Sie diese nutzen:

| | Täglich | Wöchentlich | Monatlich | Seltener als monatlich | Niemals |
|---|---|---|---|---|---|
| **Zeitung** | □ | □ | □ | □ | □ |
| **Fernsehsendungen** | □ | □ | □ | □ | □ |
| **Radiosendungen** | □ | □ | □ | □ | □ |
| **Nachrichtenseiten im Internet** | □ | □ | □ | □ | □ |
| **Mobiltelefon (WhatsApp, Telegram, etc.)** | □ | □ | □ | □ | □ |
| **Social media (Facebook, Twitter, etc.)** | □ | □ | □ | □ | □ |
| **Gespräche mit Freunden, Kollegen und Bekannten** | □ | □ | □ | □ | □ |

# Ihre politischen Einstellungen

Wenn morgen Bundestagswahl wäre, welche Partei würden Sie dann wählen?

- ☐ CDU/CSU
- ☐ Bündnis '90/Die Grünen
- ☐ SPD
- ☐ AfD
- ☐ Die Linke
- ☐ FDP
- ☐ Sonstige

Würden Sie tatsächlich wählen gehen?

- ☐ Ja
- ☐ Nein
- ☐ Unentschlossen

**Wie zufrieden sind Sie damit, wie das politische System in Deutschland heutzutage funktioniert?**
*Bewerten Sie bitte auf einer Skala von 0 bis 10, auf der 0 für „überhaupt nicht zufrieden“ und 10 für „voll und ganz zufrieden“ steht.*

überhaupt nicht zufrieden ☐—☐—☐—☐—☐—☐—☐—☐—☐—☐—☐ voll und ganz zufrieden

**Wie sehr stimmen Sie den folgenden Aussagen zu?**

| | stimme überhaupt nicht zu | | | | stimme voll und ganz zu |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| **Es geschehen viele sehr wichtige Dinge in der Welt, über die die Öffentlichkeit nie informiert wird.** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Regierungsbehörden überwachen alle Bürger genau.** | ☐ | ☐ | ☐ | ☐ | ☐ |
| **Es gibt geheime Mächte, die die Welt steuern.** | ☐ | ☐ | ☐ | ☐ | ☐ |

Wie gut beschreibt jede der nachfolgenden Aussagen Sie als Person? *Bitte antworten Sie so ehrlich und spontan wie möglich, auf einer Skala von 1 bis 7. Der Wert 1 bedeutet: Trifft überhaupt nicht zu. Der Wert 7 bedeutet: Trifft voll und ganz zu.*

**Ich bin jemand, der …**

| | Trifft überhaupt nicht zu | | | | | | Trifft voll und ganz zu |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **… gründlich arbeitet.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… kommunikativ, gesprächig ist.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… manchmal etwas grob zu anderen ist.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… originell ist, neue Ideen einbringt.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… sich oft Sorgen macht.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

**Ich bin jemand, der …**

| | Trifft überhaupt nicht zu | | | | | | Trifft voll und ganz zu |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **… verzeihen kann.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… eher faul ist.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… aus sich herausgehen kann, gesellig ist.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… künstlerische Erfahrungen schätzt.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… leicht nervös wird.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

**Ich bin jemand, der …**

| | Trifft überhaupt nicht zu | | | | | | Trifft voll und ganz zu |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **… Aufgaben wirksam und effizient erledigt.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… zurückhaltend ist.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… rücksichtsvoll und freundlich mit anderen umgeht.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… eine lebhafte Phantasie, Vorstellungen hat.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… entspannt ist, mit Stress gut umgehen kann.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

**Wie sehr treffen die folgenden Aussagen auf Ihr eigenes Verhalten zu?**
*Bitte bewerten Sie erneut auf einer Skala von 1 bis 7. Der Wert 1 bedeutet: Trifft überhaupt nicht zu. Der Wert 7 bedeutet: Trifft voll und ganz zu.*

Ich halte mindestens 1,5m Abstand zu Mitmenschen.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Ich werde mich sozial isolieren, wenn ich Kontakt hatte mit einer infizierten Person.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Ich halte mich stets auf dem Laufenden über Neuigkeiten zur Corona-Pandemie.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Ich wasche bzw. desinfiziere regelmäßig meine Hände.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Ich werde mich gegen das Coronavirus impfen lassen, wenn ein Impfstoff verfügbar ist.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Ich huste und niese in die Ellbogenbeuge.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Ich trage in der Öffentlichkeit einen Mund-Nasen-Schutz.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Ich lüfte regelmäßig durch, wenn mehrere Personen einen Raum benutzen.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Ich vermeide soziale Kontakte soweit es geht.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Ich werde Mitmenschen darüber informieren, wenn ich mich mit Corona infiziert habe.

Trifft überhaupt nicht zu □—□—□—□—□—□—□ Trifft voll und ganz zu

Die Corona-Warn-App ist eine Smartphone-App, die Nutzer informieren soll, ob sie in Kontakt mit einer infizierten Person geraten sind und daraus ein erhöhtes Ansteckungsrisiko anzunehmen ist.

Haben Sie die Corona-Warn-App auf Ihrem aktuellen Mobiltelefon installiert?
- ☐ Ja
- ☐ Nein

Für den Fall, dass Sie positiv auf Corona getestet werden würden: Wie wahrscheinlich ist es, dass Sie dies über die Corona-Warn-App melden?

Sehr unwahrscheinlich ☐—☐—☐—☐—☐—☐—☐ Sehr wahrscheinlich

**Wie sehr stimmen Sie den folgenden Aussagen zur Corona-Warn-App zu?** *Bitten antworten Sie auf einer Skala von 1 bis 7. Der Wert 1 bedeutet: Stimme überhaupt nicht zu. Der Wert 7 bedeutet: Stimme voll und ganz zu.*

**Die Corona-Warn-App …**

... hilft dabei, die Ausbreitung von Corona in Deutschland zu verlangsamen.

Stimme überhaupt nicht zu ☐—☐—☐—☐—☐—☐—☐ Stimme voll und ganz zu

... hilft dabei, die Ausbreitung von Corona in meiner Stadt zu verlangsamen.

Stimme überhaupt nicht zu ☐—☐—☐—☐—☐—☐—☐ Stimme voll und ganz zu

... hat für mich persönlich keinen großen Nutzen.

Stimme überhaupt nicht zu ☐—☐—☐—☐—☐—☐—☐ Stimme voll und ganz zu

... ist datenschutzrechtlich bedenklich.

Stimme überhaupt nicht zu ☐—☐—☐—☐—☐—☐—☐ Stimme voll und ganz zu

... ist ein gutes Mittel um Infektionsketten nachzuverfolgen.

Stimme überhaupt nicht zu ☐—☐—☐—☐—☐—☐—☐ Stimme voll und ganz zu

... wird noch nicht von genügend Menschen genutzt.

Stimme überhaupt nicht zu ☐—☐—☐—☐—☐—☐—☐ Stimme voll und ganz zu

Die sogenannte 7-Tage-Inzidenz gibt die Zahl der Corona-Neuinfektionen (d.h. positiv auf Corona getestete Personen) pro 100.000 Einwohnern innerhalb der vergangenen sieben Tage an. Sie gilt als wichtige Kennziffer zur Einschätzung der aktuellen Corona-Lage (Hinweis: Stand 17.12. liegt der Wert für Gesamtdeutschland bei 179).

Bitte schätzen Sie die 7-Tage Inzidenzrate in Ihrer Stadt. ____________

**Wie besorgt sind Sie über die Möglichkeit, dass …**

| | Überhaupt nicht besorgt | | | | | Sehr besorgt |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **… Sie selbst an COVID-19 erkranken.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… Freunde oder Verwandte an COVID-19 erkranken.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| **… andere Menschen in Ihrer Umgebung an COVID-19 erkranken.** | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

Wie hoch schätzen Sie das Risiko ein, dass Sie sich innerhalb der nächsten 3 Monate mit COVID-19 anstecken?

Sehr unwahrscheinlich ☐—☐—☐—☐—☐—☐—☐ Sehr wahrscheinlich

## Spendenmöglichkeit

Wichtig: Das folgende Szenario wird mit 25% Wahrscheinlichkeit tatsächlich umgesetzt. Sie sollten also sorgfältig überlegen, was Sie tun wollen. Es handelt sich womöglich um reale Geldbeträge.

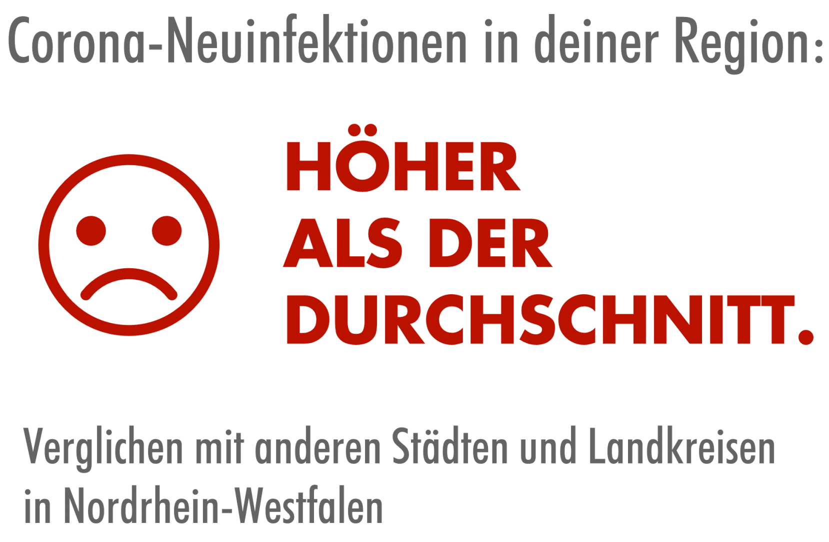
Ihnen steht ein Geldbetrag in Höhe von 1 Euro zur Verfügung. Sie können frei entscheiden, welchen Anteil davon Sie spenden wollen, und welchen Anteil Sie für sich selbst behalten. Ihre Spende wird für eine Online-Werbekampagne auf sozialen Medien eingesetzt, die mehr Menschen (u.a. in Ihrer Region) zur Nutzung der Corona-Warn-App ermutigt. In der Vergangenheit hat sich gezeigt, dass 1 Corona-Warn-App-Installation durchschnittlich knapp 50 Cent Werbeausgaben entspricht. Den Teil des Geldbetrags, den Sie nicht spenden, erhalten Sie als zusätzliche Entlohnung in Form von Panelpunkten.

Am Ende der Umfrage lost ein Zufallsgenerator aus, ob die Spende und die zusätzliche Entlohnung tatsächlich ausgezahlt werden. Bitte bewegen Sie den Schieberegler, um über Ihr Budget zu entscheiden:

**Welchen Betrag möchten Sie spenden?** Ihre Spende beträgt ______ Euro.

**Bitte schauen Sie sich das folgende Video an.**

In Kiel wurden zuletzt 117,1 Corona-Neuinfektionen pro 100.000 Einwohnern in 7 Tagen gemeldet, das ist 27% höher als der landesweite Durchschnitt. (Quelle: Robert-Koch-Institut, Stand 17.12.)



Sie nähern sich nun dem Ende des Fragebogens. Einige der folgenden Fragen werden Ihnen bekannt vorkommen. Wundern Sie sich nicht, das ist ein ganz normaler Teil der Umfrage. Bitte beantworten Sie diese Fragen genauso sorgfältig und gewissenhaft wie die vorherigen.

Für den Fall, dass Sie positiv auf Corona getestet werden würden: Wie wahrscheinlich ist es, dass Sie dies über die Corona-Warn-App melden?

Sehr unwahrscheinlich □—□—□—□—□—□—□ Sehr wahrscheinlich

**Wie sehr stimmen Sie den folgenden Aussagen zur Corona-Warn-App zu?** *Bitten antworten Sie auf einer Skala von 1 bis 7. Der Wert 1 bedeutet: Stimme überhaupt nicht zu. Der Wert 7 bedeutet: Stimme voll und ganz zu.*

**Die Corona-Warn-App …**

... hilft dabei, die Ausbreitung von Corona in Deutschland zu verlangsamen.

Stimme überhaupt nicht zu □—□—□—□—□—□—□ Stimme voll und ganz zu

... hilft dabei, die Ausbreitung von Corona in meiner Stadt zu verlangsamen.

Stimme überhaupt nicht zu □—□—□—□—□—□—□ Stimme voll und ganz zu

... hat für mich persönlich keinen großen Nutzen.

Stimme überhaupt nicht zu □—□—□—□—□—□—□ Stimme voll und ganz zu

... ist datenschutzrechtlich bedenklich.

Stimme überhaupt nicht zu □—□—□—□—□—□—□ Stimme voll und ganz zu

... ist ein gutes Mittel um Infektionsketten nachzuverfolgen.

Stimme überhaupt nicht zu □—□—□—□—□—□—□ Stimme voll und ganz zu

... wird noch nicht von genügend Menschen genutzt.

Stimme überhaupt nicht zu □—□—□—□—□—□—□ Stimme voll und ganz zu

Die sogenannte 7-Tage-Inzidenz gibt die Zahl der Corona-Neuinfektionen (d.h. positiv auf Corona getestete Personen) pro 100.000 Einwohnern innerhalb der vergangenen sieben Tage an (Hinweis: Stand 17.12. liegt der Wert für Gesamtdeutschland bei 179).

**Bitte schätzen Sie die 7-Tage Inzidenzrate in Ihrer Stadt.** ____________

**Wie besorgt sind Sie über die Möglichkeit, dass …**

| | Überhaupt nicht besorgt | | | | | Sehr besorgt |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **… Sie selbst an COVID-19 erkranken.** | □ | □ | □ | □ | □ | □ | □ |
| **… Freunde oder Verwandte an COVID-19 erkranken.** | □ | □ | □ | □ | □ | □ | □ |
| **… andere Menschen in Ihrer Umgebung an COVID-19 erkranken.** | □ | □ | □ | □ | □ | □ | □ |

Wie hoch schätzen Sie das Risiko ein, dass Sie sich innerhalb der nächsten 3 Monate mit COVID-19 anstecken?

Sehr unwahrscheinlich □—□—□—□—□—□—□ Sehr wahrscheinlich

## Erneute Spendenmöglichkeit

Sie stehen erneut der gleichen Spendenentscheidung gegenüber wie zuvor. Das Szenario auf dieser Seite wird wieder mit 25% Wahrscheinlichkeit tatsächlich umgesetzt. Maximal eine Ihrer beiden Spendenentscheidungen wird zufällig ausgewählt, nie jedoch beide gleichzeitig.

Zur Erinnerung: Ihnen steht ein Geldbetrag in Höhe von 1 Euro zur Verfügung. Sie können frei entscheiden, welchen Anteil davon Sie spenden wollen für eine Online-Werbekampagne zur Corona-Warn-App. In der Vergangenheit hat sich gezeigt, dass 1 Corona- Warn-App-Installation durchschnittlich knapp 50 Cent Werbeausgaben entspricht. Den Teil des Geldbetrags, den Sie nicht spenden, erhalten Sie als zusätzliche Entlohnung in Form von Panelpunkten.

Am Ende der Umfrage lost ein Zufallsgenerator aus, ob die Spende und die zusätzliche Entlohnung tatsächlich ausgezahlt werden. Bitte bewegen Sie den Schieberegler, um über Ihr Budget zu entscheiden:

**Welchen Betrag möchten Sie spenden?** Ihre Spende beträgt ______ Euro.

# Finale Angaben zu Ihrer Person

Fast geschafft! Zum Abschluss der Umfrage möchten wir Sie gerne noch um einige letzten Angaben zu Ihrer Person bitten.

Welches ist Ihr höchster Bildungsabschluss?

- ☐ Schule ohne Abschluss verlassen
- ☐ Volks-/Hauptschulabschluss (DDR: 8. Klasse)
- ☐ Realschulabschluss/Mittlere Reife (DDR: 10. Klasse)
- ☐ Fachhochschulreife (Abschluss einer Fachoberschule)
- ☐ Abitur/Hochschulreife
- ☐ Fachhochschule (früher: Ingenieurschule, Lehrerbildung, DDR: Ingenieur und Fachschulabschluss)
- ☐ Universitäts-, Hochschulabschluss
- ☐ Promotion
- ☐ Sonstiger Bildungsabschluss

**Wie viele Personen leben in Ihrem Haushalt (d.h. zusammen wohnende und wirtschaftende Einheit)? ____________**

Wie viele Kinder haben Sie? ____________

Was beschreibt Ihren aktuellen Erwerbsstatus am besten?

- ☐ Vollzeit angestellt
- ☐ Teilzeit angestellt
- ☐ Selbstständig
- ☐ im Studium/in Ausbildung
- ☐ Nicht erwerbstätig, nicht in Ausbildung

Welches ist Ihr höchster Bildungsabschluss?

- ☐ unter 900 €
- ☐ 900 € bis unter 1.300 €
- ☐ 1.300 € bis unter 1.500 €
- ☐ 1.500 € bis unter 2.000 €
- ☐ 2.000 € bis unter 2.600 €
- ☐ 2.600 € bis unter 3.200 €
- ☐ 3.200 € bis unter 4.500 €
- ☐ 4.500 € bis unter 6.000 €
- ☐ 6.000 € und mehr
- ☐ keine Angabe

## Vielen Dank für Ihre Teilnahme an dieser Umfrage!

Sie haben nun das Ende des Fragebogens erreicht. Sie konnten im Laufe der Umfrage zwei Mal über einen Geldbetrag von jeweils 1 Euro entscheiden. Von einem Zufallsgenerator wurde ausgelost, ob eines dieser Szenarien tatsächlich umgesetzt wird. Folgende Auszahlung wurde für Sie bestimmt:

Sie spenden ______ Euro an eine Online-Werbekampagne für die Corona-Warn-App.

Sie erhalten ______ Euro als zusätzliche Entlohnung in Form von Panelpunkten. Bitte beachten Sie, dass es 4 bis 6 Wochen dauern kann, bis diese Ihrem Konto gutgeschrieben werden..

Diese Umfrage wurde durchgeführt von Forschern der Universität Bonn. Ziel der Studie ist es, mehr über die Einstellungen zur Corona-Pandemie in Deutschland zu erfahren. Dabei ging es unter anderem auch um die Bereitschaft zur Nutzung der Corona-Warn-App. Für weitere Informationen zur App haben wir für Sie im Folgenden einige Antworten auf häufig gestellte Fragen (FAQs) zusammengestellt.

Sobald Sie fertig sind, klicken Sie bitte auf Umfrage abschließen.

## References


**Abelson, Robert P, Elliot Aronson, William J McGuire, Theodore M Newcomb, Milton J Rosenberg, and Percy H Tannenbaum.** 1968. *Theories of Cognitive Consistency: A Sourcebook.* Edited by Robert P Abelson, Elliot Aronson, William J McGuire, Theodore M Newcomb, Milton J Rosenberg, and Percy H Tannenbaum. Chicago: Rand-McNally. [110]

**Albrecht, Felix, Sebastian Kube, and Christian Traxler.** 2018. "Cooperation and Norm Enforcement - The Individual-Level Perspective." *Journal of Public Economics* 165: 1–16. https://doi.org/10.1016/j.jpubeco.2018.06.010. [108, 111]

**Alesina, Alberto, and Nicola Fuchs-Schündeln.** 2007. "Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences." *American Economic Review* 97 (4): 1507–28. [121]

**Alfaro, Laura, Ester Faia, Nora Lamersdorf, and Farzad Saidi.** 2021a. "Health Externalities and Policy: The Role of Social Preferences." *ECONtribute Discussion Paper No. 109.* [108, 111]

**Alfaro, Laura, Ester Faia, Nora Lamersdorf, and Farzad Saidi.** 2021b. "Social Interactions in Pandemic." *ECONtribute Discussion Paper No. 110,* https://doi.org/10.3386/w27134. [108, 109, 111, 122]

**Andre, Peter, Teodora Boneva, Felix Chopra, and Armin Falk.** 2021. "Fighting Climate Change: The Role of Norms, Preferences, and Moral Values." *IZA Discussion Paper 14518.* [108]

**Baron, Reuben M, and David A Kenny.** 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51 (6): 1173–82. https://doi.org/https://psycnet.apa.org/doi/10.1037/0022-3514.51.6.1173. [119]

**Barrios, John M., Efraim Benmelech, Yael V. Hochberg, Paola Sapienza, and Luigi Zingales.** 2021. "Civic Capital and Social Distancing during the Covid-19 Pandemic." *Journal of Public Economics* 193: 104310. https://doi.org/10.1016/j.jpubeco.2020.104310. [108]

**Barsbai, Toman, Dieter Lukas, and Andreas Pondorfer.** 2021. "Local Convergence of Behavior across Species." *Science* 371 (6526): 292–95. https://doi.org/10.5281/ZENODO.4159697. [124]

**Bartscher, Alina, Sebastian Seitz, Michaele Slotwinski, Nils Wehrhöfer, and Sebastian Siegloch.** 2020. "Social Capital and the Spread of Covid-19: Insights from European Countries." *ZEW Discussion Paper 20-023.* [108]

**Batson, Daniel C., and Adam A Powell.** 2003. "Altruism and Prosocial Behavior." In *Handbook of Psychology: Personality and Social Psychology,* edited by Theodore Millon and Melvin J. Lerner, 463–84. Wiley. https://doi.org/10.1002/0471264385.wei0519. [110]

**Battaglia, Michael P, David C Hoaglin, and Martin R Frankel.** 2009. "Practical Considerations in Raking Survey Data." *Survey Practice* 2 (5): 2953. [112, 159]

**Bauch, Chris T, and David JD Earn.** 2004. "Vaccination and the Theory of Games." *Proceedings of the National Academy of Sciences* 101 (36): 13391–94. [109]

**Bauer, Michal, Christopher Blattman, Julie Chytilová, Joseph Henrich, Edward Miguel, and Tamar Mitts.** 2016. "Can War Foster Cooperation?" *Journal of Economic Perspectives* 30 (3): 249–74. [123]

**Becker, Anke, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse.** 2012. "The Relationship between Economic Preferences and Psychological Personality Measures." *Annual Review of Economics* 4 (1): 453–78. https://doi.org/10.1146/annurev-economics-080511-110922. [115, 158]

**Becker, Sascha O, Lukas Mergele, and Ludger Woessmann.** 2020. "The Separation and Reunification of Germany: Rethinking a Natural Experiment Interpretation of the Enduring Effects of Communism." *Journal of Economic Perspectives* 34 (2): 143–71. [121]

**Bowles, Samuel, and Herbert Gintis.** 2011. *A Cooperative Species: Human Reciprocity and Its Evolution.* Princeton University Press. https://doi.org/doi:10.1515/9781400838837. [110]

**Branas-Garza, Pablo, Diego Andrés Jorrat, Antonio Alfonso, Antonio M. Espin, Teresa García, and Jaromir Kovarik.** 2020. "Exposure to the Covid-19 Pandemic and Generosity." *MPRA Paper No. 103389,* https://doi.org/10.31234/osf.io/6ktuz. [123]

**Brosig-Koch, Jeannette, Christoph Helbach, Axel Ockenfels, and Joachim Weimann.** 2011. "Still Different after All These Years: Solidarity Behavior in East and West Germany." *Journal of Public Economics* 95 (11-12): 1373–76. [121]

**Burks, Stephen, Jeffrey Carpenter, and Lorenz Goette.** 2009. "Performance Pay and Worker Cooperation: Evidence from an Artefactual Field Experiment." *Journal of Economic Behavior & Organization* 70 (3): 458–69. [108]

**Burks, Stephen, Jeffrey Carpenter, Lorenz Götte, and Aldo Rustichini.** 2012. "Which Measures of Time Preference Best Predict Outcomes: Evidence from a Large-Scale Field Experiment." *Journal of Economic Behavior & Organization* 84 (1): 308–20. [111]

**Caicedo, Felipe Valencia, Thomas Dohmen, and Andreas Pondorfer.** 2021. "Religion and Prosociality across the Globe." *Working Paper.* [124]

**Cameron, A. Colin, Jonah Gelbach, and Douglas Miller.** 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90 (3): 414–27. https://doi.org/10.3386/t0344. [117]

**Campos-Mercade, Pol, Armando N. Meier, Florian H. Schneider, and Erik Wengström.** 2021. "Prosociality Predicts Health Behaviors during the COVID-19 Pandemic." *Journal of Public Economics* 195: 104367. https://doi.org/10.1016/j.jpubeco.2021.104367. [108, 122, 123]

**Cappelen, Alexander W., Ranveig Falch, Erik Ø. Sørensen, and Bertil Tungodden.** 2021. "Solidarity and Fairness in Times of Crisis." *Journal of Economic Behavior & Organization* 186: 1–11. https://doi.org/10.1016/j.jebo.2021.03.017. [123]

**Carlsson, Fredrik, Olof Johansson-Stenman, and Pham Khanh Nam.** 2014. "Social Preferences Are Stable over Long Periods of Time." *Journal of Public Economics* 117: 104–14. https://doi.org/10.1016/j.jpubeco.2014.05.009. [123]

**Casoria, Fortuna, Fabio Galeotti, and Marie Claire Villeval.** 2021. "Perceived Social Norm and Behavior Quickly Adjusted to Legal Changes during the COVID-19 Pandemic." *Journal of Economic Behavior & Organization* 190: 54–65. [122]

**Chan, Ho Fai, Ahmed Skali, David A. Savage, David Stadelmann, and Benno Torgler.** 2020. "Risk Attitudes and Human Mobility during the COVID-19 Pandemic." *Scientific Reports* 10 (1): 19931. https://doi.org/10.1038/s41598-020-76763-2. [111]

**Cohn, Alain, Michel André Maréchal, David Tannenbaum, and Christian Lukas Zünd.** 2019. "Civic Honesty around the Globe." *Science* 365 (6448): 70–73. https://doi.org/10.1126/science.aau8712. [124]

**de Oliveira, Angela C.M., Tammy C.M. Leonard, Kerem Shuval, Celette Sugg Skinner, Catherine Eckel, and James C. Murdoch.** 2016. "Economic Preferences and Obesity among a Low-Income African American Community." *Journal of Economic Behavior & Organization* 131: 196–208. https://doi.org/10.1016/j.jebo.2015.11.002. [111]

**Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde.** 2008. "Homo Reciprocans: Survey Evidence on Behavioural Outcomes." *Economic Journal* 119 (536): 592–612. https://doi.org/10.26481/umaror.2008007. [108]

**Durante, Ruben, Luigi Guiso, and Giorgio Gulino.** 2021. "Asocial Capital: Civic Culture and Social Distancing during COVID-19." *Journal of Public Economics* 194: 104342. https://doi.org/10.2139/ssrn.3611606. [108]

**Epper, Thomas, Ernst Fehr, and Julien Senn.** 2020. "Other-Regarding Preferences and Redistributive Politics." *University of Zurich Working Paper No. 339,* https://doi.org/10.2139/ssrn.3526809. [108]

**Exadaktylos, Filippos, Antonio M Espín, and Pablo Branas-Garza.** 2013. "Experimental Subjects Are Not Different." *Scientific Reports* 3 (1): 1–6. [112]

**Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde.** 2018. "Global Evidence on Economic Preferences." *Quarterly Journal of Economics* 133 (4): 1645–92. https://doi.org/10.1093/qje/qjy013. [108, 112, 114, 124, 145]

**Falk, Armin, Anke Becker, Thomas J. Dohmen, David Huffman, and Uwe Sunde.** 2016. "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences." *IZA Discussion Paper No. 9674,* https://doi.org/10.2139/ssrn.2725874. [109, 112, 113, 145]

**Falk, Armin, Stephan Meier, and Christian Zehnder.** 2013. "Do Lab Experiments Misrepresent Social Preferences? The Case of Self-Selected Student Samples." *Journal of the European Economic Association* 11 (4): 839–52. https://doi.org/10.1111/jeea.12019. eprint: https://academic.oup.com/jeea/article-pdf/11/4/839/10461629/jeea0839.pdf. [112]

**Farboodi, Maryam, Gregor Jarosch, and Robert Shimer.** 2021. "Internal and External Effects of Social Distancing in a Pandemic." *Journal of Economic Theory* 196: 105293. [108, 109, 111]

**Fehr, Ernst, and Urs Fischbacher.** 2003. "The Nature of Human Altruism." *Nature* 425: 785–91. https://doi.org/10.1046/j.0013-0427.2003.00027.x. [108]

**Fehr, Ernst, and Simon Gächter.** 2002. "Altruistic Punishment in Humans." *Nature* 415 (6868): 137–40. https://doi.org/10.1038/415137a. [108, 110]

**Fehr, Ernst, and Ivo Schurtenberger.** 2018. "Normative Foundations of Human Cooperation." *Nature Human Behaviour* 2 (7): 458–68. https://doi.org/10.1038/s41562-018-0385-5. [108]

**Fenichel, Eli P, Carlos Castillo-Chavez, M Graziano Ceddia, Gerardo Chowell, Paula A Gonzalez Parra, Graham J Hickling, Garth Holloway, Richard Horan, Benjamin Morin, Charles Perrings, et al.** 2011. "Adaptive Human Behavior in Epidemiological Models." *Proceedings of the National Academy of Sciences* 108 (15): 6306–11. [109]

**Festinger, Leon.** 1957. *A Theory of Cognitive Dissonance.* Stanford University Press. [110]

**Fischbacher, Urs, and Simon Gächter.** 2010. "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments." *American Economic Review* 100 (1): 541–56. [108]

**Franzen, Axel, and Katrin Botzen.** 2011. "Vereine in Deutschland Und Ihr Beitrag Zum Wohlstand Der Regionen." *Soziale Welt* 62: 391–413. [134, 160]

**Frondel, Manuel, Daniel Osberghaus, and Stephan Sommer.** 2021. "Corona and the Stability of Personal Traits and Preferences: Evidence from Germany." *ZEW Discussion Paper 21-029,* https://doi.org/10.2139/ssrn.3820484. [123]

**Fuhrmann-Riebel, Hanna, Ben D'Exelle, and Arjan Verschoor.** 2021. "The Role of Preferences for Pro-Environmental Behaviour among Urban Middle Class Households in Peru." *Ecological Economics* 180: 106850. https://doi.org/10.1016/j.ecolecon.2020.106850. [108]

**Galbiati, Roberto, Emeric Henry, Nicolas Jacquemet, and Max Lobeck.** 2021. “How Laws Affect the Perception of Norms: Empirical Evidence from the Lockdown.” *PloS ONE* 16 (9): e0256624. [122]

**Gerlitz, Jean-Yves, and Jürgen Schupp.** 2005. “Zur Erhebung Der Big-Five-Basierten Persoenlichkeitsmerkmale Im SOEP.” *DIW Research Notes* 4: 2005. [112, 157]

**Gollwitzer, Anton, Killian Mcloughlin, Martel, Cameron: Marshall, Julia, Johanna M. Höhs, and John A. Bargh.** 2021. “Linking Self-Reported Social Distancing to Real-World Behavior during the COVID-19 Pandemic.” *Preprint.* [109]

**Harper, Craig A., Liam P. Satchell, Dean Fido, and Robert D. Latzman.** 2020. “Functional Fear Predicts Public Health Compliance in the COVID-19 Pandemic.” *International Journal of Mental Health and Addiction,* 1–14. https://doi.org/10.1007/s11469-020-00281-5. [122]

**Heider, Fritz.** 1958. *The Psychology of Interpersonal Relations.* Wiley. [110]

**Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, et al.** 2006. “Costly Punishment across Human Societies.” *Science* 312 (5781): 1767–70. https://doi.org/10.1126/science.1127333. [124]

**Jagelka, Tomas.** 2020. “Are Economists’ Preferences Psychologists’ Personality Traits? A Structural Approach.” *Working Paper.* [115]

**Jensen, Ulrich Thy.** 2020. “Is Self-Reported Social Distancing Susceptible to Social Desirability Bias? Using the Crosswise Model to Elicit Sensitive Behaviors.” *Journal of Behavioral Public Administration* 3 (2): 1–11. https://doi.org/10.30636/jbpa.32.182. [109]

**Jones, Callum, Thomas Philippon, and Venky Venkateswaran.** 2021. “Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home.” *Review of Financial Studies* 34 (11): 5188–223. [109]

**Keeling, Matt J, and Pejman Rohani.** 2011. *Modeling Infectious Diseases in Humans and Animals.* Princeton university press. [109]

**Kermack, William Ogilvy, and Anderson G McKendrick.** 1927. “A Contribution to the Mathematical Theory of Epidemics.” *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character* 115 (772): 700–721. [109]

**Khwaja, Ahmed, Frank Sloan, and Martin Salm.** 2006. “Evidence on Preferences and Subjective Beliefs of Risk Takers: The Case of Smokers.” *International Journal of Industrial Organization* 24 (4): 667–82. https://doi.org/10.1016/j.ijindorg.2005.10.001. [111]

**Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk.** 2020. “The Formation of Prosociality: Causal Evidence on the Role of Social Environment.” *Journal of Political Economy* 128 (2): 434–67. https://doi.org/10.1086/704386. [124]

**Kosse, Fabian, and Michela M. Tincani.** 2020. “Prosociality Predicts Labor Market Success around the World.” *Nature Communications* 11 (1): 5298. https://doi.org/10.1038/s41467-020-19007-1. [108, 112]

**Lades, Leonhard K., Kate Laffan, and Till O. Weber.** 2021. “Do Economic Preferences Predict Pro-Environmental Behaviour?” *Ecological Economics* 183: 106977. https://doi.org/10.1016/j.ecolecon.2021.106977. [108]

**Makridis, Christos A., and Cary Wu.** 2021. “How Social Capital Helps Communities Weather the COVID-19 Pandemic.” *PloS ONE* 16 (1): e0245135. https://doi.org/10.1371/journal.pone.0245135. [108]

**Müller, Stephan, and Holger A. Rau.** 2021. “Economic Preferences and Compliance in the Social Stress Test of the COVID-19 Crisis.” *Journal of Public Economics* 194: 104322. https://doi.org/10.1016/j.jpubeco.2020.104322. [108, 122]

**Nettle, Daniel, Agathe Colléony, and Maria Cockerill.** 2011. "Variation in Cooperative Behaviour within a Single City." *PloS ONE* 6 (10): e26922. https://doi.org/10.1371/journal.pone.0026922.t001. [124]

**Nikolov, Plamen, Andreas Pape, Ozlem Tonguc, and Charlotte Williams.** 2020. "Predictors of Social Distancing and Mask-Wearing Behavior: Panel Survey in Seven U.S. States." *IZA Discussion Paper 13745.* [114]

**Ostrom, Elinor.** 1998. "A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997." *American Political Science Review* 92 (1): 1–22. [107]

**Petherick, Anna, Rafael Goldszmidt, Eduardo B. Andrade, Rodrigo Furst, Thomas Hale, Annalena Pott, and Andrew Wood.** 2021. "A Worldwide Assessment of Changes in Adherence to COVID-19 Protective Behaviours and Hypothesized Pandemic Fatigue." *Nature Human Behaviour,* https://doi.org/10.1038/s41562-021-01181-x. [122]

**Quaas, Martin F., Japer N. Meya, Hanna Schenk, Björn Bos, Moritz A. Drupp, and Till Requate.** 2021. "The Social Cost of Contacts: Theory and Evidence for the First Wave of the COVID-19 Pandemic in Germany." *PLoS ONE* 16 (3): e0248288. [108, 109, 111]

**Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb.** 2019. "Fast and Wild: Bootstrap Inference in Stata Using Boottest." *Stata Journal* 19 (1): 4–60. https://doi.org/10.1177/1536867X19830877. [117]

**Shachat, Jason, Matthew J Walker, and Lijia Wei.** 2021. "How the Onset of the Covid-19 Pandemic Impacted pro-Social Behaviour and Individual Preferences: Experimental Evidence from China." *Journal of Economic Behavior & Organization* 190: 480–94. [123]

**Sutter, Matthias, Martin G. Kocher, Daniela Glätzle-Rützler, and Stefan T. Trautmann.** 2013. "Impatience and Uncertainty: Experimental Decisions Predict Adolescents' Field Behavior." *American Economic Review* 103 (1): 510–31. https://doi.org/10.1257/aer.103.1.510. [111]

**Torgler, Benno.** 2002. "Does Culture Matter? Tax Morale in an East-West-German Comparison." *FinanzArchiv / Public Finance Analysis* 59 (4): 504–28. [121]

**Tversky, Amos, and Daniel Kahneman.** 1973. "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology* 5 (2): 207–32. [123]

**van Hulsen, Merel, Kirsten I. M. Rohde, and Job van Exel.** 2020. "Inter-Temporal and Social Preferences Predict Compliance in a Social Dilemma: An Application in the Context of COVID-19." *Working Paper,* https://doi.org/10.2139/ssrn.3665978. [108]

**Volk, Stefan, Christian Thöni, and Winfried Ruigrok.** 2012. "Temporal Stability and Psychological Foundations of Cooperation Preferences." *Journal of Economic Behavior & Organization* 81 (2): 664–76. [123]

**Zettler, Ingo, Christoph Schild, Lau Lilleholt, Lara Kroencke, Till Utesch, Morten Moshagen, Robert Böhm, Mitja D Back, and Katharina Geukes.** 2022. "The Role of Personality in COVID-19-related Perceptions, Evaluations, and Behaviors: Findings across Five Samples, Nine Traits, and 17 Criteria." *Social Psychological and Personality Science* 13 (1): 299–310. [115]

# Chapter 3

# The Effect of Task (Mis)Matching and Self-Selection on Intrinsic Motivation and Performance*

*Joint with Jonas Radbruch and Sebastian Schaube*

## 3.1 Introduction

Employers determine crucial features of their employees' work environment, such as the task they have to work on, their schedule, where they work, and with whom they work together. As these (non-monetary) features can significantly affect workers' behavior and performance (e.g., Cassar and Meier, 2018), a well-designed work environment is important. To understand how the work environment impacts workers' behavior and performance, researchers typically compare workers' behavior across different settings. However, such comparisons abstract from the fact that different workers may prosper under different circumstances. In other words, having a particular worker work in a particular environment may entail a "match value" because different work environments cater to different workers' strengths or preferences.

In this paper, we focus on a core aspect of work environments—namely, which task individuals work on—and study how the match between workers' task and task preferences affects their performance. As firms and organizations usually have to allocate a set of tasks to a set of workers, they may fail to realize substantial match value if they implement an unsuitable allocation that forces workers to work

* We would like to thank Thomas Dohmen, Ximeng Fang, Luca Henkel, Chui-Yee Ho, Sebastian Kube, Louis Strang, as well as seminar participants in Bonn for helpful comments and suggestions. Financial support by the Bonn Graduate School of Economics and IZA is gratefully acknowledged. An early version of this research project with the same title was submitted as a master's thesis by Timo Freyer.

on tasks they are neither good at nor (intrinsically) motivated to work on.[1] A lack of intrinsic motivation may be particularly problematic in the common case where employers are devoid of comprehensive and contractible measures to incentivize effort or have other good reasons to refrain from high-powered incentives even though their employees' output is sufficiently observable (Holmström, 2017). For this reason, we study a context in which workers are remunerated with flat pay.[2] In that case, workers may prefer one task over another because they find that task more worthwhile or enjoyable. Task preferences may also be shaped by a preference for being productive if differences in task-specific ability enable workers to produce more output on one task than the other at a given level of effort. Consequently, workers may perform better on their preferred task due to a higher task-specific ability or intrinsic motivation. Thus, as a first step, we investigate how task-allocation procedures that respect workers' task preferences affect their behavior and performance relative to a task-allocation procedure that always has them working on their non-preferred task.

To implement a task allocation that respects workers' preferences, employers can either take preferences into account when they assign tasks or delegate the task choice to the workers themselves, which may be more practical given that preferences are typically unknown and may first need to be elicited. The delegation of decision rights may not only help to align workers' tasks and task preferences, but has also been hypothesized to independently increase workers' intrinsic motivation due to increased feelings of autonomy (Ryan and Deci, 1985, 2000). Hence, as a second step, we test whether it makes a difference if workers are assigned or can self-select their preferred task.

To shed light on the behavioral consequences of different task-allocation rules, we conducted a real-effort online experiment where subjects worked for a fixed payment either on a task that asked them to evaluate tweets or on a data-entry task. A crucial aspect of the experimental design is that the tasks generated outside value and, therefore, admitted intrinsic motivation: subjects contributed to creating data sets that can be used for research purposes. This setup allows us to study worker behavior in an environment where intrinsic motivation is a key determinant of effort provision. The experiment consisted of two parts. In the first part, subjects were briefly introduced to both tasks before we elicited their binary task preferences. In the second part, we varied the task-allocation procedure across subjects. Knowledge of all subjects' preferences allowed us to exogenously vary

1. Intrinsic motivation is considered to originate from a "desire to perform the task for its own sake" (Bénabou and Tirole, 2003, p. 490). We will follow this understanding and consider effort to be intrinsically motivated if it does not contribute to obtaining a separable outcome as, for instance, increased pay (Ryan and Deci, 2000).

2. If available, performance incentives offer a theoretically appealing solution to the task-allocation problem: sufficiently steep incentives induce workers to prefer the task they are more able at and provide an extrinsic motivation to exert effort.

whether subjects were assigned their preferred or non-preferred task (treatments Assigned Preferred and Assigned Nonpreferred). In a third treatment, subjects had the opportunity to actively self-select their preferred task (treatment Self-Selected). After they were assigned or had self-selected a task, subjects were asked to work for at least 10-15 minutes but could decide to work longer. While treatments Self-Selected and Assigned Preferred both change the allocation of tasks to workers relative to treatment Assigned Nonpreferred, a comparison between the two identifies the effect of allowing subjects to self-select their preferred task instead of assigning them their preferred task.

We find that the match between tasks and task preferences strongly affects performance. Subjects in Assigned Preferred produced on average 47% higher output than subjects in Assigned Nonpreferred. Similarly, subjects in Self-Selected produced 55% more output on average than subjects in Assigned Nonpreferred. If subjects worked on their preferred task, they on average worked longer and were more productive—i.e., produced more output per minute worked. Whereas the increase in productivity could result from ability-sorting, longer working times indicate that the treatments increased subjects' effort. The performance of subjects in Assigned Preferred and Self-Selected did not differ significantly. Evidence from the post-experimental questionnaire suggests a potential explanation for this: subjects in Assigned Preferred already reported a high degree of perceived choice, which was significantly but only mildly increased by explicit self-selection. In sum, we observe that assigning subjects their preferred task and giving them autonomy to self-select that task similarly increased subjects' performance relative to the case where they were assigned their non-preferred task.

Many studies document how (non-monetary) features of the work environment can affect workers' behavior and performance (e.g., Cassar and Meier, 2018). To that end, researchers have typically compared workers' behavior across different work environments, including working from home vs. at the workplace (Dutcher, 2012; Bloom et al., 2015), self-managed working time arrangements vs. fixed schedules (Beckmann, Cornelissen, and Kräkel, 2017), and working alone vs. in the presence of peers (Falk and Ichino, 2006; Mas and Moretti, 2009). It is also well known that firms can design work environments to screen workers, which has been primarily studied in the context of incentive schemes (Stiglitz, 1975; Salop and Salop, 1976; Lazear, 2000; Niederle and Vesterlund, 2007; Dohmen and Falk, 2011). Our focus is neither on selection into nor on differences in behavior across work environments. Instead, we study how the alignment of workplace features with workers' preferences affects performance. Our results show that a task-allocation procedure that respects preferences can significantly improve workers' performance by increasing both productivity and labor supply.

By allowing subjects to actively self-select their task, we provide evidence on the effects of (workplace) autonomy on performance. Prior theoretical work in economics has stressed that delegating decision rights to agents can on the one hand lead to more efficient use of asymmetric information, induce incentive effects, and ease participation constraints, while on the other hand principals give up control and autonomy might be abused (Aghion and Tirole, 1997; Dessein, 2002; Bloom and Van Reenen, 2011; Fehr, Herz, and Wilkening, 2013; Aghion, Bloom, and Van Reenen, 2014; Rohlfing-Bastian and Schöttner, 2017, 2021). A different perspective is provided by Self-Determination Theory (Ryan and Deci, 1985, 2000) which has identified autonomy as an essential pillar of human self-motivation. The theory formulates the idea that autonomy might directly increase individuals' intrinsic motivation. Empirical research in economics has documented that individuals intrinsically value autonomy (Bartling, Fehr, and Herz, 2014; Owens, Grossman, and Fackler, 2014; Ertac, Gumren, and Gurdal, 2020) and frequently finds that autonomy is associated with positive effects on effort and performance (Zuckerman et al., 1978; Bloom et al., 2015; Beckmann, Cornelissen, and Kräkel, 2017; Kiessling, Radbruch, and Schaube, 2021; Beckmann and Kräkel, 2022). We provide a clean test on the direct effect of task autonomy on performance. Studying this effect not only requires controlling for selection into environments that allow for autonomous decisions but also for selection due to autonomy, as autonomy allows individuals to self-select workplace features that make them more productive, more motivated, or reduce their effort costs. Our design accounts for both potential confounds. Whereas random treatment assignment precludes selection into autonomy, we can control for selection due to autonomy because we elicit task preferences of all participants, including those who get assigned a task.[3] Hence, we can compare subjects who actively self-selected their preferred task and subjects who were assigned their preferred task.

Related to this paper, Kamei and Markussen (2022) investigate the effects of task (mis)matching on productivity and free riding under individual and team-based performance incentives in a real-effort laboratory experiment. Whereas they focus on the effect of productivity-sorting under different remuneration schemes due to the opportunity to vote on the task, we are primarily interested in how task (mis)matching and self-selection affect intrinsic motivation. We complement their study by revealing that task (mis)matching can also be important in the absence of performance pay, as taking into account preferences can significantly increase not only productivity but also labor supply. Due to this different focus, our experimental design arguably also allows for an easier detection of a potential direct effect of autonomy because subjects' effort may be more elastic. In particular, we chose tasks that have an intrinsic value but excluded monetary incentives to

3. Methodologically, we build on an established experimental paradigm that elicits preferences or voting outcomes but randomly implements or overwrites them (Dal Bó, Foster, and Putterman, 2010; Sutter, Haigner, and Kocher, 2010; Brandts, Cooper, and Weber, 2015; Kiessling, Radbruch, and Schaube, 2021).

perform well. Moreover, we induced substantial opportunity costs of effort by allowing subjects to conclude the online experiment when they stopped working (e.g., Goerg, Kube, and Radbruch, 2019). Yet, both studies do not find evidence for a direct effect of task autonomy on performance.

Delegating decisions on the choice of workplace features in general and task allocation more specifically can potentially be leveraged to increase worker motivation in many situations. Our results suggest that taking into account workers' preferences is especially important in situations where comprehensive and contractible measures of effort and productivity are hard to come by, such that the use of pay-for-performance is not feasible. Whereas employment relationships rarely entail no incentives (Kreps, 1997), it is quite common that performance incentives are only crude or implicit, such that employers frequently rely on their employees' intrinsic motivation to a significant extent. Under these conditions, it may be critical to make sure that employees genuinely enjoy what they do, or that it is personally important to them to perform well. Our results provide evidence that simply allowing for self-selection may increase workers' motivation and performance and, thus, provides an additional tool in the toolbox of firms and organizations.

The remainder of the paper is structured as follows: Section 3.2 describes our experimental design and procedural details, Section 3.3 presents the results, and Section 3.4 concludes.

## 3.2 Experimental Design

The aim of this study is to investigate how the task-allocation procedure affects performance. Different task-allocation procedures may generally affect performance not only by implementing different allocations of workers to tasks, but also by independently affecting workers' intrinsic motivation. Hence, we purposefully designed a work environment in which workers may be intrinsically motivated to work on their task and, to be able to detect differences in motivation, in which effort is elastic. Simultaneously, these are two common aspects of many natural work environments. First, many workers are to some degree intrinsically motivated—either because they enjoy the work they do or value the output they generate. Second, the behavior of employees is usually not fully controlled by the employer, and employees may engage in alternative activities, that is, they face opportunity costs of working.

To capture these features, we implemented a real-effort online experiment where we allowed subjects to choose themselves when to stop working. This approach has been used in other studies (Abeler et al., 2011; Gneezy et al., 2017; DellaVigna et al., 2022) and addresses the issue that subjects' effort in real-effort laboratory experiments tends to be inelastic if there are no tangible opportunity costs (Corgnet, Hernán-González, and Schniter, 2015; Araujo et al., 2016; Goerg, Kube, and Radbruch, 2019), which can make even large differences in (intrinsic) motivation hard to detect. Moreover, we deliberately chose productive real-effort tasks—the output subjects generated had an outside value. This is in contrast to many tasks commonly used in real-effort experiments which are intentionally futile and try to minimize intrinsic motivation.[4]

### 3.2.1 Setup

Subjects in the online experiment participated in two parts. In the first part, they familiarized themselves with the two tasks and had to indicate which task they would prefer to work on in the second part of the experiment. One task required subjects to assess tweets on the German debt brake rule with respect to different categories (henceforth "Assess Tweets").[5] Each subtask displayed a single tweet, and subjects had to answer six questions to classify the tweet. For instance, subjects had to assess to what extent the author expressed approval of the policy and whether the tweet contained an argument to substantiate that stance. The other task was a classic data entry task that required subjects to transcribe student numbers by university from scans of German statistical yearbooks (henceforth "Data Entry").[6] Each subtask displayed the scan of student numbers for one particular year and an empty table where the data had to be entered. To reduce the scope of a single subtask, we instructed subjects to digitalize only the data for universities in a particular federal state, Rhineland-Palatinate. For each university, subjects had to transcribe the name of the university as displayed in the document, the total number of students, and the number of female students. The work screens of both tasks are displayed in Figure 3.B.1 and Figure 3.B.2 in the appendix. To

4. For an overview of real-effort tasks that have been employed in behavioral experiments and an assessment of whether they provide outside value or not, see (Charness, Gneezy, and Henderson, 2018).

5. The debt brake rule is regulated by paragraph (2) of Article 115 of the Basic Law for the Federal Republic of Germany and essentially stipulates that the state has to run a balanced budget to restrict the accumulation of public debt. In 02/2020, we scraped over a thousand recent tweets which contained the hashtag "#schuldenbremse" (German for "debt brake"). We preselected a sufficiently large subset of roughly 150 tweets that were not mere retweets, actually referred to the debt brake rule, and expressed the author's stance on the policy. This subset of tweets was used for the experiment.

6. Included scans start with the fall semester 1977/1978 and end with the fall semester 2004/2005. We only included data on fall semesters and restricted the set of included scans to this time period because the structure of statistical yearbooks changes over time but is fairly constant over this period. While the scans are available online (http://www.digizeitschriften.de/dms/toc/?PID=PPN514402342, last accessed: 02-02-21), the data for the fall semester 1998/1999 and earlier are not available in digitized format (https://www-genesis.destatis.de/genesis//online?operation=table&code=21311-0002&bypass=true&levelindex=1&levelid=1610018460550#abreadcrumb, last accessed: 02-02-21). In addition, the data could not easily be extracted from the scans with the help of conventional conversion programs.

ensure that both tasks were perceived to be productive, we emphasized in the instructions that subjects' assessments of tweets can be used to study the debate on the German debt brake rule on Twitter. Likewise, we underscored that one can use the transcribed data on historical student numbers by university to study the development of the share of female students at German universities over time.[7]

In the second part of the experiment, subjects had to work on one of the two tasks. Depending on the treatment, the task was either self-selected by subjects or exogenously assigned to subjects. Importantly, we allowed subjects to choose how long they wanted to work. We implemented this by asking subjects to work for at least 10-15 minutes. Additionally, the instructions stated that working longer would help to generate more data. Subjects could stop working after each subtask. To quit the work stage, they had to press a button on the screen. However, they were only allowed to press the button and leave the work stage after the minimum work requirement of 10 minutes had expired.

### 3.2.2 Treatments

**Table 3.1.** Summary of Treatments

| Treatment | Worked on Preferred Task | Self-Selected Task |
|---|---|---|
| Assigned Nonpreferred | No | No |
| Assigned Preferred | Yes | No |
| Self-Selected | Yes | Yes |

In a between-subjects design, we study three treatments that differ in the task-allocation procedure in the second part of the experiment. Importantly, knowledge of all subjects' task preferences allowed us to consider subjects' preferences in the allocation procedure.

In the first two treatments, we assigned subjects either their preferred (Assigned Preferred) or their nonpreferred task (Assigned Nonpreferred), based on the binary preference measure elicited in the first part of the experiment. In our setting, subjects may prefer one task over the other due to differences in task-specific ability and an inherent preference for being productive, or because they find that task more enjoyable or worthwhile and have a higher intrinsic motivation to work on that task.[8] A comparison between these two treatments

7. The tasks have several other features which made them suitable for our purpose: both tasks are tedious and can be administered in a computerized format, and manual work on the tasks is not easy to substitute. In addition, they are sufficiently different, such that we could expect the majority of subjects to have a clear preference for one task over the other.

8. "In the wild", other features of the (work) environment can influence task preferences. Most notably, performance-related pay reinforces ability-based sorting, which is reduced in our

holds constant that tasks were exogenously assigned and identifies the effect of preference-based sorting.

In the third treatment (Self-Selected), subjects had the opportunity to explicitly self-select their preferred task. A comparison between Assigned Preferred and Self-Selected holds constant that subjects worked on their preferred task but varies whether that task was assigned or self-selected. This precludes sorting and identifies the direct effect of self-selection per se, which may only affect performance by directly increasing workers' intrinsic motivation. A summary of the treatments is provided in Table 3.1.

### 3.2.3 Procedural Details

The experiment was conducted in September 2020. The experiment was split up into two parts taking place on two consecutive days to create temporal separation between the elicitation of task preferences—which took place in the first part of the experiment—and the assignment (or self-selection) of tasks, which followed in the second part.[9] Subjects had to register for both parts and were informed in the invitation that they would only receive a payment if they completed both parts. We provide a translation of the experimental instructions in Section 3.B.

After they received the participation link, subjects could start immediately with the first part of the experiment, which they had to complete until the end of the day. During the first part, we informed subjects that we would generate two different data sets—one on the debate about the German debt brake rule on Twitter, the other on the development of the share of female students at German universities over time. They were instructed that they would have to work on one of two different tasks to generate data for the corresponding data set on the next day, and that they would receive a fixed payment of 5 € upon completion of the experiment. Before we elicited preferences for both tasks, we familiarized subjects with both tasks and gave them the opportunity to inspect the work screens. Afterwards, we elicited how much they were interested in working on either one of the two tasks and in the underlying topics using two separate seven-point Lik-

setting due to flat incentives. Other potential determinants of task preferences include career concerns or selection into peer groups.

9. One might worry that some subjects' preferences are not stable but change overnight, which would lead to misclassification; for instance, a subject who is assigned the preferred task based on the preference stated the day before might actually not prefer that task anymore. Similarly, it would also be problematic if subjects frequently misreported their preferences. We can assess these issues by comparing subjects' choices in Self-Selected to the preferences they indicated before. We observe inconsistent choices for 13 out of 191 subjects. Additionally, 8 out of these 13 subjects indicated an interest in working on the two tasks which differed by at most one point on two separate 7-point Likert scales. This suggests that they were almost indifferent. Given that only 5 subjects remain, misclassification and misreporting do not pose a serious problem for our analysis.

ert scales each.[10] In addition, we elicited their binary task preferences. For that purpose, we asked subjects to state which task they would prefer to work on the next day. Finally, subjects had to answer a brief sociodemographic questionnaire including questions on age, gender, and high-school grade.

On the next day, subjects received the participation link for the second part, which they could again start immediately and had to complete until the end of the day. Depending on the treatment, subjects were either informed which task they had been assigned or had to self-select their task. Subsequently, they received detailed instructions on their task before entering the work stage. Subjects were asked to work for at least 10-15 minutes and pointed to the fact that working longer would help to generate more data. Leaving the work stage was possible by clicking on a button; yet, this was only possible after a minimum working time requirement of 10 minutes had expired. After the work stage, we elicited subjects' mood and administered a subset of questions from the Intrinsic Motivation Inventory (Ryan, 1982; McAuley, Duncan, and Tammen, 1989) to assess self-reported effort, interest and enjoyment, and perceived choice.[11]

The experiment was implemented using oTree (Chen, Schonger, and Wickens, 2016), and subjects were recruited from the BonnEconLab subject pool via Hroot (Bock, Baetge, and Nicklisch, 2014). We conducted four sessions. In total, 556 subjects participated in the first part of the experiment, and 489 subjects completed both parts. Notably, all subjects who did not finish the experiment—except for a single person—dropped out before the start of the second part, i.e., before the treatment manipulation and before they knew which task they would have had to work on. While randomization took place on those subjects who finished the first stage, we only include subjects who finished both parts of the experiment in our analysis. Because we expected a potential direct effect of self-selection to be smaller than a potential effect of task (mis)matching, we oversampled treatments Assigned Preferred and Self-Selected relative to Assigned Nonpreferred. To approximate similar treatment distributions for both tasks, we stratified subjects by their binary task preferences before we randomized them into treatments in the following way: out of every five participants who indicated a preference for a task, one was randomized into Assigned Nonpreferred, two were randomized into Assigned Preferred, and two were randomized into Self-Selected. Among those subjects that completed both parts of the experiment, 103 were in Assigned Nonpreferred, 195 in Assigned Preferred, and 191 in Self-

10. The (translated) questions can be found in the experimental instructions in Section 3.B.1.

11. For each of these three subscales of the Intrinsic Motivation Inventory we included three questions. Answers had to be indicated on 7-point Likert scales. Each sub-scale measure averages over the three corresponding questions, taking into account that some questions are reverse-coded. The (translated) questions can be found in the respective section in the experimental instructions in Section 3.B.2.

Selected. Table 3.A.1 shows that randomization is balanced on observable characteristics. Given a work stage of 10-15 minutes, completing both parts took about 30 minutes. Subjects received their fixed payment of 5 € via bank transfer. This study is registered in the AEA RCT Registry and the unique identifying number is: "AEARCTR-0006373".

### 3.2.4 Task Preferences

Among all participants, 60% preferred Assess Tweets, and 40% preferred Data Entry. Self-reported interest of 4.76 for Assess Tweets and 3.95 for Data Entry, measured on separate scales from 1 (low) to 7 (high), underscores that Assess Tweets was slightly more popular.

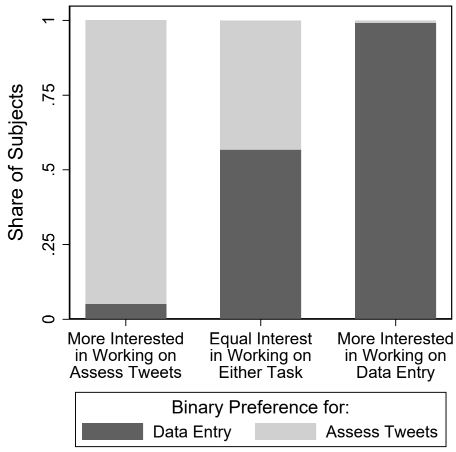
Figure 3.1a shows that relative task interest predicts binary task preferences. Conditional on whether subjects indicated to be more interested in working on Assess Tweets, Data Entry, or both tasks similarly on the 7-point Likert scale, the figure depicts how subjects' binary preferences were distributed over the two tasks. Nearly all subjects who were more interested in assessing tweets preferred to work on that task; conversely, nearly all subjects who were more interested in transcribing student numbers preferred Data Entry. Meanwhile, those subjects who were equally interested in both tasks preferred each of the two in similar proportions. These observations indicate that subjects stated their genuine task preferences, which can be considered a prerequisite to our hypothesis that subjects who work on their preferred task provide more effort: if preferences were shaped by factors unrelated to the tasks themselves, it would be less clear why subjects should be more intrinsically motivated to work on their preferred task. Table 3.A.2 in Section 3.A provides additional information on how subjects' characteristics differ conditional on task preference.

As an immediate consequence, our treatments successfully introduced variation in how much subjects were interested in working on the task they ended up with. This is illustrated in Figure 3.1b, which displays treatment averages of subjects' interest in working on the task they were assigned or had self-selected. In contrast to subjects in Assigned Nonpreferred, subjects in Assigned Preferred and Self-Selected displayed a strong interest in their respective task.

## 3.3 Results

We use the experimental variation in the task-allocation procedures to study their causal effects on performance. We first focus on subjects' output as the most comprehensive measure of performance. Output corresponds to the number of tweets

**(a)** Relative Interest Predicts Binary Task Preferences




**(b)** Manipulation Check

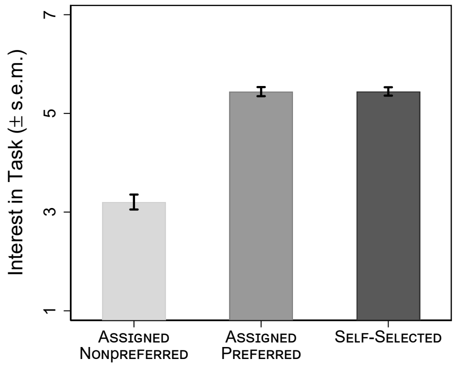
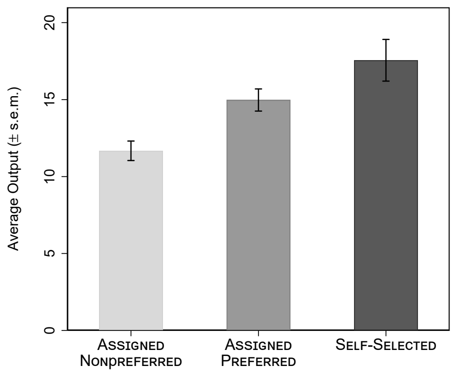
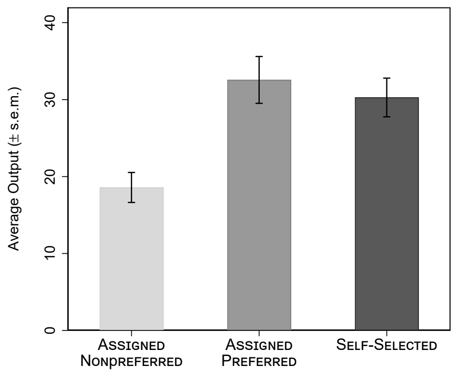



**Figure 3.1**

*Notes:* The figure in (a) depicts the shares of subjects who indicate a preference for Assess Tweets vs. Data Entry, conditional on which task they indicated to be more interested to work on. The figure in (b) shows subjects' average interest in working on the assigned/self-selected task by treatment.




**(a)** Assess Tweets




**(b)** Data Entry

**Figure 3.2.** Average Output by Treatment

*Notes:* For each task, the figure shows subjects' average output by treatment.

assessed or the number of rows transcribed correctly.[12] Figure 3.2 displays subjects' average output by treatment for both tasks separately.

12. A row is considered to be transcribed correctly if it matches the modal input for that row. The modal input of a row is determined by first gathering the answers of those subjects who worked on the respective scan and subsequently taking the input which is most frequent.

In a first step, we compare the performance of subjects in Assigned Preferred and Self-Selected—where the task-allocation procedure allows subjects to work on their preferred task—to the performance of subjects in Assigned Nonpreferred. For both tasks, we observe that the average output is higher if the task-allocation procedure takes into account subjects' task preferences. Subjects in Assigned Preferred (Self-Selected) assessed on average 3.3 (5.9) tweets more than in Assigned Nonpreferred, which corresponds to an increase of 28% (50%) relative to the baseline level of 11.7 tweets. Likewise, they transcribed 14 (11.7) additional rows of data on average, corresponding to an increase of approximately 75% (63%) compared to the average output of 18.6 rows in the baseline. All these comparisons are significant at least at the 5-percent level (Mann-Whitney U tests).[13]

To test the overall effect, we pool the data from both tasks. For this purpose, we normalize subjects' output for each task by the corresponding sample mean in Assigned Nonpreferred.[14] Table 3.2 reports the corresponding results from a regression of these normalized outcomes on treatment indicators.

Column (1) shows that, in the pooled sample, average output was 47% (55%) higher in Assigned Preferred (Self-Selected) relative to Assigned Nonpreferred, and both coefficients are significant. Adding control variables in column (2) does not affect our results.[15] We summarize these observations in the following result:

**Result 3.1.** *Subjects who worked on their preferred task (Self-Selected and Assigned Preferred) produced significantly more output than subjects who worked on their nonpreferred task (Assigned Nonpreferred). This holds for both tasks.*

Note that the comparison between Assigned Preferred and Assigned Nonpreferred isolates the sorting effect that accrues from assigning workers their preferred tasks instead of those that oppose their preferences. It holds constant the fact that tasks are assigned to subjects. The comparison of Self-Selected and Assigned Nonpreferred, however, identifies the joint effect of allocating subjects their preferred instead of non-preferred task and allowing them to actively self-select a task instead of being assigned one.

13. Assigned Preferred vs. Assigned Nonpreferred: p-value = 0.02 (Assess Tweets); p-value < 0.01 (Data Entry). Self-Selected vs Assigned Nonpreferred: p-value = 0.02 (Assess Tweets); p-value < 0.01 (Data Entry). All tests reported in this paper are two-sided.

14. This normalization procedure takes into account that output is measured on different scales across the two tasks and allows for an intuitive interpretation of treatment effects relative to the baseline in Assigned Nonpreferred. We corroborate this analysis with OLS regressions using plain outcome measures, either pooling the data from both tasks and using a task-fixed effect, or considering each task separately. The results are reported in Tables 3.A.3 - 3.A.5.

15. In Table 3.A.6, we additionally show that the results prevail if we drop subjects who might have misunderstood or not properly read the instructions and transcribed no row of data correctly.

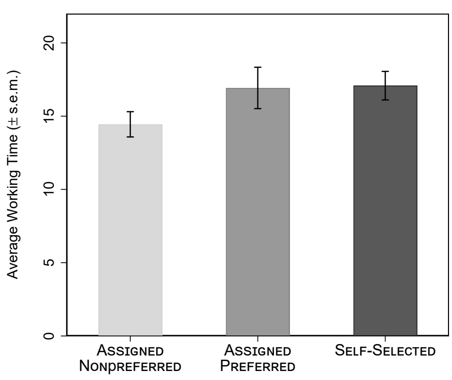
**Table 3.2.** Percentage Improvement in Average Output Relative to Assigned Nonpreferred (Pooled)

| | Output | |
|---|---|---|
| | (1) | (2) |
| Assigned Preferred | 0.47*** | 0.47*** |
| | (0.10) | (0.10) |
| Self-Selected | 0.55*** | 0.55*** |
| | (0.11) | (0.11) |
| Controls | No | Yes |
| Wald test (p-value), | | |
| $H_0$: Assigned Preferred = Self-Selected | 0.53 | 0.50 |
| $R^2$ | 0.04 | 0.07 |
| Observations | 489 | 489 |

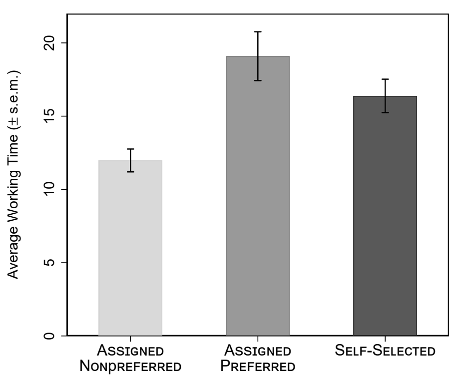
*Notes:* This table reports results from OLS regressions of normalized output on treatment indicators, pooling data from both tasks. Baseline treatment (omitted) is Assigned Nonpreferred. Normalization was conducted by dividing each subject's output by the respective sample mean in Assigned Nonpreferred. Controls include gender, age, high school degree (y/n), and high school grade (linearized). *, **, and *** denote significance at the 10, 5, and 1 percent level. Robust standard errors are reported in parentheses.

In a second step, we isolate the direct effect of self-selection by comparing the performance of subjects in Self-Selected to the performance of those in Assigned Preferred. We do not observe a consistent pattern across the two tasks. Subjects in Self-Selected assess on average 2.6 tweets more than subjects in Assigned Preferred (17% increase relative to Assigned Preferred; Mann-Whitney U test: p-value = 0.77). Contrarily, subjects who Self-Selected Data Entry transcribed on average 2.3 rows of data less than subjects in Assigned Preferred (7% decrease; Mann-Whitney U test: p-value = 0.91). Hence, the order of treatments in terms of average output is flipped across tasks, and the differences are insignificant in both cases.

Again, we can use the regression results in Table 3.2 to compare the differences pooling the data from both tasks. Comparing the coefficients on the treatment indicators reveals that the increase in average performance was larger in Self-Selected than in Assigned Preferred, but the difference is not significant (Wald test, p-value = 0.50). Overall, the regression analysis substantiates the findings from the non-parametric and task-specific analysis above. Both task-allocation procedures that allowed subjects to work on their preferred task induced similar performance increases relative to the case where subjects were assigned their non-preferred task. Active self-selection, however, did not yield an additional increase in performance. We summarize these observations in the following finding.




**(a)** Assess Tweets




**(b)** DataEntry

**Figure 3.3.** Average Working Time by Treatment

*Notes:* For each task, the figure shows subjects' average working time by treatment.

**Result 3.2.** *Average performance does not differ significantly between subjects who self-selected their preferred task (Self-Selected) and subjects who were assigned their preferred task (Assigned Preferred).*

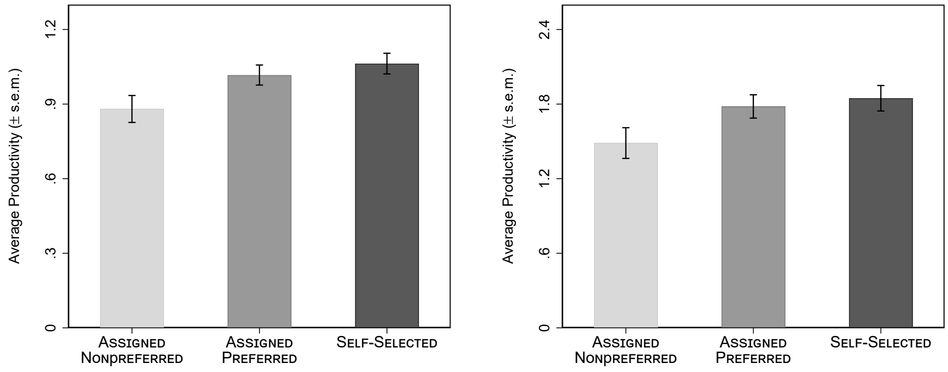
**Productivity or Labor Supply?**

Changes in performance can either result from changes in workers' productivity or the amount of time workers allocated to the task. As our experiment allowed for adjustments at both margins, we assess whether average working time and productivity differ between treatments.

We conceptualize working time as the time subjects spent on the work screen until they arrive at their final subtask.[16] Productivity is measured by dividing the output by our measure of working time and, thus, describes how many units of output a subject produced on average per minute of working time. Thus, subjects' productivity is jointly determined by subjects' ability and effort on the intensive margin, whereas working time is a measure of effort on the extensive margin.

Figure 3.3 displays subjects' average working time by treatment and task. In both tasks, subjects worked longer if the task-allocation procedure allowed them to work on their preferred task. In Assess Tweets, subjects worked about 2.6 minutes longer on average in both Assigned Preferred and Self-Selected than

16. This is our preferred measure as some subjects might have remained on the work screen until the minimum working time had expired without actually working. Alternative definitions of working time would include the time subjects remain on the work screen until they successfully quit and time until subjects first attempt to quit. All measures are highly correlated (Pearson's r > 0.86 in all cases) and using one of the alternative measures yields similar results, reported in Table 3.A.7.




**(a)** Assess Tweets

**(b)** DataEntry

**Figure 3.4.** Average Productivity by Treatment

*Notes:* For each task, the figure shows subjects' average productivity by treatment.

subjects in Assigned Nonpreferred. This amounts to an increase of roughly 17% relative to the average working time of 14.4 minutes in the latter treatment. In Data Entry, subjects in Self-Selected worked 4.4 minutes longer on average than subjects in Assigned Nonpreferred (+37%). This difference is even more pronounced in Assigned Preferred, where the average working time is 7.1 minutes longer (+59%) compared to the baseline (12 minutes). While in Data Entry the differences between Assigned Preferred or Self-Selected and Assigned Nonpreferred are statistically significant (Mann-Whitney U tests, p-value $< 0.01$ in both cases), these comparisons do not reach statistical significance in Assess Tweets (p-value $= 0.45$ and p-value $= 0.50$).[17]

Comparing the working time between Assigned Preferred and Self-Selected, we find virtually no difference in Assess Tweets (p-value $= 0.94$). In Data Entry, subjects worked slightly longer on average if they were assigned their preferred task than if they could actively self-select that task, but the difference is not significant (p-value $= 0.77$).

Figure 3.4 provides the corresponding comparison of average productivity. In both tasks, subjects' were on average more productive if they worked on their preferred instead of nonpreferred task. In Assess Tweets, subjects in Assigned Preferred and Self-Selected assessed 0.14 and 0.18 additional tweets per minute compared to Assigned Nonpreferred, which amounts to an increase of 16% and 21% given the baseline of 0.88 tweets per minute. Both differences barely

17. A look at the cumulative distribution functions (CDFs) for working time, displayed in Figure 3.A.3, provides an explanation for this observation: while in Data Entry the CDFs for Assigned Preferred and Self-Selected are uniformly shifted to the right relative to the CDF for Assigned Nonpreferred, the treatments seem to have only affected the upper third of the distribution in Assess Tweets.

fail to reach the next higher significance level (Mann-Whitney U tests, p-value = 0.11 and p-value = 0.06). In Data Entry, the pattern is very similar: subjects in Assigned Preferred and Self-Selected transcribed 0.29 and 0.36 additional rows per minute compared to Assigned Nonpreferred, corresponding to a 20% and 24% increase given the baseline of 1.49 rows per minute and at least marginally significant in both cases (p-value = 0.06 and p-value = 0.04). In the appendix, we show that these treatment effects are partially driven by subjects working more diligently if the task-allocation procedure respects their preferences.[18]

Again, we can also compare average productivity in Assigned Preferred and Self-Selected. While in both tasks average productivity slightly increased if subjects could self-select their preferred task instead of being assigned that task, these differences are small and not significant in both tasks (p-value > 0.6 in both cases).

Akin to our earlier analysis of treatment differences in average output, we normalize subjects' working time and productivity in each task by the corresponding sample mean in Assigned Nonpreferred to pool the data from both tasks in an OLS regression of the two outcomes on treatment indicators. The results, reported in Table 3.3, confirm the non-parametric tests. Working time increases for both Assigned Preferred and Self-Selected. Comparing these two treatments, the increase in Assigned Preferred is even slightly stronger, but the difference is not significant (Wald test, p-value = 0.32). We observe the same pattern for productivity—subjects in Assigned Preferred and Self-Selected produced more output per minute than subjects in Assigned Nonpreferred. The difference is,

18. For Data Entry, we can explore the share of rows that subjects enter correctly. The right panel of Figure 3.A.5 displays CDFs of the share of rows transcribed correctly for the three treatments. Relative to the CDF for Assigned Nonpreferred, the CDFs for the other two treatments are markedly shifted to the right. We can reject the hypothesis that the distributions are similar if we compare Self-Selected to Assigned Nonpreferred (Mann-Whitney U test, p-value = 0.02), and almost reject that hypothesis if we compare Assigned Preferred and Assigned Nonpreferred (p-value = 0.10). Conversely, the distributions in Assigned Preferred and Assigned Nonpreferred are quite similar (p-value = 0.33). These observations are substantiated by OLS regressions reported in Table 3.A.5.

**Table 3.3.** Percentage Improvement in Average Working Time & Productivity Relative to Assigned Nonpreferred (Pooled)

| | Working Time | | Productivity | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Assigned Preferred | 0.34*** | 0.34*** | 0.17*** | 0.17** |
| | (0.09) | (0.09) | (0.07) | (0.07) |
| Self-Selected | 0.25*** | 0.24*** | 0.22*** | 0.22*** |
| | (0.07) | (0.07) | (0.07) | (0.07) |
| Controls | No | Yes | No | Yes |
| Wald test (p-value), $H_0$: Assigned Preferred = Self-Selected | 0.32 | 0.32 | 0.39 | 0.34 |
| $R^2$ | 0.02 | 0.04 | 0.02 | 0.05 |
| Observations | 489 | 489 | 485 | 485 |

*Notes:* This table reports results from OLS regressions of normalized working time (columns (1) and (2)) and productivity (columns (3) and (4)) on treatment dummies, pooling data from both tasks. Baseline treatment (omitted) is Assigned Nonpreferred. Normalization was conducted by dividing each subject's working time (productivity) by the respective sample mean in Assigned Nonpreferred. Note that for four subjects productivity is not defined because they quit in the first round and, hence, have a working time of 0. Controls include gender, age, high school degree (y/n), and high school grade (linearized). *, **, and *** denote significance at the 10, 5, and 1 percent level. Robust standard errors are reported in parentheses.

again, insignificant (p-value $= 0.34$).[19] These observations for working time and productivity are summarized as follows:

**Result 3.3.** *Both task-allocation procedures that allow subjects to work on their preferred task increase productivity as well as labor supply. There are no significant differences between Self-Selected and Assigned Preferred.*

19. If subjects are working on their preferred task, they might be more likely to make extended breaks and resume work, which would bias their actual working time and, therefore, productivity. While such extended breaks can hardly account for the pattern we observe in the data—as treatment comparisons either reveal no differences (Assigned Preferred vs. Self-Selected) or differences in both working time *and* productivity (Assigned Preferred & Self-Selected vs. Assigned Nonpreferred)—we can further assess whether extended breaks affect our estimates by considering the time subjects require for single subtasks. A histogram of times required per subtask is provided for both tasks separately in Figure 3.A.1 in the appendix. We observe few cases in which subjects took unreasonably long to complete a subtask, such that we can conclude that there are few instances of extended breaks in the data. Additionally, we show in Table 3.A.8 that our results are not affected if we exclude subjects who paused, defined as having required at least 2 standard deviations more than the average time required per subtask on this task for at least one subtask.

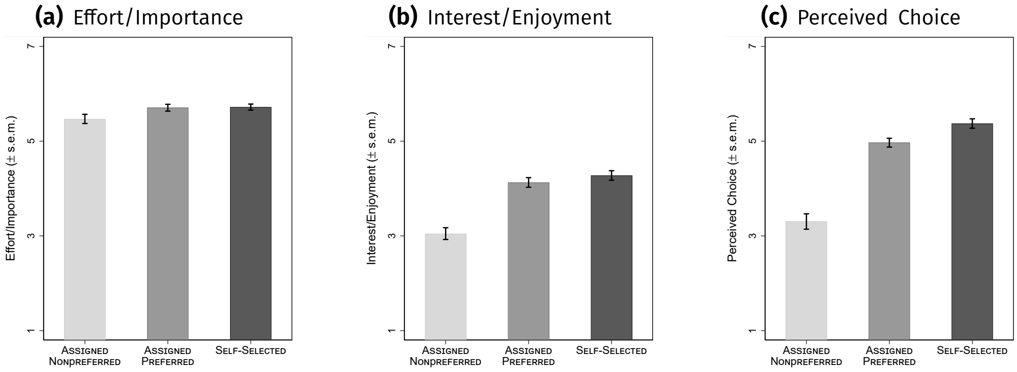



**Figure 3.5.** Evidence from the Post-Experimental Questionnaire

*Notes:* The figure in (a) shows subjects' average self-reported effort elicited in the Intrinsic Motivation Inventory (IMI) questionnaire. The figure in (b) shows subjects' average self-reported enjoyment elicited in the IMI. The figure in (c) shows subjects' average perceived choice elicited in the IMI. The corresponding regressions are reported in Table 3.A.9 in Section 3.A.

### Discussion of Results

Subjects' average performance increased if the task-allocation procedure took their task preferences into account. Subjects who worked on their preferred task produced more output per minute and worked longer on average. While the increase in average productivity could be explained by ability-based sorting alone, the effect on working time suggests that subjects were also more intrinsically motivated to provide effort if they worked on their preferred task. This interpretation is corroborated by evidence from the post-experimental questionnaire. As displayed in Figure 3.5a, self-reported effort was significantly higher for subjects who worked on their preferred task than for subjects who worked on their nonpreferred task (Mann-Whitney U test, p-value $< 0.05$ for both pairwise comparisons). In addition, Figure 3.5b shows that subjects enjoyed working more — or disliked working less — if they worked on their preferred task (p-value $< 0.01$ for both pairwise comparisons). Hence, we conclude that the observed sorting effect is not exclusively a consequence of ability-sorting but at least in part driven by increased effort due to higher intrinsic motivation.

The act of self-selection per se did not significantly affect performance in the context of our experiment. This indicates that if subjects worked on their preferred task, their intrinsic motivation did not depend on whether that task was actively self-selected or assigned, as this comparison ruled out sorting effects by design. In line with the results on subjects' behavior, self-reported effort and enjoyment were similar between subjects who were assigned their preferred task and subjects who could actively self-select their preferred task (Mann-Whitney U test, effort: p-value $= 0.82$, enjoyment: p-value $= 0.28$).

A potential explanation for the absence of a direct effect could be that subjects in Assigned Preferred might have perceived the assignment of their preferred task as a consequence of themselves indicating their preference earlier in the experiment. In that case, they could have already experienced a high level of autonomy.[20] We tried to minimize this concern while keeping the context and the assignment procedure as natural as possible by splitting the experiment into two parts. Thereby, we introduced a temporal separation between the elicitation of task preferences and the assignment of tasks. Yet, data on subjects' perceived choice elicited in the post-experimental questionnaire and displayed in Figure 3.5c indeed suggests that subjects who were assigned their preferred task already perceived a substantially higher degree of autonomy than subjects who were assigned their nonpreferred task ($d = 1.66$; Mann-Whitney U test, p-value $< 0.001$). While the level of perceived autonomy—somewhat reassuringly—further increased if subjects explicitly self-selected a task ($d = 0.41$, Mann-Whitney U test, p-value $< 0.001$), this might have compromised our ability to detect a treatment effect.

## 3.4 Conclusion

While different facets of the work environment can have a substantial influence on workers' performance, employers often lack critical information to assign "optimal" workplace features and implement fine-grained incentives to elicit effort. In the context of task allocation—a central component of work environments—this study examined whether taking into account workers' preferences can increase their effort and performance.

Our results highlight the crucial role of task (mis)matching: subjects' average output increased by almost 50% if they were assigned their preferred instead of nonpreferred task. The increase in output can be attributed to higher productivity and longer working times. This indicates that the beneficial effect of preference-based sorting was not exclusively driven by a correlation between preferences and task-specific ability, but also by higher intrinsic motivation to provide effort. The results further show that beyond inducing a suitable task matching, active self-selection of tasks did not have a significant impact. While the performance of subjects who had the opportunity to actively self-select their preferred task was markedly higher than the performance of those who were assigned their nonpreferred task, it did not differ significantly from the performance of subjects who were assigned their preferred task. Hence, we do observe a beneficial effect of

20. Ideally, to make the distinction between assignment and self-selection of a preferred workplace feature as sharp as possible, one would have to assign the preferred feature without eliciting preferences beforehand. Yet, this would not allow to identify any direct effect of self-selection as it does not allow to condition on these otherwise unobserved preferences.

self-selection in our setting, but this effect is driven by preference-based sorting and not the act of self-selecting a task per se.

These observations complement the findings in Kamei and Markussen (2022), where piece-rate incentives induce individuals to prefer the task they are more able at, but where individuals do not provide significantly more effort if they work on their preferred task.[21] The differences across the two studies can be reconciled by considering differences in the experimental design: the absence of performance incentives and the presence of more attractive outside options in our context are likely to increase subjects' elasticity of effort with respect to intrinsic motivation. Moreover, incentives and the deliberately futile tasks used in Kamei and Markussen (2022) might have crowded out intrinsic motivation altogether (Frey and Jegen, 2001). We conclude that taking into account workers' preferences may elicit most additional effort if they face high opportunity costs and are not already incentivized by performance pay.

The strong effect of preference-based sorting on performance observed in this study shows that sub-optimal task assignment within firms and organizations may lead to potentially large efficiency losses. While firms will rarely be able to improve from always assigning the non-preferred task to always assigning the preferred task (the comparison undertaken in the experiment), a back-of-the-envelope extrapolation still yields a 20% performance increase of taking preferences into account relative to a (counterfactual) baseline where the preferred task is assigned with 50% probability.[22] Thus, employers should try to consult employees' preferences, especially if they cannot comprehensively incentivize effort. Our results suggest that assigning preferred tasks after preference elicitation yields similar results as allowing for active self-selection. While task assignment has the advantage that employers retain more control, a potential drawback of asking workers for their preferences is that this might increase disappointment among those whose preferences cannot be taken into account due to organizational constraints. Studying how individuals' reaction to the assignment of non-preferred workplace features depends on whether their preferences were elicited beforehand and which strategies might mitigate negative responses seems to be an interesting avenue for future research.

Our observations on the effects of task (mis)matching may also inform management decisions on whether to invest in monitoring technologies. While monitoring technologies can allow for incentivization and mitigate moral hazard problems

21. This is evident in the fact that subjects in that study's Individual treatment do not significantly reduce the time spent on the outside option.

22. Given that the average output was about 50% higher if subjects were assigned their preferred (instead of nonpreferred) task, $\overline{y_{Preferred}} = 1.5 * \overline{y_{Nonpreferred}}$, a linear extrapolation yields $\overline{Y_{Preferred}}/\overline{Y_{Counterfactual}} = \overline{Y_{Preferred}}/(0.5 * (\overline{Y_{Preferred}} + \overline{Y_{Nonpreferred}})) = 1.2$, that is, a 20% increase relative to the counterfactual baseline in which individuals are assigned their preferred and nonpreferred task with 50% probability each.

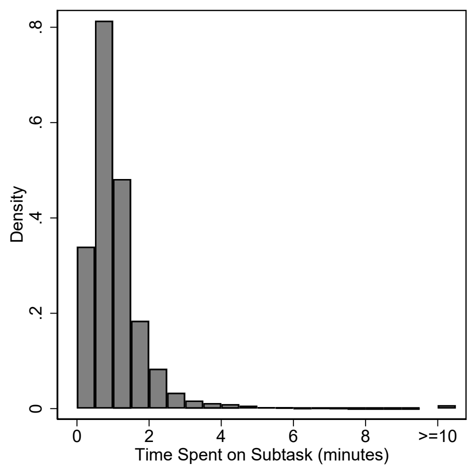
(Hubbard, 2000; Nagin et al., 2002; Duflo, Hanna, and Ryan, 2012; Pierce, Snow, and McAfee, 2015), behavioral research has shown that individuals dislike being controlled (Falk and Kosfeld, 2006). Consistent with theoretical work that investigates the interplay of monitoring technologies, performance incentives, and intrinsic motivation (Cordella and Cordella, 2017), our finding that preference-based sorting leads to significant differences in effort provision implies that it might be much more worthwhile for firms to prioritize incentivization if it is inevitable to assign unpopular tasks than if workers work on their preferred task.

In consideration of the literature on Self-Determination Theory (Ryan and Deci, 1985, 2000) and a variety of empirical studies which have found autonomy to be associated with increased effort and performance (Zuckerman et al., 1978; Bloom et al., 2015; Beckmann, Cornelissen, and Kräkel, 2017; Kiessling, Radbruch, and Schaube, 2021), the absence of a direct effect of self-selection might be surprising. As discussed in the previous section, a potential explanation is that subjects who were assigned their preferred task already perceived a high degree of autonomy, and that the treatment that had subjects explicitly self-select their preferred task was too much of a stress test. Yet, our finding is in line with Kamei and Markussen (2022), who make the randomization procedure that assigns tasks in the "no autonomy" condition explicit but also find no significant effect of self-selection per se. In conclusion, the mixed evidence calls for further research that combines the advantages of a rich-enough setting that admits intrinsic motivation and induces a high elasticity of effort on the one hand and a clean but vigorous treatment manipulation that allows to control for selection due to autonomy on the other.

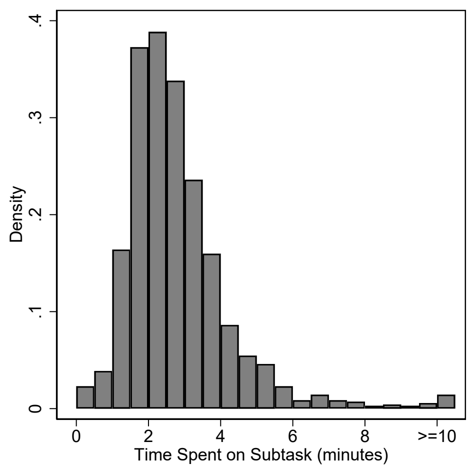
## Appendix 3.A Supplementary Figures and Tables

In this section we present supplementary figures and tables which complement the evidence from the main text. Figure 3.A.1 shows a histogram of per-subtask times for the two tasks. Figures 3.A.2 - 3.A.5 display cumulative distribution functions of the different outcome measures for the two tasks. Table 3.A.1 reports treatment averages of sociodemographic and task-related variables and constitutes both a randomization and manipulation check. Table 3.A.2 reports these variables conditional on subjects' binary task preference. Tables 3.A.3, 3.A.4, and 3.A.5 report average treatment effects from OLS regressions using plain outcome measures for the pooled sample and both tasks separately. Table 3.A.8 runs the same regressions that are reported in the main text on a restricted sample that disregards subjects who paused, defined as having required more than 10 minutes for at least one subtask. Table 3.A.6 runs the same regressions that are reported in the main text on a restricted sample that disregards subjects in Data Entry who did not transcribe any data correctly. Table 3.A.7 runs the same regressions that are reported in Table 3.3 using different definitions of working time. Table 3.A.9 reports OLS regressions of the IMI subscales on treatment indicators corresponding to Figure 3.5a in the main text.

### 3.A.1 Supplementary Figures




**(a)** Assess Tweets




**(b)** Data Entry

**Figure 3.A.1.** Histograms of Time Required per Subtask

*Notes:* This figure displays histograms of times required per subtask, separately for the two tasks.

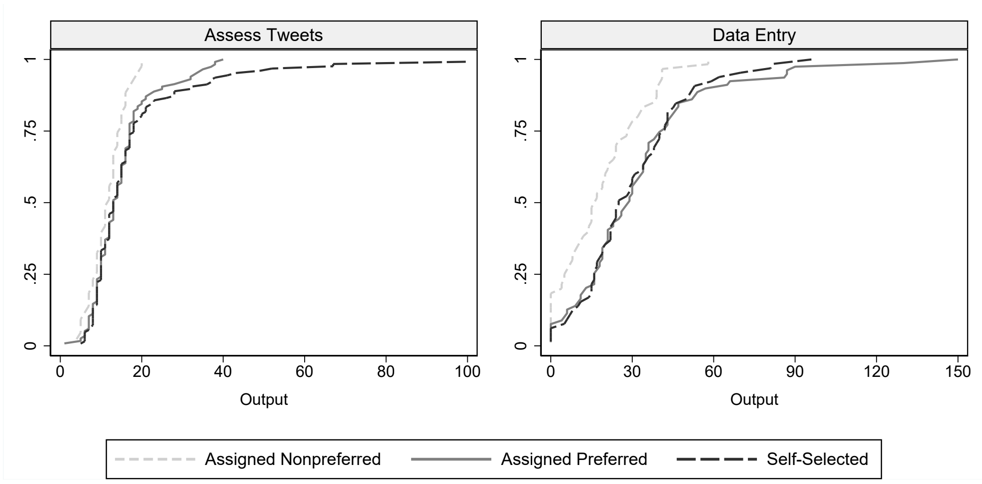



**Figure 3.A.2.** CDF of Output by Task

*Notes:* This figure displays cumulative distribution functions of subjects' output, split by treatment and separately for the two tasks.

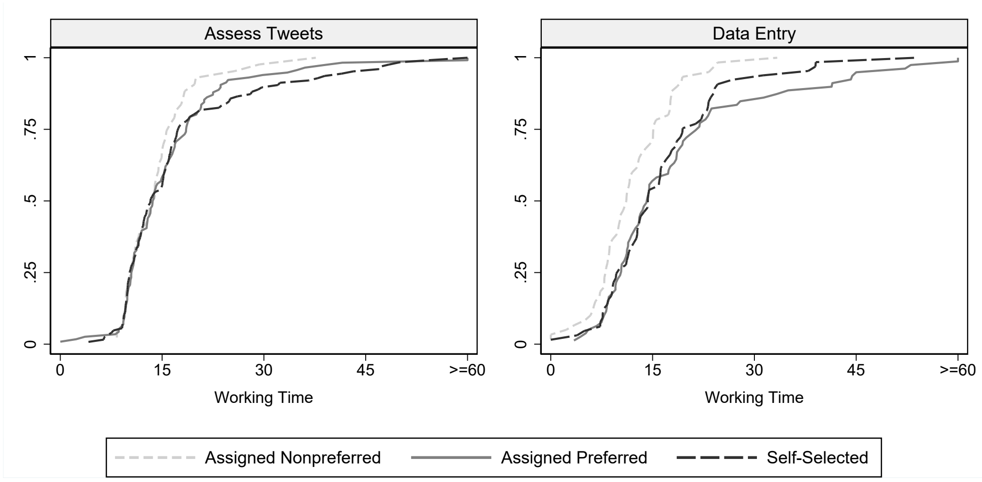



**Figure 3.A.3.** CDF of Working Time by Task

*Notes:* This figure displays cumulative distribution functions of subjects' working time, split by treatment and separately for the two tasks.

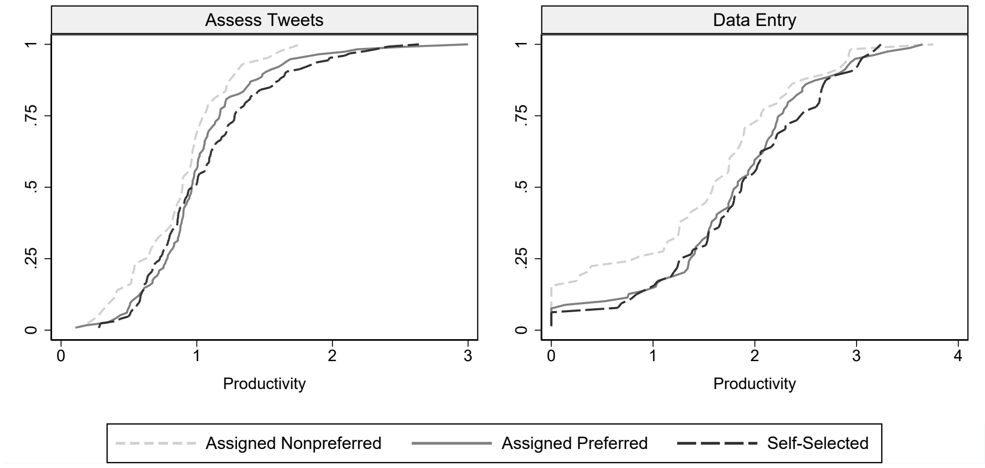



**Figure 3.A.4.** CDF of Productivity by Task

*Notes:* This figure displays cumulative distribution functions of subjects' productivity, split by treatment and separately for the two tasks.

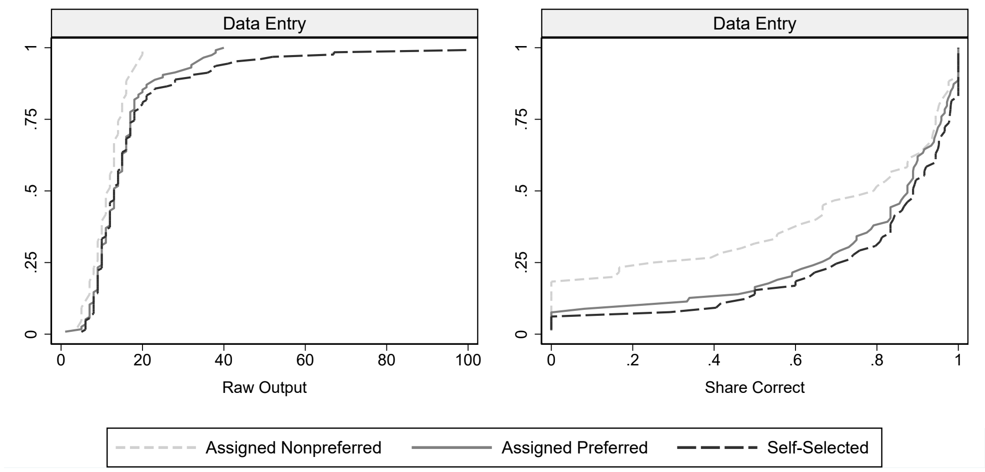



**Figure 3.A.5.** CDF of Raw Output and Share Correct in Data Entry

*Notes:* This figure displays the cumulative distribution functions of subjects' raw output and the share of rows transcribed correctly in task Data Entry, each split by treatment.

## 3.A.2 Supplementary Tables

**Table 3.A.1.** Randomization & Manipulation Check

| | (1) Assigned-NonPreferred | (2) Assigned-Preferred | (3) Self-Selected | (4) $H_0$: (1) = (2) | (5) $H_0$: (1) = (3) | (6) $H_0$: (2) = (3) |
|---|---|---|---|---|---|---|
| **Sociodemographic Variables (Randomization)** | | | | | | |
| Age | 25.31 | 25.76 | 25.82 | 0.59 | 0.55 | 0.94 |
| | (5.97) | (7.43) | (7.51) | | | |
| Female | 0.62 | 0.61 | 0.60 | 0.79 | 0.75 | 0.95 |
| | (0.49) | (0.49) | (0.49) | | | |
| **Task-Related Variables (Randomization)** | | | | | | |
| Interest in Assess Tweets | 4.67 | 4.67 | 4.91 | 0.99 | 0.21 | 0.12 |
| | (1.65) | (1.57) | (1.46) | | | |
| Interest in Data Entry | 3.94 | 3.96 | 3.94 | 0.92 | 0.98 | 0.89 |
| | (1.79) | (1.95) | (1.89) | | | |
| Interest in Assess Tweets Topic | 4.89 | 4.69 | 4.68 | 0.23 | 0.22 | 0.96 |
| | (1.42) | (1.40) | (1.40) | | | |
| Interest in Data Entry Topic | 4.81 | 4.89 | 4.99 | 0.67 | 0.28 | 0.49 |
| | (1.44) | (1.61) | (1.43) | | | |
| **Task-Related Variables (Manipulation)** | | | | | | |
| Interest in Own Task | 3.20 | 5.44 | 5.45 | 0.00*** | 0.00*** | 0.97 |
| | (1.54) | (1.29) | (1.17) | | | |
| Interest in Own Task Topic | 4.46 | 5.24 | 5.13 | 0.00*** | 0.00*** | 0.42 |
| | (1.40) | (1.34) | (1.32) | | | |
| Observations | 103 | 195 | 191 | 298 | 294 | 386 |

*Notes:* This table reports averages of subjects' characteristics by treatment. In columns (1)–(3), standard deviations are reported in parentheses. Columns (4)–(6) report p-values from two-sided t-tests for equality of means between treatments.

**Table 3.A.2.** Subjects' Characteristics Conditional on Task Preference

| | (1)<br>Assess Tweets | (2)<br>Data Entry | (3)<br>$H_0$: (1) = (2) |
|---|---|---|---|
| **Sociodemographic Variables** | | | |
| Age | 26.12 | 25.05 | 0.11 |
| | (7.71) | (6.23) | |
| Female | 0.54 | 0.71 | 0.00*** |
| | (0.50) | (0.45) | |
| **Task-Related Variables** | | | |
| Interest in Assess Tweets | 5.42 | 3.78 | 0.00*** |
| | (1.21) | (1.48) | |
| Interest in Data Entry | 2.89 | 5.53 | 0.00*** |
| | (1.47) | (1.23) | |
| Interest in Assess Tweets Topic | 5.00 | 4.32 | 0.00*** |
| | (1.39) | (1.33) | |
| Interest in Data Entry Topic | 4.48 | 5.56 | 0.00*** |
| | (1.53) | (1.21) | |
| Observations | 293 | 196 | 489 |

*Notes:* This table reports averages of subjects' characteristics conditional on task preference. In columns (1) and (2), standard deviations are reported in parentheses. Column (3) reports p-values from two-sided t-tests for equality of means between subjects who prefer Assess Tweets and subjects who prefer Data Entry.

**Table 3.A.3.** Average Treatment Effects Using Plain Outcomes (Pooled)

| | Output | | Working Time | | Productivity | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Assigned Preferred | 8.78*** | 8.69*** | 4.81*** | 4.74*** | 0.22** | 0.22** |
| | (1.89) | (1.89) | (1.24) | (1.23) | (0.09) | (0.09) |
| Self-Selected | 9.44*** | 9.40*** | 3.86*** | 3.77*** | 0.27*** | 0.28*** |
| | (1.76) | (1.75) | (0.97) | (0.94) | (0.09) | (0.09) |
| Data Entry | 13.44*** | 13.08*** | 0.09 | −0.07 | 0.74*** | 0.74*** |
| | (1.72) | (1.69) | (1.08) | (1.07) | (0.07) | (0.07) |
| Controls | No | Yes | No | Yes | No | Yes |
| Mean in Assess Tweets, Assigned Nonpreferred | 11.67 | 11.67 | 14.44 | 14.44 | 0.88 | 0.88 |
| Wald test (p-value), $H_0$: Assigned Preferred = Self-Selected | 0.71 | 0.69 | 0.47 | 0.46 | 0.42 | 0.37 |
| $R^2$ | 0.15 | 0.17 | 0.02 | 0.05 | 0.24 | 0.25 |
| Observations | 489 | 489 | 489 | 489 | 485 | 485 |

*Notes:* This table reports results from ordinary-least-squares (OLS) regressions of the main outcome measures—output, working time, and productivity—on treatment dummies and a task fixed-effect, pooling data from both tasks. Baseline treatment is Assigned Nonpreferred. Note that for four subjects productivity cannot be determined because they quit in the first round and, hence, have a working time of 0. Controls include gender, age, high school degree (y/n) and high school grade (linearized). *, **, and *** denote significance at the 10, 5, and 1 percent level. Robust standard errors are reported in parentheses.

**Table 3.A.4.** Average Treatment Effects Using Plain Outcomes (Assess Tweets)

| | Output | | Working Time | | Productivity | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Assigned Preferred | 3.30*** | 3.24*** | 2.48 | 2.24 | 0.14** | 0.13** |
| | (0.96) | (1.03) | (1.65) | (1.74) | (0.07) | (0.07) |
| Self-Selected | 5.88*** | 6.03*** | 2.64** | 2.70** | 0.18*** | 0.18** |
| | (1.49) | (1.67) | (1.30) | (1.28) | (0.07) | (0.07) |
| Controls | No | Yes | No | Yes | No | Yes |
| Mean in Assigned Nonpreferred | 11.67 | 11.67 | 14.44 | 14.44 | 0.88 | 0.88 |
| Wald test (p-value), $H_0$: Assigned Preferred = Self-Selected | 0.09 | 0.06 | 0.93 | 0.79 | 0.43 | 0.47 |
| $R^2$ | 0.03 | 0.08 | 0.01 | 0.05 | 0.02 | 0.06 |
| Observations | 285 | 285 | 285 | 285 | 284 | 284 |

*Notes:* This table reports results from ordinary-least-squares (OLS) regressions of the main outcome measures—output, working time, and productivity—on treatment dummies, using only observations from task Assess Tweets. Baseline treatment is Assigned Nonpreferred. Note that for one subject who worked on this task productivity cannot be determined because he or she quit in the first round and, hence, has a working time of 0. Controls include gender, age, high school degree (y/n) and high school grade (linearized). *, **, and *** denote significance at the 10, 5, and 1 percent level. Robust standard errors are reported in parentheses.

**Table 3.A.5.** Average Treatment Effects Using Plain Outcomes (Data Entry)

| | Output | | Working Time | | Productivity | | Raw Output | | Share Correct | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Assigned Preferred | 13.97*** | 12.94*** | 7.12*** | 7.03*** | 0.29* | 0.26* | 15.16*** | 14.41*** | 0.13** | 0.11* |
| | (3.60) | (3.48) | (1.84) | (1.80) | (0.15) | (0.16) | (3.85) | (3.70) | (0.06) | (0.06) |
| Self-Selected | 11.69*** | 10.50*** | 4.41*** | 3.86*** | 0.36** | 0.36** | 9.72*** | 8.79*** | 0.16*** | 0.15** |
| | (3.17) | (3.26) | (1.38) | (1.43) | (0.16) | (0.16) | (2.97) | (3.04) | (0.06) | (0.06) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes | No | Yes |
| Mean in Assigned Nonpreferred | 18.58 | 18.58 | 11.98 | 11.98 | 1.49 | 1.49 | 27.23 | 27.23 | 0.63 | 0.63 |
| Wald test (p-value), $H_0$: Assigned Preferred = Self-Selected | 0.56 | 0.54 | 0.18 | 0.13 | 0.64 | 0.47 | 0.20 | 0.19 | 0.42 | 0.39 |
| $R^2$ | 0.07 | 0.09 | 0.07 | 0.08 | 0.03 | 0.05 | 0.07 | 0.08 | 0.04 | 0.06 |
| Observations | 204 | 204 | 204 | 204 | 201 | 201 | 204 | 204 | 204 | 204 |

*Notes:* This table reports results from ordinary-least-squares (OLS) regressions of the main outcome measures—output, working time, and productivity—and the secondary outcome measures only available for this task—raw output and the share of rows transcribed correctly—on treatment dummies, using only observations from task Data Entry. Baseline treatment is Assigned Nonpreferred. Note that for three subjects who worked on this task productivity cannot be determined because they quit in the first round and, hence, have a working time of 0. Controls include gender, age, high school degree (y/n) and high school grade (linearized). *, **, and *** denote significance at the 10, 5, and 1 percent level. Robust standard errors are reported in parentheses.

**Table 3.A.6.** Percentage Improvement in Average Outcomes Relative to Assigned Nonpreferred (Pooled)—Dropping Subjects Who Transcribed No Row Correctly

| | Output | | Working Time | | Productivity | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Assigned Preferred | 0.40*** | 0.40*** | 0.30*** | 0.29*** | 0.11* | 0.11* |
| | (0.10) | (0.10) | (0.09) | (0.09) | (0.06) | (0.06) |
| Self-Selected | 0.46*** | 0.47*** | 0.21*** | 0.21*** | 0.15** | 0.15** |
| | (0.11) | (0.11) | (0.07) | (0.07) | (0.06) | (0.06) |
| Controls | No | Yes | No | Yes | No | Yes |
| Wald test (p-value), $H_0$: Assigned Preferred = Self-Selected | 0.61 | 0.55 | 0.37 | 0.36 | 0.49 | 0.43 |
| $R^2$ | 0.03 | 0.05 | 0.01 | 0.04 | 0.01 | 0.04 |
| Observations | 468 | 468 | 468 | 468 | 466 | 466 |

*Notes:* This table reports results from ordinary-least-squares (OLS) regressions of normalized outcomes on treatment dummies, pooling data from both tasks but dropping subjects who transcribed no row correctly. Baseline treatment (omitted) is Assigned Nonpreferred. For each task-outcome separately, normalization was conducted by dividing observed outcomes by the respective sample mean in Assigned Nonpreferred. Note that for two of the remaining subjects productivity cannot be determined because they quit in the first round and, hence, have a working time of 0. Controls include gender, age, high school degree (y/n), and high school grade (linearized). *, **, and *** denote significance at the 10, 5, and 1 percent level. Robust standard errors are reported in parentheses.

**Table 3.A.7.** Percentage Improvement in Average Working Time & Productivity Relative to Assigned Non Preferred (Pooled) — Different Definitions of Working Time

| | Default | | Time On Work Screen | | Time Until First Quit Attempt | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Assigned Preferred | 0.34*** | 0.17** | 0.19** | 0.18** | 0.31*** | 0.05 |
| | (0.09) | (0.07) | (0.10) | (0.07) | (0.09) | (0.20) |
| Self-Selected | 0.24*** | 0.22*** | 0.14* | 0.23*** | 0.22*** | 0.37 |
| | (0.07) | (0.07) | (0.08) | (0.07) | (0.07) | (0.30) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Wald test (p-value), $H_0$: Assigned Preferred = Self-Selected | 0.32 | 0.34 | 0.54 | 0.31 | 0.33 | 0.16 |
| $R^2$ | 0.04 | 0.05 | 0.02 | 0.05 | 0.05 | 0.01 |
| Observations | 489 | 485 | 489 | 489 | 489 | 489 |

*Notes:* For different definitions of working time, this table reports results from OLS regressions of normalized working time (odd columns) and productivity (even columns) on treatment dummies, pooling data from both tasks. Baseline treatment (omitted) is Assigned Non Preferred. Normalization was conducted by dividing each subject's working time (productivity) by the respective sample mean in Assigned Non Preferred. Note that, under the default definition of working time, for four subjects productivity cannot be determined because they quit in the first round and, hence, have a working time of 0. Controls include gender, age, high school degree (y/n), and high school grade (linearized). *, **, and *** denote significance at the 10, 5, and 1 percent level. Robust standard errors are reported in parentheses.

**Table 3.A.8.** Percentage Improvement in Average Outcomes Relative to Assigned Nonpreferred (Pooled)—Dropping Subjects Who Paused

| | Output | | Working Time | | Productivity | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Assigned Preferred | 0.45*** | 0.45*** | 0.27*** | 0.27*** | 0.15** | 0.15** |
| | (0.10) | (0.11) | (0.08) | (0.08) | (0.06) | (0.06) |
| Self-Selected | 0.51*** | 0.52*** | 0.20*** | 0.20*** | 0.21*** | 0.21*** |
| | (0.11) | (0.12) | (0.07) | (0.06) | (0.07) | (0.07) |
| Controls | No | Yes | No | Yes | No | Yes |
| Wald test (p-value), $H_0$: Assigned Preferred = Self-Selected | 0.61 | 0.54 | 0.41 | 0.42 | 0.22 | 0.19 |
| $R^2$ | 0.03 | 0.06 | 0.02 | 0.06 | 0.02 | 0.05 |
| Observations | 455 | 455 | 455 | 455 | 455 | 455 |

*Notes:* This table reports results from OLS regressions of normalized outcomes on treatment dummies, pooling data from both tasks but dropping subjects who paused. Subjects were classified to have paused if they required at least 2 standard deviations more than the average time required per subtask on this task for at least one subtask. Baseline treatment (omitted) is Assigned Nonpreferred. For each task-outcome separately, normalization was conducted by dividing observed outcomes by the respective sample mean in Assigned Nonpreferred. Controls include gender, age, high school degree (y/n), and high school grade (linearized). *, **, and *** denote significance at the 10, 5, and 1 percent level. Robust standard errors are reported in parentheses.

**Table 3.A.9.** Self-Reported Effort/Importance, Interest/Enjoyment, and Perceived Choice from the IMI

| | Effort/Importance | | Interest/Enjoyment | | Perceived Choice | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| AssignedPreferred | 0.24** | 0.24** | 1.08*** | 1.07*** | 1.66*** | 1.66*** |
| | (0.12) | (0.12) | (0.16) | (0.16) | (0.19) | (0.19) |
| SelfSelected | 0.25** | 0.25** | 1.23*** | 1.22*** | 2.07*** | 2.06*** |
| | (0.12) | (0.11) | (0.16) | (0.16) | (0.19) | (0.19) |
| Constant | 5.47*** | 5.57*** | 3.05*** | 3.07*** | 3.30*** | 3.10*** |
| | (0.10) | (0.38) | (0.12) | (0.51) | (0.16) | (0.64) |
| Controls | No | Yes | No | Yes | No | Yes |
| Wald test (p-value), $H_0$: AssignedPreferred = SelfSelected | 0.89 | 0.88 | 0.30 | 0.28 | 0.00 | 0.00 |
| $R^2$ | 0.01 | 0.03 | 0.11 | 0.11 | 0.24 | 0.24 |
| Observations | 489 | 489 | 489 | 489 | 489 | 489 |

*Notes:* This table reports OLS regressions of the Intrinsic Motivation Inventory (IMI) subscale measures "Effort/Importance", "Interest/Enjoyment", and "Perceived Choice" on treatment dummies. Controls include gender, age, high school degree (y/n), and high school grade (linearized). Robust standard errors are reported in parentheses.

## Appendix 3.B Experimental Instructions

This section contains the experimental instructions. Comments on the page contents or differences between treatments are *[italicized and placed within brackets]*. Input fields are indicated by `[descriptions in typewriter style within brackets]`.

### 3.B.1 Session 1

**Welcome to the Study!**

Welcome to the BonnEconLab online study. Please note that you may only take part in this study if you have registered for the study in our participation database.

In the following fields, please enter your email address with which you are registered in the BonnEconLab participation database. A payout can only be guaranteed if you enter the correct e-mail address here.

`[input field for e-mail address]`

`[input field to confirm e-mail address]`

---

*[Subjects visited the "Attention" page only if on the previous page they were detected to use a Smartphone, tablet, or a non-admissible browser. The "next" button was only displayed if—upon reloading—an admissible setup was detected.]*

**Attention**

To be able to edit the tasks, your device must be able to display PDFs, among other things. This is not easily possible on many mobile devices. It is also very helpful if you can use a mouse and a keyboard.

Please do not use a smartphone or tablet, but a desktop computer or notebook. As soon as you have changed the device, reload this page using the current URL (to be found in the address bar) and click on the "next" button.

---

**Your Bank Details**

For your completed participation in both parts of this study you will receive a fixed payoff of 5 €. This amount will be paid to you via bank transfer within the next two weeks. For this we need your bank details.

`[several input fields related to the payment procedure and subjects' bank details]`

---

*[Subjects could freely navigate between the pages "General Information", "Task 1", "Example 1", "Task 2", "Example 2", and "Further Information"—either via a navigation bar at the top end of the page, or via buttons on the bottom of the page.]*

**General Information**

Welcome to this study.

Two databases are currently being created at the Institute for Applied Microeconomics at the University of Bonn for research purposes. The first database is intended to document the historical development of student numbers by gender in Germany. The second database is intended to follow the debate on Twitter on the subject of the German debt brake rule.

Within this study you will contribute to one of these databases by adding entries.

The details of your task depend on which database you will work on. Today we will briefly introduce both tasks to you.

Tomorrow you will work on *one of the two* tasks for a period of at least 10 minutes. You are of course welcome to work longer on the task. You will receive the link to your task by email tomorrow.

The tasks:

- *Task 1: Student Numbers*

We are interested in how the proportion of women among students at German universities has changed over time. However, the corresponding data is only available in digital form for the last two decades. Before that, the numbers were recorded in statistical yearbooks. The task is to filter the required information from the scanned yearbooks and to transfer them into an input mask.

- *Task 2: Tweets*

  The public debate on the debt brake was one of the most important economic policy debates within the past year. Our goal is to follow the debate prior to the outbreak of the Covid-19 pandemic on the basis of tweets. We have already collected the relevant tweets. The task is to classify the tweets in terms of certain characteristics (e.g. status of the author, arguments used, linguistic correctness).

On the next few pages, both tasks will be presented. A more detailed description of the task you will ultimately be working on will be given to you tomorrow, just before you start to work.

If the scans or tweets are not displayed correctly, please try a different browser (preferably Chrome, otherwise Opera, Firefox or Safari) or set the security settings of your browser to "standard" and reload the page.

**Task 1**

The input mask for *Task 1: Student Numbers* is structured as follows:

- In the upper section there is the scanned document with the data. You can use a zoom function within the scan.
- In the lower section there is a table in which the data is to be entered.

First of all, you indicate which semester the data refer to, which can be inferred from the scan's table heading. Afterwards, the total number of properly enrolled students and the number of female students is to be transferred for the different universities. Your task will be limited to universities from a certain federal state (e.g. Baden-Wuerttemberg or North Rhine-Westphalia). As soon as all the necessary data for a document have been transferred, you can continue with the next document by clicking on "send data".

On the following page you can familiarize yourself with the task with the help of a sample subtask. In the example, the figures for universities in Schleswig-Holstein have been transferred. At this point you do not need to transfer any data yourself.

**Example 1**

*[Similar to Figure 3.B.2.]*

**Task 2**

The input mask for *Task 2: Tweets* is structured as follows:

- In the upper section there is a statement on the basis of which the tweet is to be assessed, as well as the tweet itself.
- In the lower section there are various questions on the tweet.

All questions aim at a categorization of the tweet with regard to certain criteria. For each tweet all questions are to be answered. As soon as all questions have been answered, you can continue with the next tweet by clicking on “send data”.

On the following page you can familiarize yourself with the task with the help of a sample subtask. At this point you do not need to answer any questions.

**Example 2**

*Similar to Figure 3.B.1.*

**Further Information**

To ensure that your browser correctly displays both the scanned data and the tweets, please answer the following two questions. You can reach the previous pages for example via the navigation bar above.

*[Subjects could only proceed to the next page if both questions were answered correctly.]*

For Example 1: how many students were there in total at the University of Bonn in the winter semester 1982/83?

`[input field for the answer]`

For Example 2: in which German city was the tweet written?

[input field for the answer]

---

You have now got to know both types of tasks. Please let us know how interesting you find working on the two tasks and which one you would prefer to work on.

How much are you interested in working on *Task 1: Student Numbers*?

[input field (7-point Likert scale)]

How much are you interested in working on *Task 2: Tweets*?

[input field (7-point Likert scale)]

Which task would you prefer to work on?

[input field for the binary preference]

---

At the end of this section, we would like to know how interesting you find the subject of the two projects.

How interesting do you find the development of student numbers in Germany?

[input field (7-point Likert scale)]

How interesting do you find the debate on the German debt brake rule?

[input field (7-point Likert scale)]

---

**Further Questions**

As a final step we would like to learn a little more about you.

`[several input fields for sociodemographic information]`

---

**See you tomorrow!**

That's it for now! Tomorrow (2:00 p.m. at the latest) you will receive an email with a link that will take you to the task that you are supposed to work on. You then have time until 11:59 p.m. tomorrow evening to complete the study.

You can now close the browser window.

---

### 3.B.2 Session 2

**Welcome Back!**

Welcome back!

Please enter the same e-mail address in the following fields which you entered in the first section. A payout can only be guaranteed if you enter the correct e-mail address here.

`[input field for e-mail address]`

`[input field to confirm e-mail address]`

*Reminder:* If you complete this second part of the study, you will receive a fixed amount of 5.00 €. This amount will be paid to you via bank transfer within the next two weeks.

---

*[Subjects visited the "Attention" page only if on the previous page they were detected to use a Smartphone, tablet, or a non-admissible browser. The "Next" button was only displayed if—upon reloading—an admissible setup was detected.]*

**Attention**

To be able to edit the tasks, your device must be able to display PDFs, among other things. This is not easily possible on many mobile devices. It is also very helpful if you can use a mouse and a keyboard.

Please do not use a smartphone or tablet, but a desktop computer or notebook. As soon as you have changed the device, reload this page using the current URL (to be found in the address bar) and click on the "Next" button.

---

*[Only subjects in treatments Assigned Preferred and Assigned Nonpreferred visited the subsequent page. "[TASK]" read "Task 1: Student Numbers" for subjects who got assigned Data Entry, but read "Task 2: Tweets" for subjects who got assigned Assess Tweets.]*

**Your Task**

We still need people to work on [TASK]. Therefore, you have been assigned this task.

Please work at least 10 minutes on [TASK]. Of course you are welcome to work longer on the task. Thereby you contribute to gathering more data.

---

*[Only subjects in treatment Self-Selected visited the subsequent page.]*

**Your Task**

We still need people to work on *Task 1: Student Numbers* as well as *Task 2: Tweets*.

Therefore, you can choose which of the two tasks you would like to work on. Please choose the task you would like to work on.

`[binary task choice via two different buttons—one for each task]`

Please work at least 10 minutes on the task you have chosen. Of course you are welcome to work longer on the task. Thereby you contribute to gathering more data.

---

*[Only subjects working on Assess Tweets visited the subsequent set of pages. Subjects could freely navigate between the pages "Instructions", "Example", and "Let's Go!"—either via a navigation bar at the top end of the page, or via buttons on the bottom of the page.]*

**Instructions**

*Reminder:* We are interested in the public debate about the debt brake—a central economic policy dispute of last year. We want to understand the debate before the COVID-19 pandemic with the help of tweets. To that end tweets are to be classified.

*Your task in detail:*

- At the top of each page you will find a statement about the debt brake and a tweet. The statement preceding the tweet always remains the same.
- Below that you will find a number of questions on this tweet. *Please answer all the questions displayed.*

For brief explanations of the individual questions, click on the respective info button.

On the following page you will find another example to help you understand the instructions and become familiar with the task. However, you do not need to answer any questions at this point.

As soon as all questions for a tweet have been answered, you can continue with the next tweet by clicking on "submit data", or you can end work by clicking on "submit data and finish work".

**Example**

*[Similar to Figure 3.B.1.]*

**Let's go!**

Please answer a final question before you start working.

How do you personally assess the basic statement above the example tweet?

`[input field (7-point Likert scale)]`

You can start working with a click on the “Let’s go!” button.

### Assess Tweets

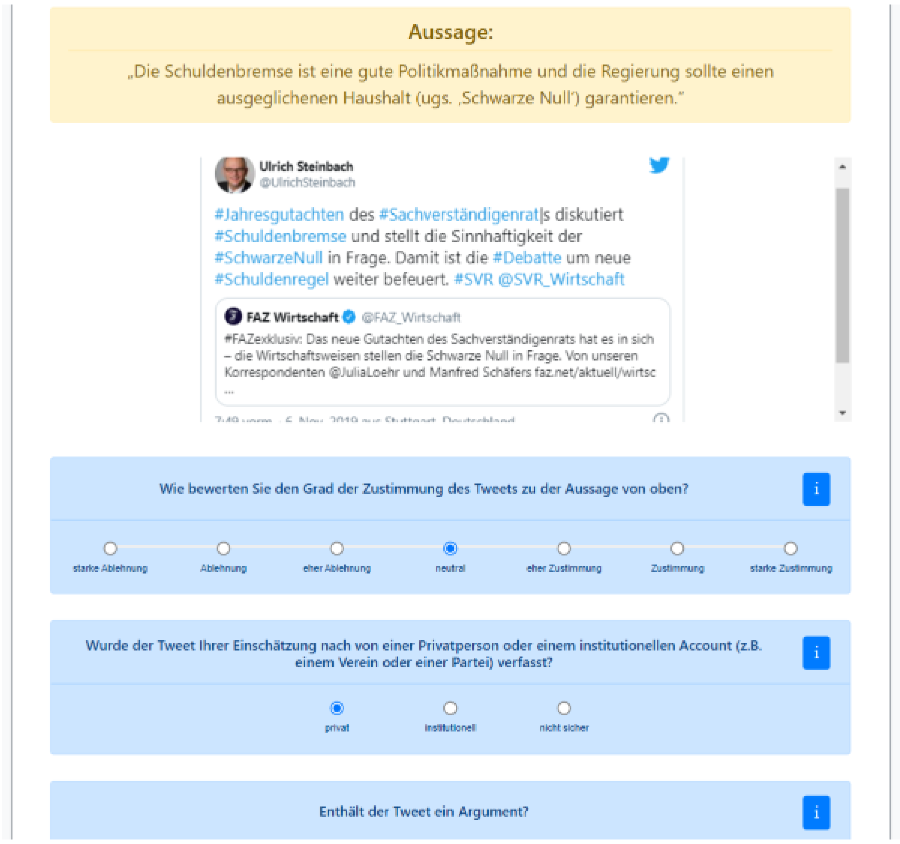



**Figure 3.B.1.** Work Screen for Task Assess Tweets

*Notes:* This figure shows a screenshot from the Assess Tweets work screen.

---

*[Only subjects working on Data Entry visited the subsequent set of pages. Subjects could freely navigate between the pages "Instructions", "Example", and "Let's Go!"—either via a navigation bar at the top end of the page, or via buttons on the bottom of the page.]*

### Instructions

*Reminder:* We are interested in how the proportion of female students at German universities has changed over time. The figures should be transferred from a scan of the statistical yearbooks into a table. You can use a zoom function within the scan.

*Your task in detail:*

- First, enter the semester in the designated field. Please use the *format "WS XXXX/XXXX"*.
- The universities are classified according to type of university and federal state. *Only transfer data for universities within the "Universities" category for the federal state of Rhineland-Palatinate.* You do not need to transfer the data for the other federal states. You can also neglect aggregated data, e.g. data which relate to the whole of Rhineland-Palatinate or to a group of universities.
- For every *university in Rhineland-Palatinate* the following information are required:
  - The name of the university as it appears in the document. *Please transfer the name of the university exactly as it is shown in the scan.*
  - The total number of all properly enrolled students and how many of them are women. *Please note that the order of the columns (and therefore the columns to be transmitted) may change from scan to scan.*

*In the example:*

On the next page you will find one of the scans for which the data has already been transferred:

- The semester (in the format "WS XXXX / XXXX")—WS 1982/1983.
- The names (exactly as printed in the scan) and the number of students (exclusively) for universities within the "Universities" category in Rhineland-Palatinate.

By clicking on the "short instructions" button, you can call up a short version of these instructions while you are working.

As soon as all relevant data have been transfered, you can continue with the next scan by clicking on "submit data", or you can end work by clicking on "submit data and finish work".

**Example**

*[Similar to Figure 3.B.2.]*

**Let's go!**

You can start working with a click on the "Let's go!" button.

### Data Entry

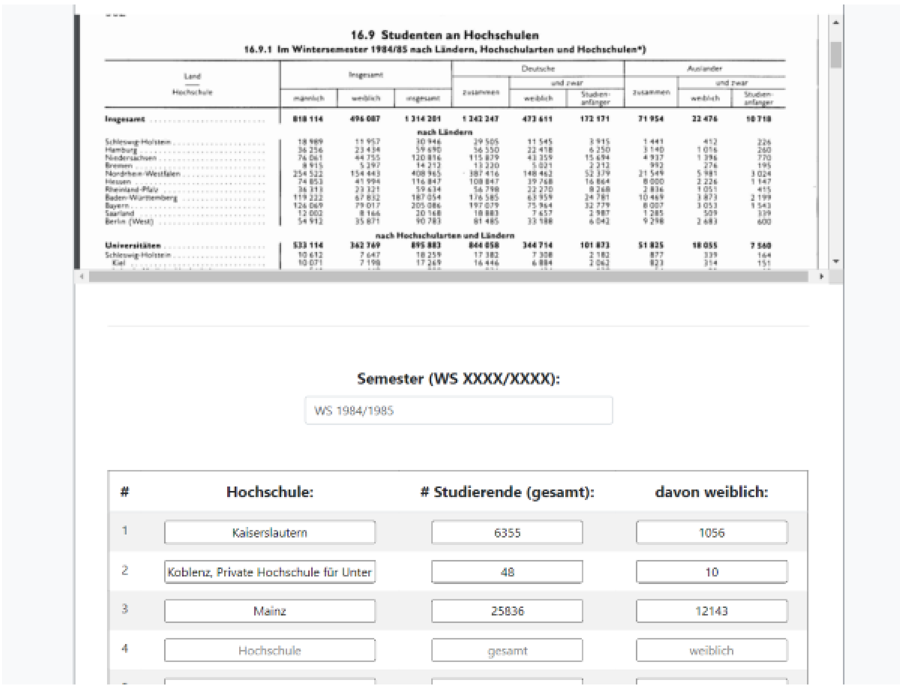



**Figure 3.B.2.** Work Screen for Task Data Entry

*Notes:* This figure shows a screenshot from the Data Entry work screen.

---

### What's Your Mood at the Moment?

Please let us know how your mood is at the moment.

`[input field (7-point Likert scale)]`

---

*[The questions on this page were displayed in random order and constitute the subsets of the "Interest/Enjoyment", "Effort/Importance", and "Perceived Choice" subscales from the Intrinsic Motivation Inventory. Below, reverse-coded questions are indicated by [R].]*

### Questionnaire

Finally, please take a moment to complete a short questionnaire.

In the following, you will be presented with a number of statements that relate to the task that you worked on during the work phase. For each of these statements, please indicate to what extent the statement applies to you personally.

*[Interest/Enjoyment]*

I enjoyed doing this activity very much.

`[input field (7-point Likert scale)]`

I thought this was a boring activity. [R]

`[input field (7-point Likert scale)]`

I would describe this activity as very interesting.

`[input field (7-point Likert scale)]`

*[Effort/Importance]*

I didn't try very hard to do well at this activity. [R]

`[input field (7-point Likert scale)]`

I tried very hard on this activity.

`[input field (7-point Likert scale)]`

It was important to me to do well at this task.

`[input field (7-point Likert scale)]`

*[Perceived Choice]*

I felt like it was not my own choice to do this task. [R]

`[input field (7-point Likert scale)]`

I did this activity because I wanted to.

`[input field (7-point Likert scale)]`

I did this activity because I had to. [R]

`[input field (7-point Likert scale)]`

---

**Thanks for Your Participation!**

That was it! Thank you for your help! As announced, your payment of 5 € will be transferred to you within the next two weeks.

You can close the browser window now.

## References

**Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman.** 2011. “Reference Points and Effort Provision.” *American Economic Review* 101 (2): 470–92. https://doi.org/10.1257/aer.101.2.470. [187]

**Aghion, Philippe, and Jean Tirole.** 1997. “Formal and Real Authority in Organizations.” *Journal of Political Economy* 105 (1): 1–29. https://doi.org/10.1086/262063. [186]

**Aghion, Phillippe, Nicholas Bloom, and John Van Reenen.** 2014. “Incomplete Contracts and the Internal Organization of Firms.” *Journal of Law, Economics, and Organization* 30 (SUPPL. 1): 37–63. https://doi.org/10.1093/jleo/ewt003. [186]

**Araujo, Felipe A., Erin Carbone, Lynn Conell-Price, Marli W. Dunietz, Ania Jaroszewicz, Rachel Landsman, Diego Lamé, Lise Vesterlund, Stephanie W. Wang, and Alistair J. Wilson.** 2016. “The Slider Task: An Example of Restricted Inference on Incentive Effects.” *Journal of the Economic Science Association* 2 (1): 1–12. https://doi.org/10.1007/s40881-016-0025-7. [187]

**Bartling, Björn, Ernst Fehr, and Holger Herz.** 2014. “The Intrinsic Value of Decision Rights.” *Econometrica* 82 (6): 2005–39. https://doi.org/10.3982/ECTA11573. [186]

**Beckmann, Michael, Thomas Cornelissen, and Matthias Kräkel.** 2017. “Self-Managed Working Time and Employee Effort: Theory and Evidence.” *Journal of Economic Behavior and Organization* 133: 285–302. https://doi.org/10.1016/j.jebo.2016.11.013. [185, 186, 203]

**Beckmann, Michael, and Matthias Kräkel.** 2022. “Empowerment, Task Commitment, and Performance Pay.” *Journal of Labor Economics,* forthcoming. https://doi.org/10.1086/718465. [186]

**Bénabou, Roland, and Jean Tirole.** 2003. “Intrinsic and Extrinsic Motivation.” *Review of Economic Studies* 70 (3): 489–520. https://doi.org/10.1111/1467-937X.00253. [184]

**Bloom, Nicholas, James Liang, John Roberts, and Zhichun Jenny Ying.** 2015. “Does Working From Home Work? Evidence From a Chinese Experiment.” *Quarterly Journal of Economics* 130 (1): 165–218. https://doi.org/10.1093/qje/qju032. [185, 186, 203]

**Bloom, Nicholas, and John Van Reenen.** 2011. “Human Resource Management and Productivity.” In *Handbook of Labor Economics,* 4: 1697–767. Elsevier. https://doi.org/10.1016/S0169-7218(11)02417-8. [186]

**Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch.** 2014. “Hroot: Hamburg Registration and Organization Online Tool.” *European Economic Review* 71: 117–20. https://doi.org/10.1016/j.euroecorev.2014.07.003. [191]

**Brandts, Jordi, David J Cooper, and Roberto A Weber.** 2015. “Legitimacy, Communication, and Leadership in the Turnaround Game.” *Management Science* 61 (11): 2627–45. https://doi.org/10.1287/mnsc.2014.2021. [186]

**Cassar, Lea, and Stephan Meier.** 2018. “Nonmonetary Incentives and the Implications of Work as a Source of Meaning.” *Journal of Economic Perspectives* 32 (3): 215–38. https://doi.org/10.1257/jep.32.3.215. [183, 185]

**Charness, Gary, Uri Gneezy, and Austin Henderson.** 2018. “Experimental Methods: Measuring Effort in Economics Experiments.” *Journal of Economic Behavior and Organization* 149: 74–87. https://doi.org/10.1016/j.jebo.2018.02.024. [188]

**Chen, Daniel L., Martin Schonger, and Chris Wickens.** 2016. “oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments.” *Journal of Behavioral and Experimental Finance* 9: 88–97. https://doi.org/10.1016/j.jbef.2015.12.001. [191]

**Cordella, Antonio, and Tito Cordella.** 2017. “Motivations, Monitoring Technologies, and Pay for Performance.” *Journal of Economic Behavior and Organization* 133: 236–55. https://doi.org/10.1016/j.jebo.2016.10.016. [203]

**Corgnet, Brice, Roberto Hernán-González, and Eric Schniter.** 2015. “Why Real Leisure Really Matters: Incentive Effects on Real Effort in the Laboratory.” *Experimental Economics* 18 (2): 284–301. https://doi.org/10.1007/s10683-014-9401-4. [187]

**Dal Bó, Pedro, Andrew Foster, and Louis Putterman.** 2010. “Institutions and Behavior: Experimental Evidence on the Effects of Democracy.” *American Economic Review* 100 (5): 2205–29. https://doi.org/10.1257/aer.100.5.2205. [186]

**DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao.** 2022. “Estimating Social Preferences and Gift Exchange at Work.” *American Economic Review* 112 (3): 1038–74. https://doi.org/10.1257/AER.20190920. [187]

**Dessein, Wouter.** 2002. “Authority and Communication in Organizations.” *Review of Economic Studies* 69 (4): 811–38. https://doi.org/10.1111/1467-937X.00227. [186]

**Dohmen, Thomas, and Armin Falk.** 2011. “Performance Pay and Multidimensional Sorting: Productivity , Preferences , and Gender.” *American Economic Review* 101 (2): 556–90. https://doi.org/10.1257/aer.101.2.556. [185]

**Duflo, Esther, Rema Hanna, and Stephen P. Ryan.** 2012. “Incentives Work: Getting Teachers to Come to School.” *American Economic Review* 102 (4): 1241–78. https://doi.org/10.1257/aer.102.4.1241. [203]

**Dutcher, E. Glenn.** 2012. “The Effects of Telecommuting on Productivity: An Experimental Examination. The Role of Dull and Creative Tasks.” *Journal of Economic Behavior and Organization* 84 (1): 355–63. https://doi.org/10.1016/j.jebo.2012.04.009. [185]

**Ertac, Seda, Mert Gumren, and Mehmet Y. Gurdal.** 2020. “Demand for Decision Autonomy and the Desire to Avoid Responsibility in Risky Environments: Experimental Evidence.” *Journal of Economic Psychology* 77: 102200. https://doi.org/10.1016/j.joep.2019.102200. [186]

**Falk, Armin, and Andrea Ichino.** 2006. “Clean Evidence on Peer Effects.” *Journal of Labor Economics* 24 (1): 39–57. https://doi.org/10.1086/497818. [185]

**Falk, Armin, and Michael Kosfeld.** 2006. “The Hidden Costs of Control.” *American Economic Review* 96 (5): 1611–30. https://doi.org/10.1257/aer.96.5.1611. [203]

**Fehr, Ernst, Holger Herz, and Tom Wilkening.** 2013. “The Lure of Authority: Motivation and Incentive Effects of Power.” *American Economic Review* 103 (4): 1325–59. https://doi.org/10.1257/aer.103.4.1325. [186]

**Frey, Bruno S., and Reto Jegen.** 2001. “Motivation Crowding Theory.” *Journal of Economic Surveys* 15 (5): 589–611. https://doi.org/10.1111/1467-6419.00150. [202]

**Gneezy, Uri, Lorenz Goette, Charles Sprenger, and Florian Zimmermann.** 2017. “The Limits of Expectations-Based Reference Dependence.” *Journal of the European Economic Association* 15 (4): 861–76. https://doi.org/10.1093/jeea/jvw020. [187]

**Goerg, Sebastian J., Sebastian Kube, and Jonas Radbruch.** 2019. “The Effectiveness of Incentive Schemes in the Presence of Implicit Effort Costs.” *Management Science* 65 (9): 4063–78. https://doi.org/10.1287/mnsc.2018.3160. [187]

**Holmström, Bengt.** 2017. “Pay for Performance and Beyond.” *American Economic Review* 107 (7): 1753–77. https://doi.org/10.1257/aer.107.7.1753. [184]

**Hubbard, Thomas N.** 2000. “The Demand for Monitoring Technologies: The Case of Trucking.” *Quarterly Journal of Economics* 115 (2): 533–60. https://doi.org/10.1162/003355300554845. [203]

**Kamei, Kenju, and Thomas Markussen.** 2022. “Free Riding and Workplace Democracy - Heterogeneous Task Preferences and Sorting.” *Management Science,* forthcoming. https://doi.org/10.1111/1468-0319.00131. [186, 202, 203]

**Kiessling, Lukas, Jonas Radbruch, and Sebastian Schaube.** 2021. “Self-Selection of Peers and Performance.” *Management Science,* https://doi.org/10.1287/mnsc.2021.4267. [186, 203]

**Kreps, David M.** 1997. “Intrinsic Motivation and Extrinsic Incentives.” *American Economic Review* 87 (2): 359–64. [187]

**Lazear, Edward P.** 2000. “Performance Pay and Productivity.” *American Economic Review* 90 (5): 1346–61. https://doi.org/10.1257/aer.90.5.1346. [185]

**Mas, Alexandre, and Enrico Moretti.** 2009. “Peers at Work.” *American Economic Review* 99 (1): 112–45. https://doi.org/10.1257/aer.99.1.112. [185]

**McAuley, Edward D., Terry Duncan, and Vance V. Tammen.** 1989. “Psychometric Properties of the Intrinsic Motivation Inventoiy in a Competitive Sport Setting: A Confirmatory Factor Analysis.” *Research Quarterly for Exercise and Sport* 60 (1): 48–58. https://doi.org/10.1080/02701367.1989.10607413. [191]

**Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor.** 2002. “Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment.” *American Economic Review* 92 (4): 850–73. https://doi.org/10.1257/00028280260344498. [203]

**Niederle, Muriel, and Lise Vesterlund.** 2007. “Do Women Shy Away From Competition? Do Men Compete Too Much?” *Quarterly Journal of Economics* 122 (3): 1067–101. https://doi.org/10.1162/qjec.122.3.1067. [185]

**Owens, David, Zachary Grossman, and Ryan Fackler.** 2014. “The Control Premium: A Preference for Payoff Autonomy.” *American Economic Journal: Microeconomics* 6 (4): 138–61. https://doi.org/10.1257/mic.6.4.138. [186]

**Pierce, Lamar, Daniel C Snow, and Andrew McAfee.** 2015. “Cleaning House: The Impact of Information Technology Monitoring on Employee Theft and Productivity.” *Management Science* 61 (10): 2299–319. https://doi.org/10.1287/mnsc.2014.2103. [203]

**Rohlfing-Bastian, Anna, and Anja Schöttner.** 2017. “Incentives and the Delegation of Task Assignment.” Working paper. https://doi.org/10.2139/ssrn.3047688. [186]

**Rohlfing-Bastian, Anna, and Anja Schöttner.** 2021. “Delegation of Job Design and the Value of Communication with Imperfect Performance Measurement.” Working paper. https://doi.org/10.2139/ssrn.3853904. [186]

**Ryan, Richard M.** 1982. “Control and Information in the Intrapersonal Sphere: An Extension of Cognitive Evaluation Theory.” *Journal of Personality and Social Psychology* 43 (3): 450–61. https://doi.org/10.1037/0022-3514.43.3.450. [191]

**Ryan, Richard M., and Edward L. Deci.** 1985. *Intrinsic Motivation and Self-Determination in Human Behavior.* Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-2271-7. [184, 186, 203]

**Ryan, Richard M., and Edward L. Deci.** 2000. “Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being.” *American Psychologist* 55 (1): 68–78. https://doi.org/10.1037/0003-066X.55.1.68. [184, 186, 203]

**Salop, Joanne, and Steven Salop.** 1976. “Self-Selection and Turnover in the Labor Market.” *Quarterly Journal of Economics* 90 (4): 619–27. https://doi.org/10.2307/1885325. [185]

**Stiglitz, Joseph E.** 1975. “The Theory of ‘Screening,’ Education, and the Distribution of Income.” *American Economic Review* 65 (3): 283–300. [185]

**Sutter, Matthias, Stefan Haigner, and Martin G. Kocher.** 2010. "Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations." *Review of Economic Studies* 77 (4): 1540–66. https://doi.org/10.1111/j.1467-937X.2010.00608.x. [186]

**Zuckerman, Miron, Joseph Porac, Drew Lathin, Raymond Smith, and Edward L Deci.** 1978. "On the Importance of Self-Determination for Intrinsically-Motivated Behavior." *Personality and Social Psychology Bulletin* 4 (3): 443–46. https://doi.org/10.1177/014616727800400317. [186, 203]