Three chapters on fairness preferences

Kjetil Røiseland Madland

Department of Economics

NHH Norwegian School of Economics

A thesis submitted for the degree of

Philosophiae Doctor (PhD)

Bergen 2022

Acknowledgements

First and foremost, I want to thank my supervisor, Bertil Tungodden, for outstanding help and guidance along the way to pursue my PhD. I am still, on a regular basis, astonished by your keen eye and quick mind. Your ability to quickly see further and dig deeper, has been an immense asset to me during these years. Thank you for your insightful questions, comments, and enthusiasm along the way. Thank you for believing in my projects even when I did not, and for pushing me to always improve. Furthermore, I want to thank my external supervisor, Botond Kőszegi, for insightful comments and guidance.

I also owe a special thank you to Erik Ø. Sørensen, for help and guidance with all things difficult. Thank you for generously lending your time and knowledge to help me with how to think about, and deal with, version control, research ethics, computer programming, and the statistical properties of estimators. These, and other important things in the life of an inexperienced researcher, would have been even more painful without you help.

To Eirik André Strømland, my co-author, friend, and long-time mentor, thank you for sharing the burden, and for fun and productive collaboration on the second chapter of this thesis. Thank you mentoring me all the way from my Bachelor's degree, and for motivating me to apply for a PhD position in the first place.

I would like to thank the entire administration team at FAIR and at the Department of Economics, for smoothing out the bumps along the way. Without you, life as a PhD student would have been even more frustrating. A particular thank you to Sebastian

Fest, for practical assistance with data collections.

I consider myself privileged to have been given the opportunity to pursue a PhD at the Department of Economics at NHH. The inclusive and inspiring environment at FAIR and at the department at large has meant a lot. I would like to express my sincere gratitude to all who have contributed to making this such a great environment to work and to do research. I want to express a special thanks to the other PhD students in my cohort: Charlotte, Mads, Mirjam and Oda. Your support and friendship have been absolutely invaluable through this process. I really have missed you a lot in my final year.

To Ada, thank you for your endless love and support. Pursuing a PhD has felt more like teamwork than a one-man show, and I consider myself extremely lucky to be on your team. Lastly, I am grateful to our wonderful children, Sverre, Eilif and Agnes, all born after I started the PhD program in 2017. Thank you for all the laughter, crying, early mornings, and paternity leave. Thank you for forcing me to practice a sound work-life balance, and for reminding me of what is most important in life.

Contents

The purely rational economic man is, indeed, close to being a social moron.

— Amartya Sen

INTRODUCTION

This thesis studies moral preferences and distributive decision-making. In almost all aspects of human life, we are faced with decisions involving fairness judgements and distribution. They often involve deciding what constitutes a fair allocation of scarce resources between individuals. Equally often, fairness considerations must be balanced against other considerations, such as self-interest or economic efficiency. In the early history of economics as a discipline, the economic agent was regularly presented as being fully rational and completely selfish. If asked to make a fairness judgement, this *homo economicus* would simply be indifferent. If pressed to balance concern for fairness, or the welfare of others, against their own self-interest, self-interest would prevail. Outside the discipline of economics, it is still a common misconception that the notion of a completely rational agent is the antithesis of an altruistic agent who cares about the well-being of others.

In the last few decades, great efforts have been made to rectify the image of the economic agent, and to present a more realistic one. Most of the influential work in this field does not attack the established models, but rather endeavors to build a more nuanced image of the economic agent. This work can be roughly split into two branches: First, there is a broad body of literature dedicated to the study of social preferences.[*] This literature has contributed to a richer understanding of the economic agent and their multiple objectives. The literature on social preferences does not reject the notion of optimizing behavior. Instead, it has taken advantage of the flexibility of

[*]See, e.g., Bolton and Ockenfels (2000); Camerer (2011); Fehr and Schmidt (1999); Rabin (1993)

the traditional models of economic choice. By allowing for richer objective functions, the new models of economic choice allow the economic agent to be both self-interested and altruistic, to care about their own consumption as well as that of others, and to also care about the difference between the two.

Second, there is a broad body of literature, in both economics and psychology, that deals with deviations from rational economic behavior. A vast body of literature has emerged following the introduction of experimental methods in behavioral sciences in general, and in economics in particular.[2] This literature aims to explain how and why agents in the real world deviate from the behavior prescribed by the axioms rationality in environments without frictions. This literature has largely focused on consumer decisions. Thus, our understanding of what affects distributive decisions and moral preferences remains limited.

This thesis consists of three individual chapters, which all contribute to understanding how fairness decisions are made. Through a series of four controlled experiments, studying a total of over 5000 individual responses, the three chapters study what affects our fairness preferences, what we focus on when making fairness decisions, and how fairness decisions differ from consumption decisions. The thesis endeavors to build a bridge between the literature on social preferences and the literature on deviations from assumptions of frictionless rationality, by studying deviations from rationality in distributive settings. The thesis contributes to our understanding of whether, and when, we should expect to see the same behavioral biases in fairness decisions as we often observe in consumer choices.[3] This is achieved by studying some known behavioral biases in distributive settings. The chapters included in this thesis

[2]See Chamberlin (1948); Huber and Puto (1983); Huber et al. (1982); Simonson (1989); Smith (1962, 1964, 1965) for some early applications of experimental methods in economics.

[3]See Bordalo et al. (2013); Dertwinkel-Kalt et al. (2022); Desvousges et al. (1993); Frederick et al. (2014); Heath and Chatterjee (1995); Huber and Puto (1983); Huber et al. (1982); Kahneman (2003); Ratneshwar et al. (1987); Simonson (1989); Simonson and Tversky (1993); Somerville (2020); Tversky (1972, 1977); Tversky and Kahneman (1981); Wedell and Pettibone (1996) for some empirical examples of behavioral biases in consumer choices.

ask the following questions:

- Chapter 1: Are preferences for fairness and efficiency path-dependent?
- Chapter 2: How does information about the social norms held by others affect people's distributive decisions?
- Chapter 3: How does the range of potential outcomes for individuals in a choice set affect which distributions are preferred?

The three chapters have some common features. For example, all three chapters make use of controlled, large-scale experiments. In all of the experiments, decision-makers take the role of impartial third-party spectators, who make decisions that affect unknown stakeholders. Spectator experiments have become popular in recent years, because they allow the researcher to study people's moral preferences directly, in situations where participants have nothing to gain from deviating from their fairness ideal.[4]

The experiments reported in this thesis also have a between-subject design in common. In all the experiments, allocation to treatment groups is randomized between individuals (Charness et al., 2012). This was an important feature of all the experiments reported here, due to fear of contamination. Because all four experiments study quite subtle effects, it has been critically important to keep as many variables as possible constant, and to avoid having respondents make several counting decisions consecutively. Together, the three chapters contribute to a better understanding of what affects distributive decision-making.

[4]See Aguiar et al. (2013); Almås et al. (2020); Cappelen et al. (2007, 2013, 2022c); Konow (2009); Konow et al. (2020); Lane (2022); Mollerstrom et al. (2015).

Chapter 1: Habits of equality: An experimental study of path dependence in fairness preferences

This chapter studies path dependence in distributive decision-making and moral preferences. The chapter contributes to both the literature on social preferences and the literature on behavioral anomalies, by studying whether or not distributive preferences are affected by previously encountered economic environments. Understanding whether moral preferences are path-dependent is important to the interpretation of many repeated economic games in the lab, as well as interactions in the real world. The chapter reports from two large-scale experiments with responses from a total of 2800 individual decision-makers. The participants in both experiments act as impartial spectators and make distributive decisions that affect the payments to stakeholders.

The first experiment studies whether or not people's preferences for efficiency and equality depend on the efficiency-equality trade-offs in prior decisions. In this experiment, 1400 spectators make eleven consecutive distributive decisions that affect bonus payments to anonymous pairs of workers. Spectators make binary choices. In all cases, one alternative is more equal than the other. For the first ten decisions, spectators face either a high efficiency cost of equality or no efficiency cost. The eleventh decision situation is the same for all spectators and entails a moderate efficiency cost for choosing the more equal alternative. Spectators who face a high efficiency cost implement less equality in the first ten decisions than compared to those who face no efficiency cost. This difference is partly maintained in a subsequent distributive situation, where the efficiency-equality trade-off is the same for all spectators. This shows that preferences for efficiency and equality depend on previously encountered efficiency costs of equality.

The second experiment studies how fairness preferences are affected by previously encountered sources of inequality. It bears some resemblance to the first experiment, but with some important distinctions. Instead of varying the efficiency cost between treatment groups, this experiment varies the source of inequality between the affected stakeholders. In the second experiment, there is no efficiency cost in any of the decisions. Spectators are allocated to one of two treatment groups for the initial decision rounds. In one treatment group, the *merit group*, the spectators are told that one of the workers performed better than the other, and is therefore given a higher bonus. In the other treatment, the *luck group*, spectators are told that both workers performed equally well, and that the allocation of the bonus is decided by luck. In a subsequent decision round, spectators in both groups face the same decision setting, which is equal to the merit setting. This experiment shows two things: First, spectators are much more willing to accept inequalities that are due to merit rather than luck. Second, spectators who are used to luck being the source of inequality are more likely to distribute bonus payments equally in a subsequent decision, where the source of inequality is merit.

Both of the experiments reported in this chapter identify path dependence in distributive preferences consistent with a strong habit effect of equality. This has implications for how we interpret repeated economic games in the lab, as well as interactions in the real world. The effects presented here are short-term effects that do not necessarily have much impact in the long run. How quickly path-dependence effects decay, and to what extent path dependence in distributive choices has lasting impacts on moral preferences, is a promising avenue for future research.

Chapter 2: Fairness of the crowd: An experimental study of social spillovers in fairness decisions

The second chapter is co-authored with Eirik André Strømland. It studies the effects of social norms on distributive decision-making. The motivation for this chapter is that people often wish to conform to the social groups that they are part of, and find it hard to behave in ways that stand out from the majority. Even in situations without a fixed group, or with weak social ties, simply doing what everybody else does may seem like the safe choice. The desire to conform to the views and values of our peers shapes our decisions in various domains. Because social norms vary across societies, identifying situations where social norms are important may help us to understand the sources of variation in behavior and institutions across societies.

The chapter reports from an incentivized spectator experiment with responses from a sample of 983 participants that is representative of the US population. The spectators distribute bonus earnings between a pair of workers. Before making their decision, the spectators receive information about the choices of a reference group of spectators from Almås et al. (2020), who have faced the same distribution decision. By drawing randomly from the data in Almås et al. (2020), we achieve random variation in the signal about the prevailing social norm that is provided to our spectators. We are thereby able to causally study the effect of a social norm signal on distributive decisions.

We find that spectators' distributive decisions are positively affected by the number of payoff-equalizing spectators in their reference group. This shows that people are affected by the social norms of others, even in an anonymous setting with weak and non-salient social ties. We also find that the redistribution choice has a strong positive correlation with spectators' initial beliefs about choices in the reference group. This is consistent with the false consensus effect. Our findings fit with the general idea that people may be sensitive to cues that their actions are socially acceptable, which is consistent with the findings of Bursztyn et al. (2020a,b). One possible interpretation of our findings is that people may care about their self-image - how they perceive themselves - in addition to their social image.

Chapter 3: Fairness and attribute range: An experimental study of range-based context effects on fairness decisions

The final chapter studies range-based context effects on distributive decisions. In standard economic models, preferences are assumed to be stable, and unaffected by the structure of the choice set. If an agent is observed to prefer alternative A over alternative B, this preference relation should be the same regardless of whether or not an alternative C is available. Understanding what agents consider to be fair and unfair, and which processes govern fairness judgements, is crucial in order to understand and predict a wide range of social interactions. Choice-set effects can have implications for how we study voting behaviour, political economy more generally, and charitable giving, among other topics. This paper studies the effects of focusing, attention, and relative thinking on people's fairness considerations.

The chapter reports from an online spectator experiment, motivated by the *Focusing model* of Kőszegi and Szeidl (2012) and the *Relative thinking model* of Bushong et al. (2021). The models predict that expanding the range of potential outcomes for one stakeholder will make the spectator emphasize the utility of this specific stakeholder more and less, respectively. The predictions from the two models are tested in a large-scale, incentivized spectator experiment in which 1400 spectators make distributive decisions for a group of real stakeholders. In four different treatment conditions, I add alternatives to the choice set, which expand the range of outcomes for one stakeholder in the group. All spectators provide complete preference rankings of all available choice alternatives. Thus, I am able to identify how the addition of an alternative C affects the relative rankings of alternatives A and B.

This chapter contributes to the behavioral literature on fairness and rationality, by testing well-documented models of consumer choice in distributive settings. Contrary to the predictions of both the focusing model and the relative thinking model, I observe no difference in which of the stakeholders is favored by spectators in either of the treatment conditions. My results suggest that either people are not as susceptible to choice-set effects when making fairness decisions as they are when making consumption decisions, or the choice-set effects only occur when choice dimensions are both clearly defined and not easily comparable. More research is needed in order to conclude. A lot still remains to be studied in order to fully understand how fairness decisions are made, and what decision-makers focus on in distributive decisions.

Chapter 1

Habits of equality: An experimental study of path dependence in fairness preferences

Kjetil Røiseland Madland*

Abstract

This paper studies the effect of choice-set history on distributive decisions made by impartial spectators. It presents findings from two large-scale incentivized spectator experiments, where participants distribute money between anonymous pairs of workers. Study 1 shows that willingness to pay a moderate efficiency cost for equality is negatively affected by the experienced efficiency cost in previous, unrelated decisions. This finding is consistent with a strong habit effect of implementing equality in previous decisions. Study 2 shows that people's willingness to accept inequalities that are due to merit

*FAIR, Department of Economics, NHH Norwegian School of Economics
I would like to thank Eirik André Strømland, Erik Ø. Sørensen, and Bertil Tungodden for their helpful comments and suggestions. I also owe a huge thanks to Sebastian Fest for invaluable help with the data collection. I would also like to thank FAIR, the Department of Economics at NHH, and the Research Council of Norway for the funding that has enabled this research. Study 1 was pre-registered in AEA's RCT registry on February 18, 2022 with ID AEARCTR-0008997. The data collection for the first experiment was completed March 17, 2022. Study 2 was pre-registered in AEA's RCT registry with ID AEARCTR-0010180. The pre-analysis plan was first registered October 6, 2022, and data collection was completed October 31, 2022. Both experiments were approved by the Institutional Review Board at NHH Norwegian School of Economics before the data collection.

depends on the source of inequality in prior decisions. I find that spectators who are used to luck as the source of inequality are more likely to distribute money equally in a subsequent merit setting, compared to spectators for whom merit is the source of inequality in all decisions. This is also consistent with a habit effect. Together, the two experiments show that fairness preferences are path dependent, and that the habit effect in the current distributive settings is stronger than the salience effect predicted by many models of reference-dependent behavior.

1.1 Introduction

How do the history of choices we make and the choice alternatives we face affect our subsequent decisions? This question is key to understanding how preferences and behavior develop over time. Understanding whether moral preferences are path dependent is important to the interpretation of many repeated economic games in the lab, as well as interactions in the real world. Path dependence in moral preferences can also potentially help us understand stable cross-country differences in economic inequality (Atkinson et al., 2011; Piketty et al., 2014), tax levels and institutions (Modica et al., 2018), and inter-generational transmission of fairness ideals (Almås et al., 2017).

There are several possible explanations for cross-country differences in inequality levels, including differences in preferences for redistribution (Almås et al., 2020; Falk et al., 2018; Henrich et al., 2005) and different beliefs about social mobility (Alesina et al., 2018). Lindbeck et al. (1999) and Alesina and Angeletos (2005) propose that countries may operate in different social equilibria, where beliefs about the sources of inequality are self-fulfilling. Another explanation for cross-country differences in inequality levels is that preferences are in part shaped by history and experiences, and that previously encountered distributive decisions affect future decisions. This avenue has been studied by Alesina and Giuliano (2011), and by Giuliano and Spilimbergo

(2008), who show that people exposed to macroeconomic volatility at a young age are more supportive of governmental redistribution, and attribute a greater portion of success to luck rather than to merit.

The present paper experimentally studies the role of path dependence in fairness preferences in two large-scale experiments with real incentives. In the two experiments Study 1 and Study 2, decision-makers are placed in an impartial spectator position.[2] In both studies, the spectators make eleven distributive decisions where they decide how to distribute money between pairs of stakeholders (hereinafter workers). A key feature of a spectator design is that it gives a direct expression of moral preferences, thus providing an opportunity to study whether moral preferences are path dependent. The two studies each investigate a distinct path-dependency of fairness preferences. One is related to how the trade-off between fairness and efficiency is handled. The other relates to how the source of inequality affects what is considered a fair outcome.

Study 1 identifies the role of choice-set history on trade-offs between efficiency and equality. This is done by varying the efficiency cost associated with implementing equality among anonymous workers. Throughout this paper, choice-set history refers to the menu of alternatives that have been available in previous decisions. In Study 1, spectators face either a high efficiency cost of equality, or no efficiency cost at all, in the first ten rounds of distribution. In the eleventh, final round, all spectators face a moderate efficiency cost of equality. In Study 1, the source of inequality is luck in all rounds for both treatment groups. Therefore, I expect that most of the spectators view equality as the fair outcome.

Study 1 addresses the following question:

[2]See e.g., Aguiar et al. (2013); Almås et al. (2020); Cappelen et al. (2007, 2013, 2022a,c); Croson and Konow (2009); Konow (2009); Konow et al. (2020); Lane (2022); Mollerstrom et al. (2015) for studies that use a spectator design.

- How does choice-set history affect spectators' trade-offs between efficiency and equality?

Study 2 investigates path-dependence in how the source of inequality determines what is considered a fair outcome. In this study, the difference between the two treatment groups is the source of inequality the spectators are presented with in the first ten distributive decisions. In Study 2, efficiency cost is kept at zero in all rounds for both treatments. In this study, spectators are assigned to either a *luck* treatment or a *merit* treatment group. The source of inequality between the workers is merit in the final round for both treatments in Study 2. The spectators are told that one worker has been allocated a higher bonus payment because they performed better than the other worker. In the first ten rounds, the source of inequality is luck in the luck treatment and merit in the merit treatment. When luck is the source of inequality, spectators are told that luck decided the bonus allocations.[3]

Study 2 addresses the following question:

- How does a previously encountered source of inequality affect spectators' fairness preferences?

Table 1.1 gives an overview of the key elements of the different treatments in Study 1 and Study 2.

Study 1 provides evidence for path dependence in fairness preferences, shaped by the previously experienced efficiency cost of equality. As expected, I firstly find that a high efficiency cost causes spectators to implement less equality in the initial rounds. In the final round, I identify path-dependent preferences for equality and efficiency, by

[3]There is no deception in either of the studies. Spectators get truthful information about the mechanisms that determine payment to the workers.

Table 1.1: Study design

	Study 1		Study 2	
Treatment group	**Low**	**High**	**Luck**	**Merit**
Efficiency-cost first ten rounds	**Zero**	**High**	Zero	Zero
Efficiency-cost final round	Moderate	Moderate	Zero	Zero
Source of inequality first ten rounds	Luck	Luck	**Luck**	**Merit**
Source of inequality final round	Luck	Luck	Merit	Merit

Note: The table gives an overview of the efficiency cost and the source of inequality in the choice situations. These are shown across decision rounds for each of the treatment groups in Study 1 and Study 2.

showing that those who faced a high efficiency cost in initial rounds also implement less equality than those who faced no efficiency cost, in a subsequent round, where efficiency cost is the same for all spectators.

Study 2 identifies path dependence in fairness preferences, shaped by the spectator's experienced source of inequality in previous decisions. In Study 2, the alternative worker pay-offs are identical between treatment groups in all eleven decisions. However, in the initial rounds, the source of inequality is different between the two treatment groups In the final decision round, the source of inequality between the workers is the same for both groups. As expected, inequality acceptance in the first ten rounds is much higher in the merit treatment than in the luck treatment. However, this difference partly carries over to the final round, in which the source of inequality is the same for both treatments. Hence, Study 2 identifies that the source of inequality experienced in previous choice settings can establish path dependence in fairness preferences.

The two studies investigate path dependence from different angles. In Study 1, the source of inequality is luck in all decisions. Here, I assume that spectators consider equality as the fair distribution. In this study, the efficiency cost of equality is different for the two treatment groups in the initial rounds. The results from Study 1 identify a path dependence in the weight the spectators place on fairness relative to efficiency. In Study 2, the efficiency is kept constant throughout. By changing the source of inequality in the initial rounds, I potentially manipulate the extent to which the spectators consider merit to be a fair source of inequality in the final round. Thus, Study 2 identifies path dependence in fairness perceptions, which in this setting specifically concerns which distribution is considered fair. Together, Study 1 and Study 2 provide compelling evidence that fairness decisions are path dependent.

In Study 1, I envisage two potential channels for path dependence: (i) A *salience effect*, as proposed by Bordalo et al. (2020). In their model, choice attributes such as price or quality are compared to previously encountered levels. Choice attributes that deviate more from the levels previously encountered are more salient than those that deviate less. Further, more salient attributes are over-weighted in decision-making. Because choice attributes are evaluated in relative terms, decision-makers overreact to sudden changes. In Study 1, the salience effect predicts that efficiency cost becomes salient in the final round, because it deviates much from previously encountered efficiency costs. Since efficiency cost is salient, spectators overreact to efficiency cost in the final round. In Study 2, the source of inequality is salient in the final round for spectators in the luck treatment, and is therefore over-weighted by these spectators.

(ii) A *habit effect*. Repeatedly implementing equality could create a habit of implementing equality. By getting used to a certain level of equality, this level is established as a fairness norm, making deviations from it more costly over time (Alvarez-Cuadrado et al., 2004; Becker and Murphy, 1988; Carroll et al., 2000; Constantinides, 1990; Dynan, 2000; Loewenstein et al., 2003; Messinis, 1999; Peysakhovich and Rand, 2016). It would be natural to expect that in the Study 1 group where equality is costly, less equality is implemented. Thus, the habit effect predicts that spectators in this group will continue to implement less equality than others, even when efficiency costs are equalized between the groups. In the initial rounds in Study 2, more equality is implemented in the luck treatment than in the merit treatment. This can create a habit for equality, which can carry over to the final round.

There are two main findings in this paper: First, Study 1 shows, in a third-party spectator game, that inequality acceptance is affected by previously encountered efficiency costs of implementing equality. Second, Study 2 shows that previously encountered source of inequality affects inequality acceptance in a subsequent fairness decision where merit is the source of inequality. Both findings are consistent with a strong habit effect. This does not necessarily rule out a salient effect but in the current experimental settings, the habit effect dominates.

To my knowledge, this paper is the first to study path dependence in the context of distributive decision-making. The paper contributes to several important literatures. Firstly, the paper contributes to the literature on path dependence in decision-making (Epstein et al., 2016; Peysakhovich and Rand, 2016; Peysakhovich et al., 2014; Reigstad et al., 2017). The paper that is closest to the present study is Peysakhovich and Rand (2016), who find evidence of path dependence in social preferences, specifically norms of cooperation. The authors construct environments in the lab that support cooperation and defection, respectively, in a prisoner's dilemma game. Their experiments show that the imposed structure affects play in subsequent one-shot games. They find that cooperation in strategic games can be sustained even after removing the structures that initially promoted cooperation. The present paper adds to these findings by showing path dependence in distributive preferences in the absence of self-interest and strategic interaction. Using a spectator design, Study 1 allows for clean identification of the effect of previously encountered efficiency costs on subsequent distributive decisions. Study 1 shows that the economic environment in previous settings can affect the relative importance of fairness and efficiency in a subsequent distributive setting. Study 2 shows that what is considered the most fair distribution is also path dependent. Specifically, I demonstrate that it can be affected by previously encountered sources of inequality.

The paper also contributes to a broad literature on fairness views and moral preferences (Alesina and Angeletos, 2005; Alesina and Giuliano, 2011; Almås et al., 2020, 2010; Andreoni and Miller, 2002; Bolton and Ockenfels, 2000; Cappelen et al., 2007, 2013, 2019, 2022c; Charness and Rabin, 2002; Croson and Konow, 2009; Durante et al., 2014; Engelmann and Strobel, 2004; Falk et al., 2018; Fehr and Schmidt, 1999; Fehr and Gächter, 2000; Henrich et al., 2005; Konow, 2000, 2009; Konow et al., 2009; Rabin, 1993) by showing that fairness preferences are path dependent. Further, I find that decision-makers who act as impartial spectators react strongly to the source of inequality, but only modestly to the efficiency cost of equality. This is consistent with the findings of Almås et al. (2020) and Engelmann and Strobel (2004).

Finally, this paper contributes to the literature on memory and associative recall (see, e.g., Bodoh-Creed (2020); Bordalo et al. (2020); Enke et al. (2020); Mullainathan (2002); Schwartzstein (2014); Wachter and Kahana (2019)), and to the literature on habit formation (Alvarez-Cuadrado et al., 2004; Becker and Murphy, 1988; Carroll et al., 2000; Constantinides, 1990; Dynan, 2000; Loewenstein et al., 2003; Messinis, 1999), by studying how spectators react to the choice-set history in distributive choice settings. The theoretical framework outlined in section 1.2 of this paper presents mechanisms that are close in spirit to the habit model of Carroll et al. (2000) and the model of salience and associative recall of Bordalo et al. (2020). An important prediction by Bordalo et al. (2020) is that agents overreact when prices are much higher than they are used to. Thus, in a redistributive setting, their theory predicts that if spectators have repeatedly faced either a high efficiency cost, or no efficiency cost, those who have faced a high cost are more likely to accept a moderate cost in subsequent decisions. This is because a moderate efficiency cost compares favorably to the high cost previously encountered. The predictions based on the theory of Bordalo et al. (2020) are supported by Simonsohn and Loewenstein (2006), who show path dependence in willingness to pay for housing. Their findings are consistent with both salience theory and other models of reference dependence.[4] The experiments reported in this paper do not provide any support for salience theory.

Various models of habit-dependent utility assume that the utility from present consumption is augmented by previous consumption. I adapt this framework to fit a context where spectators make distributive decisions for anonymous workers. I show that my findings are consistent with a static model in which the spectator's fairness ideal is determined by choices implemented in previous fairness decisions.

The rest of the paper is organized as follows: Section 1.2 outlines a theoretical framework to guide the analysis. Section 1.3 gives a thorough overview of the experimental design and the participants involved. Section 1.4 presents the experimental strategy and identification, with a description of the regressions specified in the pre-analysis plan. Section 1.5 presents the results from both of the experiments. Lastly, section 1.6 concludes.

1.2 Theoretical framework

This section provides a brief theoretical framework from which to organize the idea of path dependence in distributive decisions made by a third-party spectator. As a starting point, I present a path-independent model of moral preferences, and then add to it two opposing path-dependent forces. I borrow the path-independent model of

[4] See e.g., DellaVigna et al. (2017); Kahneman and Tversky (1979); Kőszegi and Rabin (2006); Smith (2019); Sugden (2003)

moral preferences from the spectator framework presented in Cappelen et al. (2013) and Almås et al. (2020). Their model assumes that the spectator has preferences for efficiency and fairness, and chooses a distribution that represents an optimal trade-off between the two.

Assume that a spectator decides how to distribute money between a pair of workers, A and B. The spectator chooses a distribution, allocating a non-negative share y to worker B, and $1 - y$ to worker A. In all situations, let A be the worker with the highest initial income. Let $c(j)$ be the efficiency cost of redistributing income to worker B in treatment j. In any given distributive choice setting, $m(j)$ is what the spectator considers fair to give to B in treatment j. The spectator trades off deviations from what they consider a fair distribution and the associated efficiency cost. The spectator's utility is captured by the following function:

$$V(y) = -\frac{\beta}{2}(y - m(j))^2 - c(j)y, \tag{1.1}$$

where β is a non-negative weight on fairness relative to efficiency, and $c(j)y$ is the efficiency cost of giving share y to worker B.

The interior solution of the path-independent spectator's optimization problem is to give the following share y to worker B:

$$y^* = m(j) - \frac{c(j)}{\beta} \tag{1.2}$$

Because decisions are binary in both studies, the spectator chooses the distribution which yields the highest utility of the two. In Study 1 in this paper, all spectators face an efficiency cost $c(j)y$ in the final round. According to the path-independent model, spectators choose the distribution that is closest to an optimal trade-off between equality and efficiency, regardless of the choice sets they faced in previous decisions.

In Study 2, there is no efficiency cost, so $c(j) = 0$ in all decisions. Here, the path-independent model predicts that spectators optimally choose the distribution they regard as most fair.

I depart from the model employed in Cappelen et al. (2013) and Almås et al. (2020) by introducing two opposing, path-dependent forces: a *habit effect* and a *salience effect*. The path-dependent model presented here is a static model. It describes spectator choice at a specific point in time, where habit and salience are determined by prior choice sets and decisions. The *habit effect* works as follows: Spectators who have given a certain share y to worker B in the past develop a "habit stock" for what is considered B's fair share $m(j)$. The habit stock of $m(j)$ can be thought of as a weighted average of previous allocations to worker B. This habit stock represents a level of $m(j)$ from which the spectator becomes reluctant to deviate. If the spectator has habit-dependent utility of y, they experience a utility loss when giving a share y to worker B that deviates from their habit stock, which I define as y_h. Following Carroll et al. (2000), assume that the habit stock develops over time according to the following function:

$$\dot{y_h} = \rho(y - y_h) \tag{1.3}$$

where $\dot{y_h}$ is the change in the habit stock over time. ρ is a parameter that determines the weight of the share given to B at different times. A high ρ represents a high weight on the share given to B in the more recent past. It is assumed that $0 \leq \rho \leq 1$, i.e., that the habit effect decays over time.

The *salience effect* works as follows: When a choice attribute, such as efficiency cost, fairness, or the source of inequality, deviates from a familiar range, it becomes more salient. More salient attributes are given higher decision weights than less salient attributes. In the model presented above, this can work through two channels: First, salience can affect β, and second, it can affect the weighting of dimensions that determine worker

B's fair share $m(j)$, including the source of inequality.[5] In the specific context of this paper, I make the following assumptions: 1) The value of a choice attribute is evaluated not only in absolute terms, but also relative to previously encountered values of that attribute. 2) Choice attributes are more salient when their values deviate more from previously encountered values for the same attributes.[6]

In the final distributive decision in Study 1, all spectators face a moderate efficiency cost. The spectators in the high treatment are familiar with a high efficiency cost, while the spectators in the low treatment are familiar with no efficiency cost. Since the efficiency cost in the final round is very different from what the spectators are familiar with from the first ten rounds, efficiency cost is salient in the final round for the spectators in both treatment groups. The salience of efficiency cost has different predicted effects for the two treatment groups. For spectators in the high treatment, the moderate efficiency cost in the final round compares favorably to the recently encountered high efficiency costs. For spectators in the low treatment, the moderate efficiency cost in the final round compares unfavorably to the zero efficiency cost in the first ten rounds. By comparison with previously encountered values, the salience effect predicts that spectators in the high treatment implement more equality in the final round than the spectators in the low treatment.

In Study 2, the source of inequality is salient in the final round for spectators in the luck treatment, but not for spectators in the merit treatment. Since merit is salient in the final round for the luck treatment, spectators in this treatment overreact to merit as the source of inequality. In the model, this causes $m(j)$, the fair share to the worst-performing worker (B), to decrease. As such, salience predicts that spectators in the luck treatment in Study 2 implement less equality in the final round compared to spectators in the merit treatment. This is consistent with Cappelen et al. (2022b).

[5]See Bordalo et al. (2012, 2013, 2020); Bushong et al. (2021); Enke et al. (2020); Kahneman and Miller (1986); Kőszegi and Szeidl (2012); Kőszegi and Rabin (2006) for models of attribute salience.

[6]This understanding of attribute salience is closest to that described in Bordalo et al. (2020).

They find that, when inequality is caused partly by luck and partly by merit, spectators overreact to just a little bit of merit.

I extend the fairness model presented in Cappelen et al. (2013) and Almås et al. (2020) by including the habit effect and the salience effect in a static model. I assume that choosing a distribution with an equality level that deviates from the level dictated by the habit stock for fairness is costly. The habit stock is indicated by y_h, which is the share given to B dictated by the habit stock of fairness, disregarding all other considerations. y_h is the share allocated to B that gives no utility loss from the habit effect. The salience effect is a determinant of the relative weight on fairness, β, and of worker B's fair share $m(j)$. Define a salience vector $\vec{s} = s_1, ..., s_K$, where s_k indicates the salience of choice attribute k. In the Study 1 setting, β is affected by the relative salience of efficiency cost and fairness. Specifically, β increases in the salience of fairness relative to efficiency cost. The fair share for worker B, $m(j)$, is affected by the specific setting in treatment j, but also by the salience of this setting. In Study 2, this includes the source of inequality between the workers. If merit as the source of inequality becomes more salient, spectators accept more inequality when it is due to merit rather than luck. This reduces worker B's fair share $m(j)$. For a spectator with path-dependent preferences for fairness and efficiency, the utility is captured by the following function:

$$V(y) = -\frac{\beta(\vec{s})}{2}(y - m(j, \vec{s}))^2 - c(j)y - \frac{\gamma}{2}(y - y_h)^2, \tag{1.4}$$

where $0 \leq \beta(\vec{s}),\ \ 0 \leq \gamma$.

Here, $\beta(\vec{s})$ is a function of attribute salience $\vec{s}$. Whether $\beta(\vec{s})$ is greater or smaller than β from the path-independent model, depends on which attributes are salient in the specific context. γ is the relative weight on deviations from the share to worker B dictated by the habit stock of fairness, y_h. B's fair share $m(j, \vec{s})$ is a function of both the treatment j and the attribute salience $\vec{s}$. In a given distributive setting, the spectator with path-dependent preferences optimally allocates the following share to worker B:

$$y_{pd}^* = \frac{\beta(\vec{s})m(j,\vec{s}) + \gamma y_h - c(j)}{\beta(\vec{s}) + \gamma} \tag{1.5}$$

y_{pd}^* is the optimal share to worker B for a spectator with path-dependent preferences. A spectator whose preferences depend on a habit stock for fairness, but not on salience, deviates more from the fair share $m(j)$ and cares less about efficiency cost $c(j)$. When the fair share $m(j,\vec{s})$ and its relative weight $\beta(\vec{s})$ depend on attribute salience, the distribution also depends on which attribute is salient in that context. In Study 1 in this paper, efficiency cost is made salient in the final round by deviating from previously encountered values. This is the case for both treatments. Increasing the salience of efficiency cost lowers $\beta(\vec{s})$, which is the relative decision weight on fairness. In Study 2, the source of inequality is made salient for spectators in the luck treatment, but not for those in the merit treatment. For those in the merit treatment, the final round is familiar in all respects, and thus no choice attribute is particularly salient. For spectators in the luck treatment, merit as the source of inequality is unfamiliar, and deviates from what they have encountered in the previous ten rounds. This makes the source of inequality stand out as salient. Since the source of inequality is made salient, spectators focus more on the fact that worker A is allocated a higher bonus because A performed better than B. Therefore, spectators are more willing to accept inequality that favors worker A. This entails a decrease in B's fair share $m(j,\vec{s})$.

Consider the following example: A spectator chooses between the following two distributions of money to worker A and worker B:

- 70 to A and 10 to B
- 40 to A and 20 to B

Assume that the spectator has made a series of distributive decisions prior to this one. A spectator in the high treatment in Study 1 have made decisions where the efficiency cost has been higher than in the present setting. Since the spectator in the high treatment has faced a high efficiency cost in prior decisions, this spectator has implemented less equality, i.e., a lower y, and thus has a lower habit stock y_h, compared to the spectator in the low group. Because the habit stock is lower, the habit effect dictates that the spectator in the high treatment allocates less to worker B in the final round, i.e., implements less equality. The salience effect goes in the opposite direction. The efficiency cost in the final round is salient because it deviates from the efficiency costs the spectator is used to in similar settings. Further, for the spectator in the high treatment, the efficiency cost in the final round compares favorably to previously encountered efficiency costs. This makes the spectator more likely to pay the moderate efficiency cost to equalize earnings in the current setting. For a spectator in the low treatment, the efficiency cost in the current setting is much higher than the spectator is used to. Again, the efficiency cost is salient, but for a spectator in the low treatment, it compares unfavorably to the efficiency cost previously encountered. Salience dictates that the spectator from the low treatment in Study 1 is less likely to choose the more equal distribution.

Consider now the following setting in Study 2: A spectator chooses between the following two distributions of money to worker A and worker B:

- 70 to A and 10 to B
- 40 to A and 40 to B

Here, both distributions are equally efficient. The spectator is told that the first distribution is the default because worker A performed better than worker B. Given that inequality acceptance is higher when merit rather than luck is the source of inequality, spectators in the luck treatment equalize more in the first ten rounds. The habit effect therefore dictates that the spectator in the luck treatment has a higher habit stock y_h from prior decisions. Thus, through the habit effect, spectators in the luck treatment are more likely to choose the equal outcome in the current setting, compared to spectators in the merit treatment. Again, the salience effect works in the opposite direction. For a spectator in the merit treatment, the choice setting in the final round is familiar in terms of fairness, source of inequality, and the absence of efficiency cost. Thus, none of these choice attributes are salient in the current setting. For a spectator in the luck treatment, the source of inequality is salient in the current setting. In the prior decisions made by spectators in the luck treatment, the inequality is caused by luck. For this reason, merit as a source of inequality stands out in the current setting. I assume that greater salience of merit as the source of inequality makes spectators consider it fair to give an even higher amount to the best-performing worker. This implies a reduction in B's fair share $m(j, \vec{s})$ compared to previous decisions.

1.3 Experimental design and participants

I conducted two incentivized spectator experiments with a between-subject design. In both experiments, the participants acted as third-party spectators and made a series of eleven distributive decisions for pairs of anonymous stakeholders (workers). In each decision, the spectators chose between two alternatives, where one was a very unequal distribution and the other was closer to equality.

This section begins by describing the different participant groups in the two experiments, hereinafter Study 1 and Study 2. Subsections 1.3.3 and 1.3.4 provide a detailed outline of Study 1 and Study 2, and their respective treatments. In both experiments, participants were recruited from online labor platforms, workers from Amazon Mechanical Turk (MTurk) and spectators from Prolific. For simplicity, and particularly to avoid language issues, I only recruited respondents with an American IP address.

The workers and spectators were directed to a Qualtrics survey, where one was for the spectators and one for the workers. In both experiments, the worker part of the experiement was conducted a few days before the spectator part. This was necessary to ensure the spectators were given truthful information about the workers, and to have the spectator decisions dictate payment for the workers. The spectator decisions were implemented in a probabilistic way, such that 10 percent of the spectators (140 from each study) had one of their decisions implemented for a real pair of workers. The spectators were told that there was some probability that one of their choices would be implemented for a real pair of workers. They were also told that each decision was equally likely to be chosen for payment. Finally, they were asked to regard each decision as involving a new, potentially unique, pair of workers.

1.3.1 The workers

For Study 1, 280 workers were recruited from MTurk to take part in a Qualtrics survey. The workers in Study 1 answered a survey containing hypothetical distributive decisions. The task was of no significance, other than creating real incentives for the spectators by having real stakeholders between whom the money was to be distributed. The workers received $\$1$ as payment for participation, and were told that they could earn an additional bonus that would be decided by chance and by someone who was not taking part in the same survey. On average, the workers in Study 1 received $\$4,23$ in bonus payment, bringing the total to $\$5,23$.

For Study 2, I recruited 279 workers from MTurk. The Study 2 workers performed a set of slider tasks. Workers were randomly allocated to either a luck group or a merit group. Workers in the luck group were told to set all the sliders to their correct values. They were further told that they would be randomly matched with another worker, and that one of them would be allocated a high bonus and the other a low bonus. Workers in the luck group were told that the bonus allocation was decided by luck. Workers in the merit group were asked to set the sliders to their correct values as quickly as possible. They were further told that they would be randomly matched with another worker, and that the worker with the quickest time would earn a high bonus while the other would earn a lower bonus. All workers were told that someone from outside the study would have an opportunity to alter the bonus distribution. The full instructions given to the workers in Study 1 and Study 2 are provided in Appendix 1.B and Appendix 1.D, respectively. The workers in Study 2 received $\$1$ as payment for participation, and earned on average $\$4.29$ in bonus payment, bringing the total to $\$5.29$.

1.3.2 The spectators

For the spectators in Study 1, I recruited 1,396 individuals from Prolific, an online labor platform that specializes in surveys and research. To ensure high data quality, I included an attention checker in the experiment, where subjects had to pick one out of five cities in accordance with the instructions in the question text. Thirty-three of the spectators, 2.4% of the sample, got the attention check wrong. These were dropped from the main analysis. Closer investigation strengthened the suspicion that these subjects more or less randomized their answers throughout the experiment. Two more observations were dropped because of missing treatment variable information.

The experiment in Study 2 has a spectator sample size of 1,388 participants, all of whom passed an attention check. Age data is missing for ten of the spectators. In both studies, the spectators received $\$1.10$ for participating. The complete instructions given to the spectators in Study 1 and Study 2 are presented in Appendix 1.A and Appendix 1.C, respectively.

Table 1.2 presents sample characteristics for the spectators in both Study 1 and Study 2, compared to the general population in the USA. Treatment is randomized between subjects, and all observable characteristics are balanced across treatment groups in both experiments. The sample in Study 2 was stratified to be representative of the grown-up population in the US on age, sex, and ethnicity. The Study 1 sample is more highly educated and has a lower female share than the general population. Data for political views are from different questionnaires, and are therefore hard to compare to the general population.

Table 1.2: Sample characteristics

	Study 1	Study 2	US population
Age median	35	46	38.3
Female share	0.37	0.51	0.51
College share	0.82	0.48	0.43
Conservative share	0.44	0.43	0.36
Observations	1,361	1,377	

Note: Subjects failing the attention check are dropped from the sample. "Age median" is median age in years. "Female share" is the proportion of females. "College share" is the share with at least a two-year college degree. Age, gender, and education data for the US population are from the US Census Bureau (2021). "Conservative share" in Study 1 and Study 2 samples indicates an answer of 4 or 5 on a five-point scale to the question "Would you describe yourself as politically on the "left" (i.e. a liberal) or on the "right" (i.e. a conservative)?", where 1 is very liberal and 5 is very conservative. "Conservative share" in US population is from Gallup, asking "How would you describe your political views – [very conservative, conservative, moderate, liberal or very liberal]?", including answers "conservative" and "very conservative" (Gallup, 2022).

1.3.3 Study 1 design

The spectators in Study 1 were divided randomly into two treatments, a high treatment and a low treatment. Spectators in both treatments made eleven binary decisions where they decided how to distribute experimental tokens between a pair of anonymous workers. One token equates to 10 cents in final payment to the workers. It was specifically stressed to the spectators that their decisions involved distributing real money to real people, and that it was important that they took the decisions seriously. In the high treatment group, there was a high efficiency cost of choosing the more equal distribution in the first ten rounds, while in the low treatment group, there was no efficiency cost in the first ten rounds. The eleventh, final, decision was the same for both treatments.

In all eleven decisions, alternative 1 is a distribution with high inequality and is the same in both treatments. In the first ten decisions, alternative 2 is more equal than alternative 1 for both treatments. In the high treatment, alternative 2 is also much less efficient in terms of total payment to the two workers combined. Specifically, the more equal distributions represent a five-for-one trade-off relative to the more unequal one. In other words, going from the most unequal to the most equal distribution means that the worker with the highest endowment loses five tokens for every token gained by the other worker. Full equality is possible in just a few of the choice settings. The spectators can never reverse the relative ranking of the two workers in terms of bonus income. In the low treatment, the two distributions in each choice setting are equally efficient in terms of total pay to the workers for all the first ten decisions. Tokens are distributed between workers one-for-one. To illustrate, consider the following example from Study 1. In the eleventh choice setting, spectators in both treatments face the same alternatives with the same efficiency cost of choosing the more equal alternative. The spectators' instructions were: "Please select your preferred distribution of tokens to worker A and worker B." They were given the following distributions to choose from:

- Alternative 1: 90 to A and 15 to B
- Alternative 2: 45 to A and 30 to B

Here, the more equal alternative 2 represents a three-for-one trade-off relative to alternative 1. Efficiency, defined as the total amount of tokens given to the two workers, is 105 in alternative 1 and 75 in alternative 2 in the final round. The difference between the two treatment groups is the alternatives they had been given in previous decisions. The following is an example of one of the ten previous decisions:

- Alternative 1 (the same for both treatments): 80 to A and 10 to B
- Alternative 2 for high treatment: 30 to A and 20 to B
- Alternative 2 for low treatment: 50 to A and 40 to B

In the first ten rounds, the average total pay-off is 93 in the least equal alternatives, and is the same in both treatments. In the low treatment, the average total bonus payment is the same for the more equal distribution alternatives. In the high treatment, total efficiency in the most equal distribution alternatives is on average 61.[7]

The spectators were told that the pair of stakeholders had performed a number of unspecified tasks and had been told that they could earn a bonus, which was to be decided by chance and by somebody else.[8] The spectators were also told that each worker in any given pair had performed equally well, but received no further information about the workers. Without more information, it is reasonable to assume that most spectators consider equality as the most fair distribution. For example, Cappelen et al. (2013) and Almås et al. (2020) find that meritocracy is by far the most prevalent fairness ideal in the US, and that most people consider inequalities due to luck to be unacceptable. In this case, B's fair share $m(j)$ is half of the total in the path-independent model. If preferences are path dependent, the fair share $m(j, s)$ is also a

[7]In the final decision, the Gini coefficients for the two alternatives are 0.36 and 0.1, respectively. The average Gini coefficient for alternative 1, the most unequal distribution, in the first ten decisions is 0.34. This is the same for both treatments. For alternative 2, the Gini coefficient is 0.12 in the high treatment and 0.05 in the low treatment.

[8]It was not specified to either the workers or the spectators exactly how "chance" was to be understood in the specific context.

function of attribute salience. Salience can affect the fair share upwards or downwards, depending on which choice attribute is salient in the current choice context.

The aim of this design is to, as cleanly as possible, study whether distributive choices are path dependent in the sense that present decisions are affected by past decisions. The design of Study 1 also aims to specifically capture the trade-off between fairness and efficiency, and how this may be path dependent.

1.3.4 Study 2 design

Study 2 aims to investigate the effect of habits and salience on what the spectator thinks is a fair distribution, keeping efficiency considerations the same across treatments. Specifically, Study 2 addresses how spectators think about the source of inequality between workers, and whether or not this is affected by previously experienced source of inequality. An effect in Study 1 can be driven by either a change in $\beta(\vec{s})$, the relative weight on fairness, or by a change in what the spectator thinks is worker B's fair share, $m(j, \vec{s})$. Study 2 is set up such that efficiency is the same for all spectator decisions. Thus, a treatment effect can only go through the spectator's fairness ideal. The Study 2 experiment was run on the same platforms as Study 1. Workers/stakeholders were recruited from MTurk and spectators were recruited from Prolific.

Study 2 is set up in the same way as Study 1 in that the spectators made a series of eleven distributive decisions for pairs of anonymous workers. The main difference between the studies is that instead of varying the efficiency cost in the first ten rounds, I vary the source of inequality in Study 2. There is no efficiency cost of equality for either group of spectators, as shown in Table 1.1. The spectators were told that the bonus had initially been divided unequally between the pair, and the spectators could choose to keep the unequal distribution or to equalize bonus earnings between the pair of workers. In Study 2, one of the alternatives in each decision represented equal payment to each worker in the pair. These are two additional differences from Study 1.

The spectators were divided into two treatment groups. The difference between the two groups is the source of the initial inequality between the workers in the first ten rounds. Half of the spectators were placed in the *merit* treatment. Here, worker A is allocated a larger bonus because they performed better than worker B. The other half of the spectators were placed in the *luck* treatment. In the luck treatment, the source of inequality is luck in the first ten decisions and merit in the final decisions. In the final decision round, both treatment groups were given the following task:

"In this pair, worker A performed better than worker B, and is allocated a higher bonus. Please select your preferred distribution of tokens to worker A and worker B."

- Alternative 1: 70 to A and 10 to B
- Alternative 2: 40 to A and 40 to B

In the first ten rounds, spectators in both the merit treatment and the luck treatment were given the same pay-off alternatives, but the sources of inequality were different between treatments. The merit group received the following information:

"In this pair, both workers performed equally well. By luck, worker A was allocated a high bonus of 60, while B gets 20. Please select your preferred distribution of tokens to worker A and worker B."[9]

- Alternative 1: 60 to A and 20 to B
- Alternative 2: 40 to A and 40 to B

[9]The order in which worker A and B appear in the to participants, and whether A or B is initially allocated the higher bonus, is randomized. For simplicity, in this paper A is always the worker with the highest income.

Spectators in the merit group were given the following information: "In this pair, worker A performed better than worker B, and therefore is allocated a higher bonus of 60, while B gets 20. Please select your preferred distribution of tokens to worker A and worker B."

- Alternative 1: 60 to A and 20 to B
- Alternative 2: 40 to A and 40 to B

The only difference between the two treatment groups is the source of inequality between the workers in the first ten rounds. Study 2 aims to capture (i) the habit effect. Given that people accept more inequality when inequality is due to merit rather than to luck, spectators in the luck group equalize more in the first ten rounds. Thus, entering the final round, the luck group has a higher habit stock y_h than the merit group. (ii) The salience effect. Because the merit in the final round is more salient to the luck group than to the merit group, the salience effect predicts that the luck group place more emphasis on merit in the final round. This decreases their $m(j, \vec{s})$, and makes them implement less equality than the merit group. Since there is no efficiency cost involved, all numbers are equal for the two treatment groups in all eleven decisions. This allows for a clean test of the effect of manipulating the source of inequality in the first ten rounds.

1.4 Empirical strategy and analysis

1.4.1 Main analysis

This section outlines the experimental strategy and the regressions specified in the pre-analysis plan. I analyze the data from each experiment separately, using the same regression specifications. In both experiments, the two treatment groups face choices where equality, efficiency, and source of inequality are different. Thus, they make different decisions in the first ten rounds. This first stage difference is tested in Section 1.5. In the final round, spectators in the two treatments were given the exact same choice problem. The outcome variable of interest is an indicator for choosing the most equal distribution in the final decision round. Because treatment status is randomized between subjects, the two groups in each study differ only in the choices they faced and the decisions they made in the first ten rounds. The probability that spectator i chooses the most equal distribution in the final round is estimated by the following regression equation:

$$Y_i = \alpha + \beta T_i + \gamma \mathbf{X_i} + \epsilon_i \tag{1.6}$$

where the treatment indicator T_i is equal to one if spectator i is in the high-cost treatment in Study 1. In Study 2, T_i is equal to one if spectator i is in the merit group. $\mathbf{X_i}$ is a vector of control variables age (in years), gender, education, and political views. Education is measured with an indicator equal to one if the spectator has at least a two-year college degree, and zero otherwise. Political views are measured with an indicator variable for identifying as a political conservative. This is equal to one if the spectator answers four or five on a five-point scale from very liberal (1) to very conservative (5). α_i is a constant and ϵ_i is a mean-zero error term. Equation 1.6 is estimated with OLS, both with and without the background controls $\mathbf{X_i}$. The latter is considered to be the main test, as stated in the pre-analysis plan.

1.4.2 Heterogeneity analysis

I study heterogeneous treatment effects by interacting the treatment indicator with the indicator variables for gender, education, and political views, specified as follows:

$$Y_i = \alpha + \beta T_i + \gamma \mathbf{X_i} + \kappa(T_i \times group_i) + \epsilon_i, \quad (1.7)$$

where the regression is run separately for each of the three interacted *group* indicators female, college, and conservative. The heterogeneity analysis is included to explore the generalizability of the results, and whether a treatment effect is driven by particular subgroups of participants. As stated in the pre-analysis plan, the heterogeneity analysis is to be regarded as an exploratory analysis. The power calculations in the pre-analysis plan are for the main regression, and the experiment is, as most economic experiments, underpowered to study interaction effects.

1.5 Results

This section presents the main results. The section starts by presenting the results from Study 1. I begin by providing a descriptive analysis, then turn to the main regression results. I then go on to provide an exploratory heterogeneity analysis, before presenting the results from Study 2.

1.5.1 Study 1: Results

I find that, as expected, spectators are less willing to choose the more equal distribution when there is an efficiency cost involved in doing so. This is an important first stage of Study 1. Nevertheless, in line with previous literature on fairness preferences (see e.g., Engelmann and Strobel (2004) and Almås et al. (2020)), I find that willingness to equalize is relatively insensitive with respect to efficiency costs. In 61 percent of the choice settings with a high efficiency cost, the spectator is willing to take five tokens from one worker to give one token to the other. Figure 1.1 presents the distribution of the number of times the most equal distribution is chosen in the first ten rounds for each treatment group. In the low group, more than half of the spectators choose the most equal distribution in all the first ten rounds. In the high group, the same is true for about one third of the spectators. The average share that chooses the more equal option in the first ten rounds is 0.78 and 0.61 in the low treatment and high treatment, respectively. The difference between the groups is statistically significant ($p \leq 0.001$).

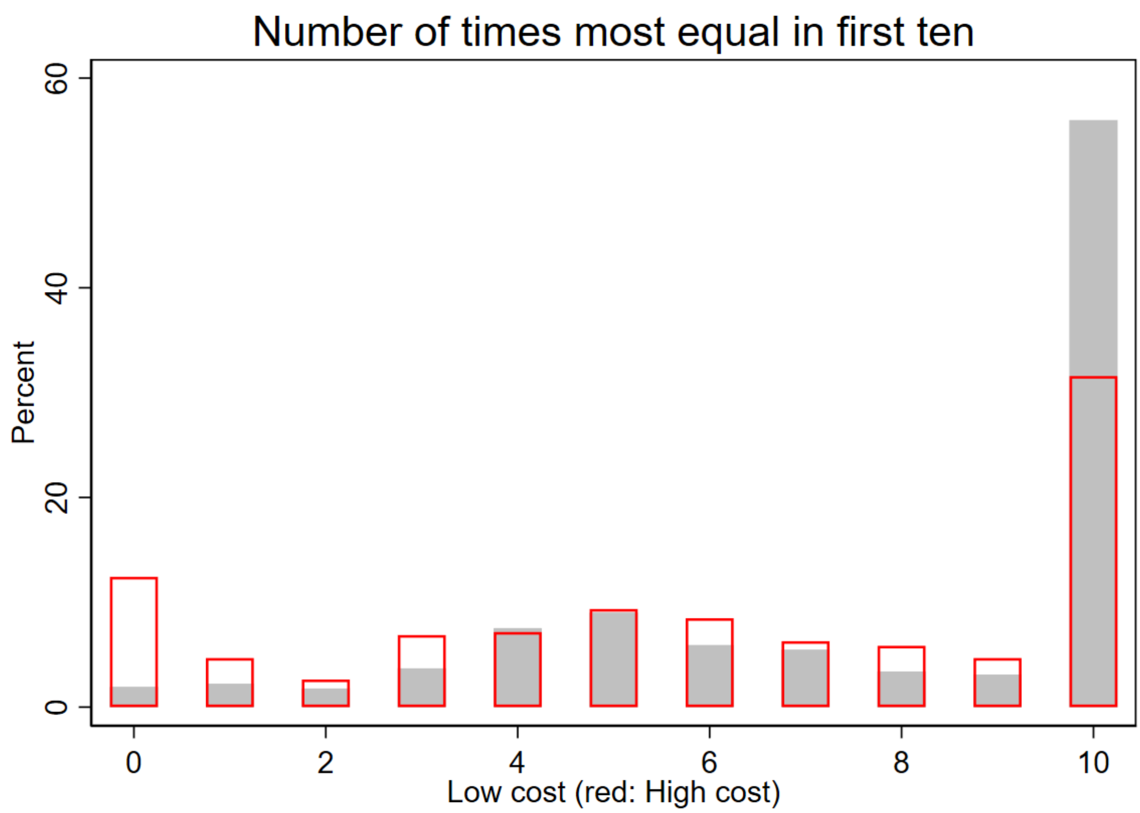


Figure 1.1: The histogram shows the distribution of the number of times the spectators choose the most equal distribution in the first ten rounds. The shaded bars indicate spectators in the low treatment. The bars with red frames indicate spectators in the high treatment.

In the eleventh decision, the share that chooses the most equal distribution is 0.74 in the low treatment and 0.63 in the high treatment, as shown in Figure 1.2. Thus, both groups seem to react to the change in efficiency cost, but not to the extent predicted by the path-independent model of Cappelen et al. (2013) and Almås et al. (2020), which predicts no difference in the final round, irrespective of what happens in the first ten rounds. This prediction is clearly rejected by the data from the Study 1 experiment ($p < 0.001$).

RESULT 1. *Willingness to pay an efficiency cost for equality depends on previously encountered efficiency costs of equality.*

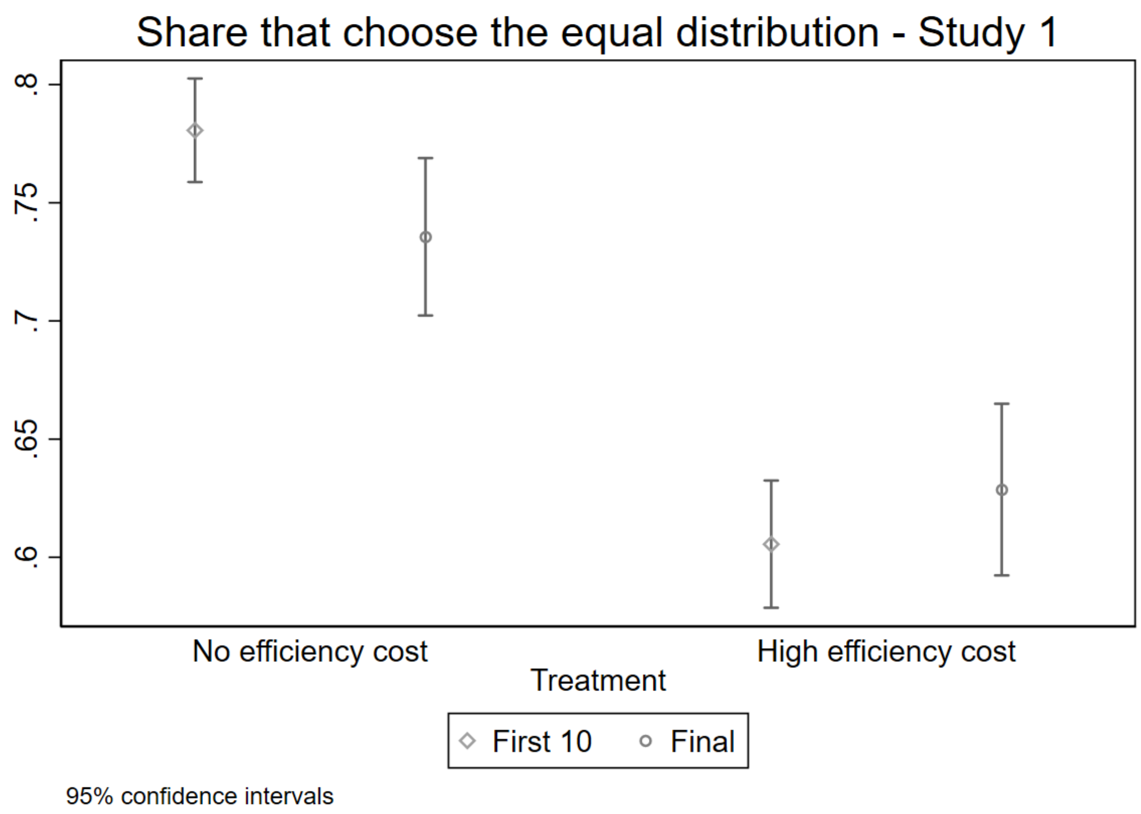


Figure 1.2: Share, by treatment group, that chooses the most equal distribution in the final round compared to the average of the first ten rounds. Path independence is rejected with $p < 0.001$.

Figure 1.2 clearly shows a partial convergence of group means in the final decision compared to the mean of the first ten decisions in each treatment group. This partial convergence is predicted by the habit effect, and is inconsistent with a strong salience effect. The latter predicts an overreaction, i.e., that the point estimates in Figure 1.2 switch place in the final decision relative to the mean of the first ten decisions. Figure 1.2 clearly shows that the habit effect dominates in the Study 1 data. I compare choices in the final round to choices in the first ten rounds for both treatment groups. Defining the difference between the average of the first ten rounds and the final round, I test the difference-in-difference between the two treatment groups by regressing the difference on treatment status. This test confirms that the decisions of the high group and the low group move in opposite directions when comparing the average of the first ten rounds to the final round. Choices in the final round are not significantly different from choices in the first ten rounds for either group.

Table 1.3 reports the corresponding regression analysis, as specified in Equation 1.6. Compared to the spectators in the low treatment who face no efficiency cost in the first ten rounds, spectators in the high treatment, who face a high efficiency cost of equality, redistribute 10.7 percentage points less in the final decision round. The estimated treatment coefficient is highly significant ($p < 0.001$), and robust to the inclusion of control variables, as seen from comparing the two columns in Table 1.3. The treatment effect is comparable in magnitude to the difference between males and females, and to the difference between people to the right of the political spectrum and those to the left and in the centre.

1.5.2 Study 1: Heterogeneity analysis

Here, I discuss the demographic explanatory variables presented in Table 1.3, and present a further analysis of treatment effects for different demographic groups.

As shown in Table 1.3, there is a positive and significant age coefficient in the data, meaning older spectators are more likely to choose the more equal outcome in the final

Table 1.3: Main regression — Study 1

	(1)	(2)
High cost	-0.107^{***}	-0.108^{***}
	(0.0251)	(0.0246)
Age		0.00279^{**}
		(0.00109)
Female		0.117^{***}
		(0.0251)
College		-0.0599^{**}
		(0.0298)
Conservative		-0.107^{***}
		(0.0255)
Constant	0.736^{***}	0.684^{***}
	(0.0170)	(0.0511)
Observations	1361	1361

Note: Outcome variable is an indicator for choosing the most equal distribution in the final decision. Columns 2 and 1 are with and without background characteristics. "Age" is in years, "Female" is an indicator for being female, and "College" is an indicator for having at least a two-year college degree. "Conservative" indicates answering 4 or 5 on a five-point scale to the question "Would you describe yourself as politically on the "left" (i.e. a liberal) or on the "right" (i.e. a conservative)?", where 1 is very liberal and 5 is very conservative. Robust standard errors in parentheses. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

decision. The coefficient for college is negative, suggesting that those with at least a two-year college degree are less likely to choose the more equal outcome. Both the age coefficient and the college coefficient are significant at the five percent level. Females are more likely to choose the more equal distribution than males. All control-variable coefficients are in the expected directions.

Table 1.4 presents results from the heterogeneity analysis specified in equation 1.7. This regression is run separately with each of the three group indicator variables interacted with the treatment indicator, including only one interaction term at a time. Each column shows the results from a regression of the indicator variable for dividing equally in the final round on treatment status, all background characteristics and the treatment variable interacted with the relevant group indicator, as specified in Equation

Table 1.4: Heterogeneity — Study 1

	Female	College	Conservative
High cost	-0.120^{***}	-0.294^{***}	-0.155^{***}
	(0.0323)	(0.0507)	(0.0312)
Group × High cost	0.0332	0.227^{***}	0.108^{*}
	(0.0493)	(0.0578)	(0.0504)
Group	0.101^{**}	-0.177^{***}	-0.162^{***}
	(0.0330)	(0.0338)	(0.0343)
Treatment effect on group	-0.0865^{**}	-0.0667^{**}	-0.0474
	(0.0372)	(0.0277)	(0.0393)
Observations	1361	1361	1361

Note: The table presents the treatment effects for different groups, estimated as in Equation 1.7. The regression is run separately with each of the column labels "Female", "College", and "Conservative" being entered as "Group" in Equation 1.7. The "High cost" coefficient is the estimated treatment effect from being in the high treatment group, i.e., facing a high efficiency cost of equality in the first ten rounds. "Group" is one if a participant is a part of the group, and zero otherwise. "Treatment effect on group" is the linear combination "High cost" + "Group" × "High cost".
Robust standard errors in parentheses. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

1.7. In the first column, the treatment variable is interacted with "Female", in the second column with "College", and in the third column with "Conservative". "Treatment effect on group" is the linear combination of the estimated treatment coefficient and the coefficient for the treatment interacted with the relevant group. The coefficients in the top and bottom rows indicate the treatment effects for the six (overlapping) groups defined by the control variables: males, females, people with a two-year college degree or more, those with less than a two-year college degree, political conservatives, and non-conservatives.

The treatment effect is negative for all groups, and with a $p < 0.05$ for all but political conservatives. The treatment effect is especially strong for those with less than a two-year college degree, with a coefficient of -29.4 percentage points $(p \leq 0.01)$. The coefficient for the interaction between college and the high treatment is positive and significant, and the total treatment effect on those with a college degree is only marginally negative. Political conservatives equalize less than liberals, but seem virtually unaffected by the treatment. The negative treatment effect in this data is driven disproportionately by political liberals and moderates, and by participants with less than two years of college.

RESULT 2. *The treatment effect on willingness to pay an efficiency cost for equality has the same direction for all groups, but with large heterogeneity in the magnitude.*

1.5.3 Study 2: Results

This section presents the results from the Study 2 experiment. The difference between the luck treatment and the merit treatment is the source of inequality between the pair of workers in the first ten rounds of redistribution. This experiment allows me to cleanly identify a path-dependency in the effect of the source of inequality on fairness perceptions, by having no efficiency cost of equality, thus ruling out β from equation 1.4.

Figure 1.3 shows the distribution of the number of times the equal distribution is chosen in the first ten rounds. I find that almost 80 percent of the spectators in the luck treatment choose the equal distribution in all of the first ten rounds, while only around ten percent of spectators in the merit group do the same. In the merit group, almost half of the spectators never choose the equal distribution. The difference between treatment groups in the average share that distributed equally in the first ten rounds is large, and statistically significant with $p \leq 0.001$. The same is true for the difference-in-difference between treatment groups for the first ten rounds versus the final round.

Figure 1.4 further shows that in the first ten rounds, the luck group is very reluctant to accept inequalities. This figure shows the average of the first ten rounds compared to the final round for both treatment groups. In the luck treatment, the average

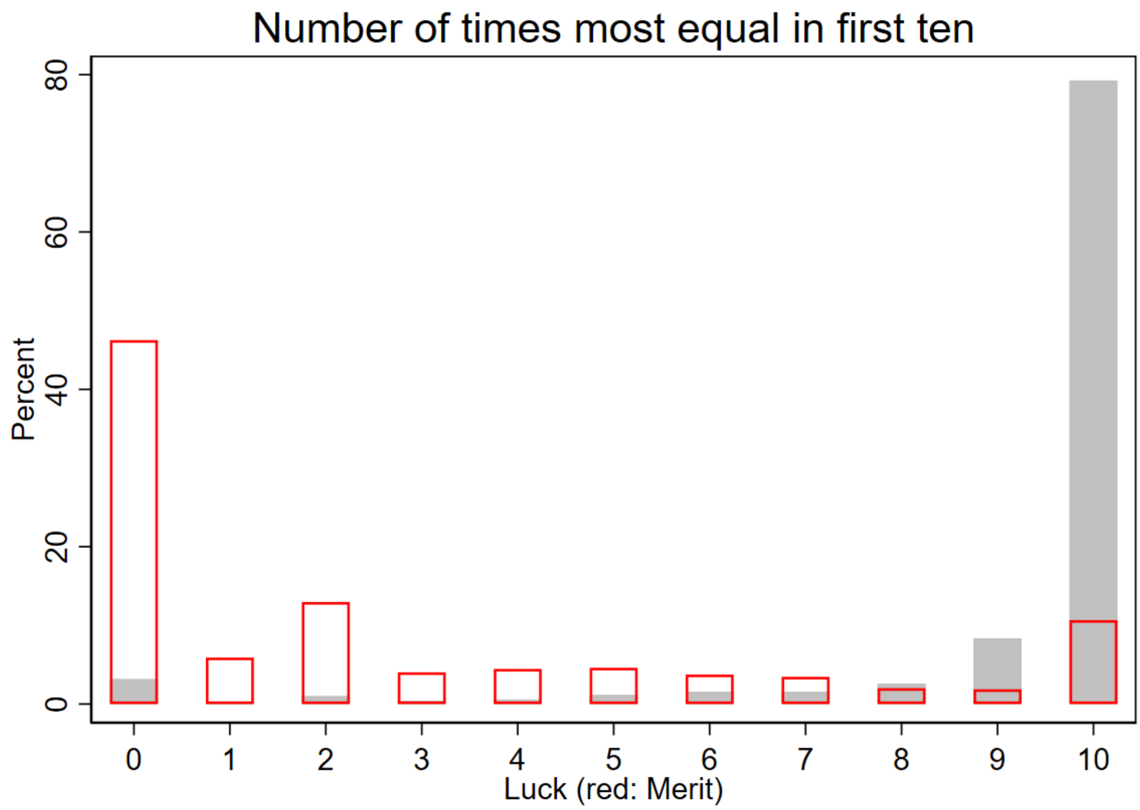


Figure 1.3: The histogram shows the distribution of the number of times spectators choose the most equal distribution in the first ten rounds. The shaded bars indicate spectators in the luck treatment. The bars with red frames indicate spectators in the merit treatment.

share that divide the payment equally in the first ten rounds is around 71 percent. However, when faced with the final decision, where the inequality is due to merit, inequality acceptance increases dramatically, and only 39 percent choose to equalize bonus payments between the pair of workers. The difference is statistically significant ($p \leq 0.001$). In the merit treatment, the share of spectators that divide the bonus equally is almost exactly the same in the final decision as the average of the first ten, at 32 and 31 percent, respectively. In the merit group, the inequality in initial allocations is due to merit in all eleven decision settings. Here, there is no discernible difference in willingness to redistribute in the final decision compared to the average of the first ten.

In the final decision, which is identical for both treatments, spectators in the merit treatment are less likely to equalize bonus earnings between workers, compared to

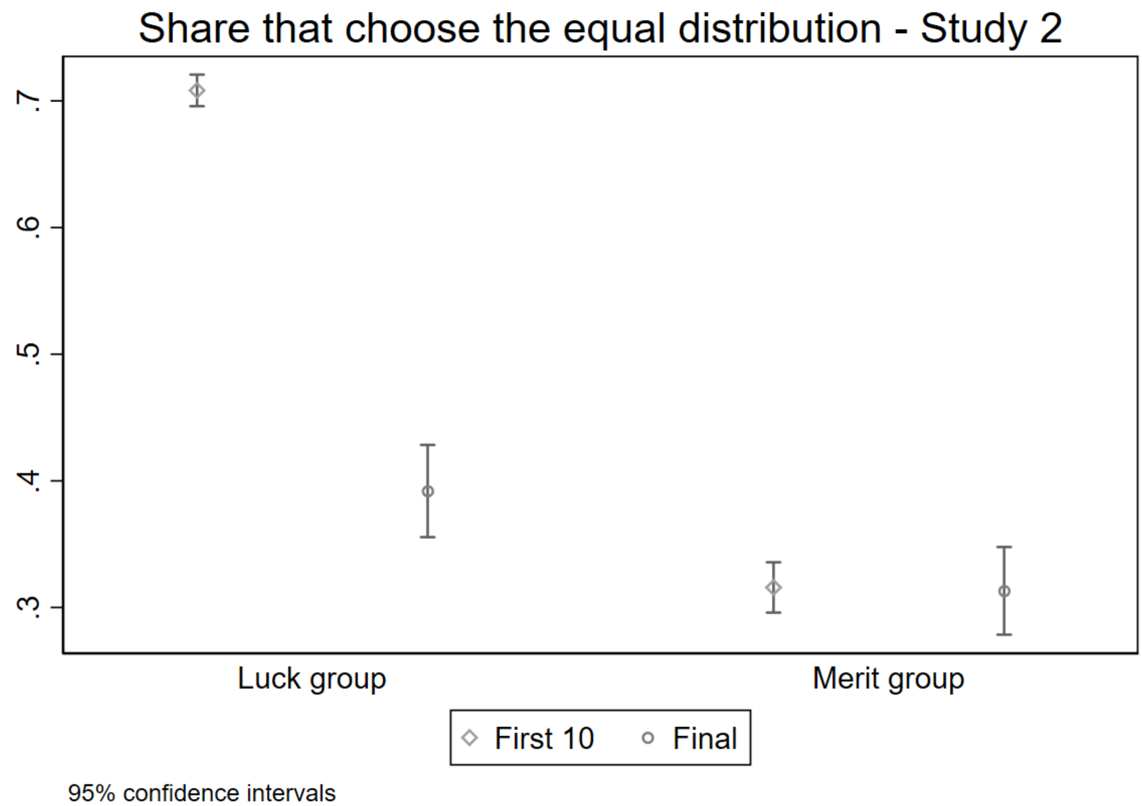


Figure 1.4: Share, by treatment group, that chooses the equal distribution in the final round compared to the average of the first ten rounds. Path independence is rejected with $p < 0.002$.

spectators in the luck treatment. This effect goes in the same direction as the effect observed in Study 1.

Table 1.5 provides the corresponding regression results for Study 2. The coefficient for "Merit group" is the estimated treatment effect of being in the merit treatment, relative to the luck treatment. The regressions show that those who see merit as the source of inequality in prior rounds implement 7.9 percentage points less equality in the final round than those who see luck as the source of inequality in prior rounds $(p = 0.002)$. The estimated treatment effect is comparable in magnitude with the coefficient for identifying as politically conservative. The treatment effect is robust to controlling for background characteristics, as seen from the comparison of the two columns in the table.

Table 1.5: Main regression — Study 2

	(1)	(2)
Merit group	-0.0788^{***}	-0.0789^{***}
	(0.0256)	(0.0256)
Age		-0.00192^{**}
		(0.000825)
Female		0.00853
		(0.0257)
College		-0.0444
		(0.0273)
Conservative		-0.0816^{***}
		(0.0292)
Constant	0.392^{***}	0.525^{***}
	(0.0185)	(0.0459)
Observations	1387	1377

Note: Outcome variable is an indicator for choosing the most equal distribution in the final decision. Columns 2 and 1 are with and without background characteristics. "Merit group" is an indicator for being in the merit treatment."Age" is in years, "Female" is an indicator for being female, and "College" is an indicator for having at least a two-year college degree. "Conservative" indicates answering 4 or 5 on a five-point scale to the question "Would you describe yourself as politically on the "left" (i.e. a liberal) or on the "right" (i.e. a conservative)?", where 1 is very liberal and 5 is very conservative.
Robust standard errors in parentheses. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

RESULT 3. *Willingness to accept inequality due to merit is affected by the source of inequality faced in previous distributive decisions.*

1.5.4 Study 2 heterogeneity analysis

Here, I discuss the demographic explanatory variables presented in Table 1.5, and present a further analysis of treatment effects for different demographic groups.

Column 2 of Table 1.5 shows a negative age coefficient. The coefficients for "Female" and "College" are not significantly different from zero. There is a negative coefficient for the variable "Conservative". Political conservatives are 8.2 percentage points less likely to implement equality in the final round, compared to liberals and moderates

$(p \leq 0.01)$. A similar effect is shown in Table 1.3 for Study 1. The age coefficient goes in opposite directions in the two studies. Willingness to pay an efficiency cost to equalize earnings increases with age (Study 1), as does willingness to accept inequalities that are due to merit. Table 1.3 shows a strong gender effect on the estimated probability of choosing the more equal distribution. The same is not found in Table 1.5 for Study 2. Here, I find no gender differences in willingness to accept inequalities that are due to luck. The estimated effect of having a college degree is similar in the two studies, but not statistically significant in Study 2.

Table 1.6 shows a heterogeneity analysis for Study 2. Each column shows the results from a regression of the probability of dividing the payment equally in the final round on treatment status, all background characteristics, and the treatment variable interacted with the relevant group indicator, as specified in Equation 1.7. Table 1.6 follows the same pattern as Table 1.4. The direction of the treatment effect is the same (negative) for all spectator groups. The direction is also the same as in Study 1 for all subgroups. This could give some indication that the two effects are driven by the same underlying mechanism(s). However, in Study 2, the treatment effect is only statistically significant for females, people with at least a two-year college degree, and political conservatives. Thus, the treatment effect in Study 2 is somewhat less robust across subgroups than the effect observed in Study 1. In Study 2, the treatment coefficient is virtually the same for the groups represented by the bottom row in Table 1.6. None of the interaction effects are statistically significant.

RESULT 4. *The treatment effect on willingness to accept inequality due to merit has the same direction for all subgroups, but is not always statistically significant.*

Table 1.6: Heterogeneity — Study 2

	Female	College	Conservative
Merit	-0.0388	-0.0200	-0.0680^{*}
	(0.0367)	(0.0442)	(0.0299)
Group × Merit	-0.0783	-0.0910	-0.0449
	(0.0512)	(0.0541)	(0.0572)
Group	0.0476	0.000563	-0.0595
	(0.0372)	(0.0387)	(0.0427)
Treatment effect on group	$-.117^{***}$	$-.111^{***}$	$-.113^{**}$
	(0.0357)	(0.0313)	(0.0489)
Observations	1377	1377	1377

Note: The table presents the treatment effects for different groups, estimated as in Equation 1.7, as well as interactions. The regression is run separately with each of the column labels "Female", "College", and "Conservative" being entered as "Group" in Equation 1.7. The "Merit" coefficient is the estimated treatment effect from being in the merit group in the first ten rounds. "Group" is one if a participant is part of the group, and zero otherwise. "Treatment effect on group" is the linear combination "Merit" + "Group" × "Merit".
Robust standard errors in parentheses. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

1.6 Discussion and conclusion

This paper presents the results from two large-scale, incentivized spectator experiments. Both experiments show that decision-makers acting as third-party spectators exhibit path-dependent fairness preferences. Study 1 shows that spectators who have faced a high efficiency cost of equality in prior decisions are less likely to accept a moderate efficiency cost in a subsequent decision, compared to spectators who have faced no efficiency cost of equality in prior decisions. Study 2 manipulates the source of inequality in prior decisions and identifies the path dependence of spectators' fairness ideals. I show that spectators in identical distributive contexts make different decisions, depending on previously encountered sources of inequality. In the two studies, spectators react to changes in efficiency costs and changes in the source of inequality, respectively. However, both of the experiments reported here show that the spectators adapt their decisions less than that predicted by path-independent models. The main results from both experiments are consistent with habit-dependent preferences for fairness. Both experiments also indicate that the habit effect dominates the salience effect in distributive decisions.

To my knowledge, this paper is the first to study the effect of choice-set history on distributive choices made by impartial spectators. The experiments presented show that distributive choices depend on the available alternatives in previous decisions, and that decision-makers develop habits that carry over to subsequent decisions. This paper has documented that such habits can have a strong impact on distributive decisions, even when other salient features of the choice context are rigged to counteract the effect of habits. The magnitude of the effects reported here are comparable with differences in distributive behavior across the political spectrum. Path-dependence in fairness preferences has implications for how we study behavior in repeated distributive decisions, and particularly for how we interpret results from repeated economic games.

Understanding how habits for fairness are formed, and how they develop over time, appears to be a promising avenue for future research. In order to fully understand how habits affect distributive behavior, we must first understand how habits are formed and how broadly they apply. The habit effects documented in the present paper arise in a series of decisions that look very similar, and where most aspects of the distribution are kept constant. They are also made over a brief span of time and in a confined choice context. Studying to what extent habits carry over to similar or dissimilar decisions, and how quickly the habit effects decay over time, would help to improve our understanding of path dependence in decision-making more generally. If habits for equality are shown to have a long-term impact on distributive behavior, this could even help us to understand why some countries have more redistributive institutions than others, and why family background matters for moral preferences.

1.A Spectator instructions — Study 1

Welcome!

Please note that your participation will be registered on the following Amazon Mechanical Turk worker ID:

[worker ID]

The worker ID was retrieved automatically when you clicked on the link that brought you here. This step is necessary for assigning payments to the right account and to ensure that you only participate in this study once.

Introduction

Welcome to this research project! We very much appreciate your participation.

This is a study about the economics of decision-making. Several research institutions have provided funds for this research.

Payment

Your payment for taking the HIT will be sent to you shortly after the completion of this HIT.

Procedures

The study consists of two parts and you will be given instructions on your screen before every single part of the survey. Please always make sure to read the instructions carefully before you continue.

Participation

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate.

Confidentiality

Your worker ID will only be used for payment purposes, and will not be stored with the data collected. Anonymized data from the HIT might be shared in open science repositories.

Verification

At the end of this survey, you will be given a completion code. You will need to copy this code to the survey code field on the AMT web page that directed you here at the beginning.

Questions about the Research

If you have questions regarding this study, you may contact

thechoicelab@nhh.no

[*New page*]

I have read and understood the above consent form and desire to participate in this study.

Yes/No

[*New page*]

You will now be asked to choose how to distribute tokens, each equating 10 cents in final payment, between several pairs of anonymous workers, who have completed the same set of tasks equally well. These are real people. With a certain probability one of your decisions in this survey will be implemented to decide the payout for a real pair of workers. Each decision is equally likely to be chosen, using a random number generator. Your decisions affect real people, so it is important that you take this seriously.

In addition to a participation fee, workers are told that they may get a bonus. They know that this bonus is decided by someone else, but they have no further information.

For simplicity, we call the workers worker A and worker B. Each decision involves a new pair of real workers, so A and B are not necessarily the same people throughout.

[*New page*]

Before proceeding to the main questions, you must answer a simple question about cities. Here, you must simply select 'London'. This is an attention check.

Based on the text above, which city have you been asked to enter?

London

Paris

Madrid

Rome

Frankfurt

The group with a high efficiency-cost of equality answer ten questions similar to the one below:

Please select your preferred distributions of money to worker A and worker B

- $20 to A and $80 to B
- $30 to A and $30 to B

The group with no efficiency-cost of equality answer ten questions similar to the one below:

Please select your preferred distributions of money to worker A and worker B

- $20 to A and $80 to B
- $50 to A and $50 to B

Then, all participants answer the same distributive question:

Please select your preferred distributions of money to worker A and worker B

- $15 to A and $90 to B
- $30 to A and $45 to B

Some background questions at the end:

You have completed the first part of the survey. We would now like to ask you five more questions before we conclude this survey.

What is your gender?

- Male
- Female
- Other/do not wish to reply

How old are you?

[Open textbox]

What is the highest level of education you have completed?

- Less than High School
- High School / GED
- Some College
- 2-year College Degree

- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

Would you describe yourself as politically on the "left" (i.e. a liberal) or on the "right" (i.e. a conservative)?

1 - Very liberal	2	Neutral	4	5 - Very conservative

Finally, if you have any comments or suggestions related to this study please write them down in the field below. Your feedback is very important to improve our research.

[Open textbox]

1.B Worker instructions — Study 1

Welcome!

Please note that your participation will be registered on the following Amazon Mechanical Turk worker ID:

[worker ID]

The worker ID was retrieved automatically when you clicked on the link that brought you here. This step is necessary for assigning payments to the right account and to ensure that you only participate in this study once.

[*New page*]

Introduction

Welcome to this research project! We very much appreciate your participation.

This is a study about the economics of decision-making. Several research institutions have provided funds for this research.

Payment

Your payment for taking the HIT will be sent to you shortly after the completion of this HIT.

Procedures

The study consists of two parts and you will be given instructions on your screen before every single part of the survey. Please always make sure to read the instructions carefully before you continue.

Participation

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate entirely without jeopardy to future participation in other studies conducted by us.

Confidentiality

Your worker ID will only be used for payment purposes, and will not be stored with the data collected. Anonymized data from the HIT might be shared in open science repositories.

Verification

At the end of this survey, you will be given a completion code. You will need to copy this code to the survey code field on the AMT web page that directed you here at the beginning.

Questions about the Research

If you have questions regarding this study, you may contact

thechoicelab@nhh.no

[*New page*]

I have read and understood the above consent form and desire to participate in this study.

Yes/No

[*New page*]

You will now be asked to make some hypothetical decisions about how to distribute tokens between several pairs of anonymous workers. Imagine that they have completed the same set of tasks and performed equally well. Imagine they get a participation fee, and that they are also told that they may get a bonus. They know that this bonus is decided by someone else, but they have no further information.

For simplicity, we call these workers worker A and worker B. Each decision involves a new pair of workers, so A and B should not be thought of as the same people throughout. Although these decisions do not affect real people, it is vital to our research that you try to answer each question as if it were a real situation.

Thank you for taking this seriously.

You may yourself earn a bonus from this HIT. This is decided by chance and somebody else not in this study. They will not know what you answer, and your answers in this study will not in any way affect your potential bonus.

[*New page*]

The group with a high efficiency-cost of equality answer ten questions similar to the one below:

Please select your preferred distributions of money to worker A and worker B

- 20 to A and 80 to B

- 30 to A and 30 to B

The group with no efficiency-cost of equality answer ten questions similar to the one below:

Please select your preferred distributions of money to worker A and worker B

- 20 to A and 80 to B
- 50 to A and 50 to B

Then, all participants answer the same distributive question:

Please select your preferred distributions of money to worker A and worker B

- 15 to A and 90 to B
- 30 to A and 45 to B

Some background questions at the end:

You have completed the first part of the survey. We would now like to ask you five more questions before we conclude this survey.

What is your gender?

- Male
- Female

- Other/do not wish to reply

How old are you?

[Open textbox]

What is the highest level of education you have completed?

- Less than High School
- High School / GED
- Some College
- 2-year College Degree
- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

Would you describe yourself as politically on the "left" (i.e. a liberal) or on the "right" (i.e. a conservative)?

1 - Very liberal	2	Neutral	4	5 - Very conservative

Finally, if you have any comments or suggestions related to this study please write them down in the field below. Your feedback is very important to improve our research.

[Open textbox]

1.C Spectator instructions — Study 2

The introduction and the demographics questions are identical to those from study 1, see appendix 1.A. The only differences are in the instructions and the decision part, as shown below.

A few days ago, a group of workers performed a set of tasks for us. In addition to a participation fee, workers were told that they might get a bonus. They were told that this bonus is decided by performance, luck and by someone else, but they got no further information. Workers were paired together. In some pairs, each worker in any given pair performed equally well. In these pairs, one worker gets a higher bonus than the other one by chance of luck. In other pairs, one worker performed better than the other. In these cases, the best-performing worker is allocated a higher bonus than the other worker.

Bear in mind that these are real people. All amounts are given in experimental tokens, each worth 10 cents in final payment. For simplicity, we call the workers worker A and worker B. Each decision involves a new pair of real workers, so A and B are not necessarily the same people throughout.

You will now be shown allocations of bonus for some paired of workers, and you may choose to redistribute the bonus between the two workers as you see fit.

With a certain probability, one of your decisions in this survey will be implemented to decide the payout for a real pair of workers. Each decision is equally likely to be chosen, using a random number generator. Your decisions affect real people, so it is important that you take this seriously and read the instructions carefully.

[*New page*]

Spectators are randomized into either a Merit treatment or a Luck treatment, each with probability one-half. In the Merit treatment, spectators make ten decisions where the information is the same, but the numbers and the order of worker A and worker B varies:

In this pair, worker B performed better than worker A, and therefore is allocated a higher bonus 60, while A gets 20.

Please select your preferred distribution of tokens to worker A and worker B

- 20 to A and 60 to B
- 40 to A and 40 to B

[*New page*]

In the Luck treatment, spectators make ten decisions where the information is the same, but the numbers and the order of worker A and worker B varies:

In this pair, both workers performed equally well. By luck, worker B was allocated a high bonus of 60, while B gets 20.

Please select your preferred distribution of tokens to worker A and worker B

- 20 to A and 60 to B
- 40 to A and 40 to B

The numbers in the first ten decisions are the same for both groups.

[*New page*]

The final decision is the same for both group:

In this pair, worker B performed better than worker A, and is allocated a higher bonus.

Please select your preferred distribution of tokens to worker A and worker B.

- 10 to A and 70 to B
- 40 to A and 40 to B

1.D Worker instructions — Study 2

Welcome!

Please note that your participation will be registered on the following Amazon Mechanical Turk worker ID:

The worker ID was retrieved automatically when you clicked on the link that brought you here. This step is necessary for assigning payments to the right account and to ensure that you only participate in this study once.

New page

Introduction

Welcome to this research project! We very much appreciate your participation.

This is a study about the economics of decision-making. Several research institutions have provided funds for this research.

Payment

Your payment for taking the HIT will be sent to you shortly after the completion of this HIT.

Procedures

You will be given instructions along the way. Please always make sure to read the instructions carefully before you continue.

Participation

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate.

Confidentiality

Your worker ID will only be used for payment purposes, and will not be stored with the data collected. Anonymized data from the HIT might be shared in open science repositories.

Verification

At the end of this survey, you will be given a completion code. You will need to copy this code to the survey code field on the AMT web page that directed you here at the beginning.

Questions about the Research

If you have questions regarding this study, you may contact

thechoicelab@nhh.no

Write ACCEPT in the form below if you have read and understood the above consent form and desire to participate in this study.

[Open textbox]

New page

Worker are randomized into either a merit group or a luck group

Merit group:

In this assignment, you have the opportunity to earn a bonus.

Your aim in this assignment is to correctly solve a set of slider tasks as quickly as possible. After completion, you will be randomly paired with one other worker, and the one with the quickest time is allocated a winner bonus. The other one is allocated a lower bonus.

The exact allocation is variable, and depends on chance, as well as the choices of a third party outside of this study.

Your finally decided bonus will be paid within two weeks once the third party has made their decision.

New page

Task: Set the sliders to their prescribed values as quickly as possible.

[20 sliders to be set to the correct value. Visible timer at the bottom of the page.]

New page

You have completed all the tasks. Your final payment will be decided by your relative performance and by a third party outside of this study, and paid to you within two weeks. Thank you.

Luck group:

In this assignment, you have the opportunity to earn a bonus.

Your aim in this assignment is to correctly solve a set of slider tasks. After completion, you will be randomly paired with one other worker, and, by luck, one of you will be allocated a high bonus and the other one is allocated a lower bonus.

The exact allocation is variable, and depends on chance, as well as the choices of a third party outside of this study.

Your finally decided bonus will be paid within two weeks once the third party has made their decision.

New page

[20 sliders to be set to the correct value. No visible timer at the bottom of the page.]

New page

You have completed all the tasks. Your final payment will be decided by your relative performance and by a third party outside of this study, and paid to you within two weeks. Thank you.

Chapter 2

Fairness of the crowd: An experimental study of social spillovers in fairness decisions

Kjetil Røiseland Madland and Eirik André Strømland *

Abstract

This paper reports from a large-scale experiment conducted to study the effects of social norms on distributive decision-making. In an incentivized spectator experiment, subjects chose how to divide bonus earnings between a pair of stakeholders. Before choosing a distribution, our spectators stated their beliefs about, and received a signal about, the share of payoff-equalizing spectators in a reference group, randomly drawn from a previous study with the same distributive setting (Almås et al., 2020). This draw gives random variation in the intensity of the signal about the norm that applies in the

*Madland: FAIR, Department of Economics, NHH Norwegian School of Economics; Strømland: HVL Western Norway University of Applied Sciences.
Both authors would like to thank Mathias Ekström, Samuel Hirshman, Laura Khoury, Fanny Landaud, Akshay Moorthy, Osama Moeed Nawab, Hallgeir Sjåstad, Oda Sund, Erik Ø. Sørensen, Bertil Tungodden, Heidi Christina Thysen, Alexander Willén, Mads Fjeld Wold, and Nina Serdarevic for helpful comments and suggestions. We also owe huge thanks to Sebastian Fest for invaluable help with the data collection. We also want to thank FAIR, the Department of Economics at NHH, the Department of Economics at HVL, and the Research Council of Norway for the funding that has enabled this research. Our experiment was approved by the Institutional Review Board at NHH Norwegian School of Economics, and pre-registered at AEAs RCT Registry with RCT ID AEARCTR-0010023.

current setting. We find a statistically significant but small effect of the number of payoff-equalizing spectators in the reference group on the probability of equalizing payoffs in the current setting. The redistribution choice is strongly correlated with spectators' initial beliefs about the reference group. The effect of the signal about redistribution is primarily driven by the subgroup of participants who receive a large shock to their initial beliefs.

2.1 Introduction

Have you ever agreed with something just because you thought it was what others wanted to hear? People often wish to conform to the social groups that they are a part of, and find it hard to behave in ways that stand out from the majority. Even in situations without a fixed group, or with weak social ties, simply doing what everybody else does may seem like the safe choice. The desire to conform to the views and values of our peers shapes our decisions in various domains. Since social norms vary across societies, identifying situations where social norms are important may help us understand the sources of variation in behavior and institutions.

Societies across the world differ vastly in their levels of economic inequality (Piketty et al., 2014). A recent study by Almås et al. (2020) (hereinafter ACT) found that the difference in inequality levels between the US and Norway fits quite well with the differences in preferences for inequality between the two countries – the US has more libertarians while Norway has more egalitarians. While variation in preferences probably plays an important part in explaining cross-country variation in inequality, it is also possible that fairness views are shaped by social norms that regulate behavior. If social norms depend on how many people are observed to act in a certain way, the prevalence of libertarians in the US and the prevalence of egalitarians in Norway potentially suggest a different social norm for redistribution in the two countries. This fairness norm could be strongly shaped by the majority choice in the respective societies. Thus, once initially established, a social norm may be self-reinforcing through the mechanism whereby people tend to prefer to conform with the prevailing social norm (Alesina and Angeletos, 2005; Gunnthorsdottir et al., 2007).

In this paper, we conduct a randomized, incentivized online spectator experiment to study whether a social norm of equalizing income can be self-reinforcing in a setting where people obtain information about the norm by observing previous redistribution choices made by their peers. In our experiment, impartial spectators choose how to distribute a bonus between two unknown stakeholders. Before making their distributive choice, the spectators were informed about the choices made by a reference group of spectators who have made the same choice prior to this study. Each spectator views the choices made by a random group of three spectators drawn from the data of ACT. Specifically, spectators are informed about how many of the spectators in the reference group from ACT chose to equalize income between stakeholders. We hypothesize that the information about the reference group of spectators serves as a signal about which social norm applies in the current situation, thereby tilting the spectator in the direction of the perceived norm. Before giving information about the reference group, we elicited spectators' beliefs about the choices made by the reference group. In order to be able to study whether a norm of equal redistribution is self-reinforcing, we use an empirical setting where we know from research that there is a relatively balanced distribution of egalitarians and libertarians, i.e., a US sample (Almås et al., 2020). A high prevalence of libertarians creates a high potential for participants to receive a strong signal. On the other hand, since we draw our signals from a distribution of choices made in a previous study, we need a high prevalence of egalitarians so that we can randomly vary the information provided to participants and avoid deception.

Our study allows for clean identification of the "descriptive-norm" effect (Cialdini et al., 1990) of fairness preferences in a spectator environment. The descriptive norm effect refers to a phenomenon whereby people tend to prefer to act in ways that are consistent with the behavior of others. By drawing randomly from the data in ACT, we obtain random variation in the signal provided to each spectator, without deceiving our participants. The signal contains information about the distributive actions of a reference group of spectators from the same US population. By removing all interaction, and providing minimal information about the reference group, our study contributes to the part of the social-norm literature that shows that people often tend to simply do "what everybody else does". The signal provided is the only piece of information that can inform participants, in addition to their prior beliefs, about the prevailing descriptive norm in the reference population - presumably the general US population. Moreover, if there is a descriptive norm effect, participants' prior beliefs should also be correlated with the probability of equalizing earnings.

Sociologists have long held that our choices are shaped by social norms shared by those around us. In recent decades, the role of norms has gained increasing attention in economics since social norms are closely linked to economic institutions and societal characteristics.[2] Lindbeck et al. (1999) model labour participation in a framework where living on benefits induces a utility cost by deviating from a social norm. Benabou and Tirole (2011) model the interplay between social norms and material incentives, and how laws affect not only material incentives, but also individuals' perceptions about society's moral values. In these models, law and order are not only upheld by hard sanctions, but also by the social cost associated with norm-breaking. Keizer et al. (2008) find that observing others violating social norms and legitimate rules makes people more likely to violate other norms and rules themselves, causing disorder to spread.

In the models presented by Lindbeck et al. and Benabou and Tirole, norms are understood to work through non-pecuniary incentives, such as feelings of guilt, pride, shame, and other social repercussions, rather than through preferences. Understood

[2]See Elster (1989) for a discussion of social norms in sociology and economics

in this context, using norms to encourage socially desirable behavior and discourage socially undesirable behavior can be costly in terms of welfare. This is shown in empirical work by DellaVigna et al. (2012), who find that door-to-door fund-raising lowers the utility of potential donors because of social pressure. Their findings are in contrast to the concept of *homo economicus*, for whom the possibility of donating money to fund-raisers is seen as a real option with a non-negative value. Funk (2010) find suggestive evidence that reducing the cost of voting has a bigger impact in smaller communities, where the chance of being observed is larger. Gerber et al. (2008) find that increasing people's beliefs that others will know whether they have voted or not substantially increases voting. DellaVigna et al. (2016) also show that being able to tell others, or not having to lie about having voted, is an important motivation for voting.

A large body of empirical literature shows that people tend to adapt their actions to the social norms of others, both injunctive norms, i.e., what others approve of, (see, e.g., Bursztyn et al. (2020a,b); Cialdini et al. (1990); Pryor et al. (2019)) and empirical norms, what most others actually do (see, e.g., Bicchieri and Xiao (2009); Bott et al. (2020); Del Carpio (2013); Fellner et al. (2013); Krupka and Weber (2009); Pryor et al. (2019); Schultz et al. (2007); Wenzel (2005)). Bicchieri and Xiao (2009) compare the effects of injunctive norms and descriptive norms, and find that norm-compliance is mainly driven by the latter, i.e., that actions speak louder than words. Krupka and Weber (2009) suggest that norm compliance is driven by two primary mechanisms, focusing and informational influence. Focusing implies that directing agents' focus toward a social norm increases norm compliance independently of whether or not other people adhere to that norm. Informational influence means that adherence to a norm increases when a lot of other people are also observed following the norm. Our paper contributes to this strand of the social-norm literature by showing that some spectators are swayed by the majority choice of a group of people with whom they only share nationality, and even if their actions are not observed by others.

A vast literature has studied the effects of social identity, self-categorization and group-belonging on the social norms that guide behavior (Akerlof and Kranton, 2000, 2005; Austen-Smith and Fryer Jr, 2005; Ben-Ner et al., 2009; Bénabou and Tirole, 2006; Benjamin et al., 2010; Bicchieri et al., 2021; Charness and Chen, 2020; Chen and Li, 2009; Fang and Loury, 2005; Klor and Shayo, 2010; Koszegi et al., 2019; Kranton, 2016; Peysakhovich and Rand, 2017; Shayo and Zussman, 2011; Tajfel et al., 1979, 1971; Wenzel, 2004). Self-categorization theory (Tajfel et al., 1979) predicts that actions will be positively affected by people in an agent's in-group, and negatively affected by people in the agent's out-group. The present paper contributes to this literature by showing that people are affected by the norms of others, even when social ties are weak and non-salient. Thus, our study relates most closely to Peysakhovich and Rand (2017), who show that in-group favoritism can arise even under minimal conditions, such as team color in the game Pokémon Go. Our study also relates to Pryor et al. (2019), who find that "a general desire to conform with others may out-power the common in-group vs out-group mentality".

The rest of the paper is organized as follows: Section 2.2 provides a thorough description of the experiment and the participants included in our sample. Section 2.3 describes the empirical analyses we pre-registered, as well as exploratory results. Section 2.4 presents regression results, while section 2.5 concludes.

2.2 Experimental design and participants

This paper reports on an incentivized online experiment, where one group of participants act as *stakeholders/workers*, who receive a bonus payment for work performed, and one group act as impartial *spectators* who decide how the bonus should be distributed between the stakeholders. The experimental design largely follows that of ACT.

The role of the workers in our experiment is to provide real incentives for the spectators, whose decisions are of first-order interest. The workers work on sentence unscrambling tasks for about five minutes. There are no right or wrong answers, so performance on the tasks is not measured. After completing the tasks, each worker is paired with another worker, and one of them is allocated a bonus of $6, while the other one gets nothing. The bonus allocation is decided by luck. The spectators are informed about the allocation of bonus payments to the workers. They are also informed that both workers performed equally well, and that the distribution was decided by luck. The spectators may choose to keep the allocation as it is, or choose a different allocation.

We recruited 983 spectators from the online labour platform Prolific and 200 respondents as workers through MTurk. The spectator sample was stratified to be representative of the US population in terms of age, sex, and ethnicity.[3] Both the worker part and the spectator part of the experiment were run as online experiments using Qualtrics. The experiment was conducted in September 2022.

Table 2.1 presents sample characteristics for our spectator sample. The sample is representative of the grown-up US population in terms of age, sex, and ethnicity, but our sample is more highly educated and less politically conservative.

In order to study whether a social norm of equal distribution of income can be self-reinforcing, we constructed an experimental setting in which all spectators received information about what a group of spectators (whom we refer to as a "reference group") did in a previous experiment by ACT. Each of these reference groups consisted of three spectators randomly drawn from the dataset collected by ACT. All spectators in our study were given information about the number of people in the reference group who chose to distribute the bonus equally between workers. By drawing reference groups

[3]In the pre-analysis plan, we originally aimed for 1000 spectators and 200 workers. The data collection was stopped at 983 because it proved hard to recruit more participants using Prolific's representative sample.

Table 2.1: Sample characteristics

	Sample	US population
Age median	45	38.3
Female share	0.51	0.51
College share	0.68	0.43
Conservative share	0.23	0.36
Observations	983	

Note: "Age median" is the median age in years. "Female share" is the proportion of females. "College share" is the share with at least a two-year college degree. Age, sex, and education data for the US population from the US Census Bureau (2021). "Conservative share" in the Study 1 and Study 2 samples indicates 4 or 5 on a five-point scale for the question "Would you describe yourself as politically on the 'left' (e.g., a liberal) or on the 'right' (e.g., a conservative)?", where 1 is very liberal and 5 is very conservative. "Conservative share" of the US population is taken from Gallup, asking "How would you describe your political views – [very conservative, conservative, moderate, liberal or very liberal]?", including the answers "conservative" and "very conservative" (Gallup, 2022).

randomly, we achieved exogenous variation in the number of subjects in the reference group who chose to equalize earnings between workers. This allows us to causally examine how a signal about a social norm of equal distribution of income influences redistributive choices.

2.2.1 Spectator decisions and treatment

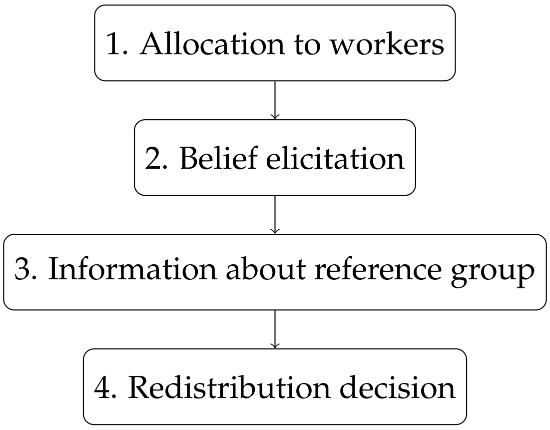
Our setup draws on the *luck* treatment in ACT. In this treatment, spectators are told that the initial inequality between workers is due to luck, and they are given the chance to redistribute earnings. In our experiment, all spectators are given information about the decisions of a randomly generated group of $N = 3$ from the US sample in the luck treatment in the ACT data. Specifically, our spectators are informed about the share of egalitarians in the reference group, i.e., how many in the reference group that chose to equalize bonus earnings between the two workers. Randomly drawing three spectators from the luck sample in ACT gives us the following probabilities of having

x members of the reference group divide the money equally: $Pr(x = 0) = 0.10, Pr(x = 1) = 0.35, Pr(x = 2) = 0.40$ and $Pr(x = 3) = 0.15$. These numbers are rounded with two decimals precision to the closest exact percentage to the original distribution. By randomizing between four different information screens with these probabilities, we achieve a distribution which gives an accurate representation of the data in ACT.

In our study, the spectators were first informed about the worker part and were told that workers are paired together such that each worker in a pair has performed equally well. The workers have performed the same set of tasks and, as far as the spectators are concerned, only differ in the bonus earnings they are allocated. The initial allocation of bonus payments to the pair of workers is ($6, 0), i.e., one worker in the pair gets $6 and the other gets nothing. The spectators are informed that workers know that the bonus is decided by luck and potentially also by someone else. As in ACT, the spectators in our study choose between the following distributions: (6, 0), (5, 1), (4, 2), (3, 3), (2, 4), (1, 5), and (0, 6). Both in ACT and in our experiment, the first alternative was framed as choosing not to redistribute earnings, while the others are framed as redistributions.

Before receiving information about the reference group from ACT, our spectators are told that "a large sample of Americans" have previously faced the same distributive decision, and that they themselves will be presented with information about the choices of a random sample of these Americans. Our spectators are asked to make an incentivized guess about how many in their reference group chose to distribute the bonus equally between a pair of workers. We define this as the spectators' prior beliefs. Those who state correct beliefs about the number of spectators in their reference group who divide the bonus equally earn an extra 10 cents. After stating their beliefs, our spectators were immediately told how many in their reference group chose to equalize bonus earnings. Finally, our spectators were matched with a real pair of workers, and asked to make their own distribution decision. The matching was done in a probabilistic way, such that ten percent of our spectators had their decisions implemented for a real pair of workers. The exact instructions for the workers are provided in appendix 2.B, and the instructions for the spectators are provided in appendix 2.C. Figure 2.1 presents the complete chain of events for our experiment.

Figure 2.1: Experiment chain of events



Our experimental design allows us to answer the following research questions:

- How does information about what others have done influence fairness views in a spectator setting?
- How does the impact of information about choices made by a reference group depend on our spectator's initial beliefs about redistribution in the reference group?

Our main hypothesis is that the decisions made by people in the reference group serve as a signal about the social norm, and that there is a cost associated with deviating from the social norm. Thus, we hypothesize a positive correlation between decisions to redistribute equally and the number of payoff-equalizing spectators in the reference group. Drawing a reference group of $N = 3$ ensures that we achieve wide dispersion of choices in a group, and acceptable statistical power even for the extreme values, without having to deceive the spectators. Although the actions of a group of three people need not be representative of the average sentiment in the complete sample, we argue that this treatment should have an effect for two reasons: First, an extensive literature has documented that people tend to extrapolate small sample properties to large samples (Kahneman and Tversky, 1973; Rabin, 2002; Tversky and Kahneman, 1971), and that they tend to neglect base rates (Benjamin et al., 2019; Pennycook and Thompson, 2016). Second, because the spectators in our experiment have very limited information, even a fully Bayesian updater should give positive weight to the signal. A key feature of our design is that we have exogenous variation in the share of members in the reference group who distribute the bonus equally between workers. This allows us to study, in a clean way, how increasing the strength of the signal about an underlying norm of equalizing income influences the spectators' choices.

2.3 Empirical strategy and analysis

This section outlines the identification strategy and main regressions. Some of the analyses were considered to be of primary interest, while others were pre-registered as exploratory tests, either because they were considered less important or because of statistical power. As is often the case with pre-registered studies, some additional analyses were thought of ex post. These are considered as exploratory analyses.

2.3.1 Pre-registered analysis

We define the outcome variable Y_i as a dummy variable equal to one if spectator i divides the bonus equally between the two workers in the pair, and zero otherwise. Spectator $i's$ initial beliefs about the number of payoff-equalizing members in the reference group is defined as $N_{0,i} \in [0, 3]$, and the information signal received by spectator i is $N_{1,i} \in [0, 3]$.

To estimate a linear effect of the share of the reference group that equalizes earnings on the probability of distributing the bonus equally between stakeholders, we run the following OLS regression:

$$Y_i = \alpha + \beta N_{1,i} + \gamma X_i + \epsilon_i, \tag{2.1}$$

where X_i is a vector of controls, and ϵ_i is a mean-zero error term. The causal interpretation of Equation 2.1 rests on the assumption that the error term is uncorrelated with the social-norm signal:

$$E[\epsilon_i | N_{1,i}] = 0 \quad \forall \, N_{1,i},$$

which is satisfied by our random allocation to treatment. Further, for the main regression, we need a linearity assumption that is not guaranteed by randomization alone.

The average effect from Equation 2.1 could potentially mask two opposing effects: i) an effect for those who receive a positive shock to their initial beliefs, and ii) the effect for those who receive a negative shock to their beliefs. For a non-parametric estimate of the effects of different social norm signals, $N_{0,i}$, we also run the following regression:

$$Y_i = \alpha + \beta_1(N_{1,i} = 1) + \beta_2(N_{1,i} = 2) + \beta_3(N_{1,i} = 3) + \gamma X_i + \epsilon_i, \tag{2.2}$$

where $N_{1_i} = 0$ is the omitted value. Equation 2.2 estimates the average effect of each level of the signal on the likelihood of equalizing income. We assume that people experience a cost of deviating from what they perceive as the social norm. In such case, people are positively (negatively) affected by a positive (negative) signal. $N_{1_i} = 0$

is always a weakly negative signal; it affects spectator $i's$ beliefs about the share of egalitarians either downward or not at all, but never upwards. $N_{1_i} = 3$ is always a weakly positive signal, whereas the direction of the signals $N_{1_i} = 1$ and $N_{1_i} = 2$ depends on the spectator's prior belief. For spectators who are given the positive signal, $N_{1,i} = 3$, and those who are given a negative signal, $N_{1,i} = 0$, an assumption about a desire to conform with others gives a clean prediction about the direction of the behavioral effect.

Equations 2.1 and 2.2 are estimated as OLS regressions, using robust standard errors. We run regressions 2.1 and 2.2 both with and without control variables X_i. As stated in the pre-analysis plan, we regard the regression results without controls as the main test. However, the results with controls are also interesting in themselves, and potentially yield more precise estimates.

Pre-specified control variables are: age, sex, education, and political orientation. Age is given in years, and sex is coded binary with a female indicator variable. Education is measured by an indicator variable that is equal to one if the respondent has at least a two-year college degree. Political orientation is measured by an indicator that is equal to one if the respondents describes themselves as somewhat conservative or very conservative (4 or 5 on a scale from 1 to 5. See the full question text in appendix 2.C.)

2.3.2 Heterogeneity analysis

We run two sets of heterogeneity analyses: i) heterogeneity with respect to initial belief, N_0, and ii) heterogeneity with respect to background characteristics.

To gain a better understanding of who is affected by the treatment, we interact the information treatment, N_1 with the spectator's initial belief, N_0. This gives us an overview of how spectators with different initial beliefs are affected by the information treatment. This is analysed using the following regressions:

$$Y_i = \alpha_i + \beta N_{1,i} + \rho N_{0,i} + \sigma(N_{1,i} \times N_{0,i}) + \gamma X_i + \epsilon_i \tag{2.3}$$

as well as by running regression 2.2 separately for all $N_{0,i} \in [0, 3]$.

As before, we consider the regression without control variables to be the main test. Interacting the social norm signal with spectators' prior beliefs, we expect to see a larger treatment effect for spectators who receive a stronger belief shock from the signal. By belief shock, we mean the difference between prior belief $N_{0,i}$ and the signal $N_{1,i}$.

To test for heterogeneity in the treatment effect with respect to background characteristics, we interact the indicator variables in X_i from Equation 2.1 with the treatment, $N_{1,i}$. In this regression, our age variable is replaced by a dummy, which is equal to one if the spectator is older than the median age in the US, around 39 years. This regression is specified as follows:

$$Y_i = \alpha_i + \beta N_{1,i} + \gamma X_i + \kappa(N_{1,i} \times x_{k,i}) + \epsilon_i \tag{2.4}$$

This regression is run separately for each $x_{k,i}$, where k indicates the specific variable to be interacted with the treatment. In other words, the treatment is interacted with *either* age, sex, education, or political orientation in each regression. In this way, the relevant comparison group is the rest of the sample for the alternative value of one variable at a time, i.e., young vs. old, instead of "young males on the political left, with less than two years of college".

2.4 Results

2.4.1 Descriptive statistics

Figure 2.2 shows the distribution of initial beliefs about the number of payoff-equalizing members in the reference group in our study, compared to the true distribution of signals. Around half of the spectators in our sample believe that 2 out of 3 members of the reference group chose to equalize earnings between a pair of workers in the study by ACT. Less than 20 percent believe that either 0 or 1 member of the reference group chose to equalize earnings. Comparing Figure 2.2 to the true probabilities of $Pr(x = 0) = 0.10, Pr(x = 1) = 0.35, Pr(x = 2) = 0.40$ and $Pr(x = 3) = 0.15$, we see that the participants as a group generally overestimate the share of egalitarians in their reference group. However, almost half of the spectators state beliefs that match the mode in the true distribution. [4]

The majority of participants receive a shock to their initial belief (i.e., the difference between prior belief and the information signal) through the information they receive, as can be seen in Figure 2.3. The figure shows the distribution of individual-level belief shocks. Only about 30 percent of our participants have correct beliefs about their reference group. The majority of spectators receive a negative information shock. This is the result of a general tendency to overestimate the share of egalitarians in the reference group.

Figure 2.4 shows the distribution of choices made in our experiment.[5]The x-axis indicates the amount given to worker B. Worker B is the one who is initially allocated

[4]The Pearson correlation between the individual guess and the actual number of egalitarians in the individual's reference group is 0.051. This is low, as is to be expected when the signal is randomly drawn.

[5]The distribution of bonus earnings allocated to worker B in Almås et al. (2020) is presented in appendix 2.A.

Figure 2.2: This figure shows the distribution of initial beliefs about the share that divided the bonus equally in the reference group. The distribution of beliefs is skewed upwards relative to the true distribution.

0, while worker A receives \$6. Only 1% of spectators allocate more of the bonus to worker B than to worker A. This is strong evidence against spectators randomizing in their distribution choice. A substantial majority of 75% of our spectators distribute the money equally between the two workers. This indicates that spectators are generally unwilling to accept inequalities that are due to luck, supporting the findings of Konow (2000) and ACT. The share that equalizes bonus payments is higher than in ACT (53%). Some of the difference could be explained by the fact that all the spectators in the present study, regardless of the signal they receive, are potentially aware of the social norm of distributing the bonus equally between workers (Krupka and Weber, 2009). Because all spectators in our experiment are asked about how many in the reference group they think distributed the bonus equally, the norm of equal distribution may be

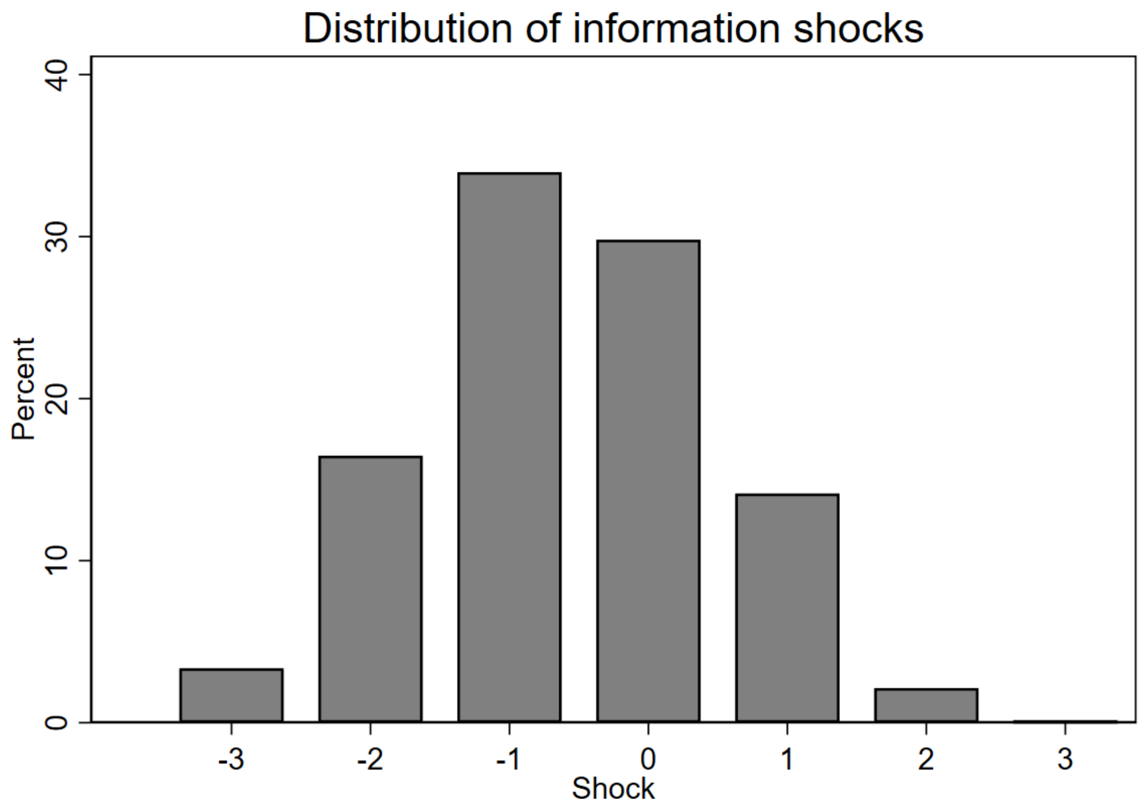


Figure 2.3: Distribution of information shocks. Information shock is defined as the difference between the signal (how many in the reference group that divided the bonus equally) and the initial belief.

more salient in our experiment than in that of ACT.

Figure 2.5 shows, separately for each information group, the share of spectators that distribute the bonus equally. The share is high for all signal groups, between 71 and 81 percent. Those who receive the strongest signal, in terms the number of payoff-equalizing members in the reference group, equalize more. This is consistent with spectators conforming to the perceived social norm followed by others. Receiving a signal that everyone in the reference group splits the bonus equally between workers, tells the spectator that there is a higher probability that distributing the money equally between workers is an empirical norm in the reference population.

Figure 2.6 shows the relationship between what our spectators do and their prior beliefs about what others have done. There is a clear positive relationship here, although

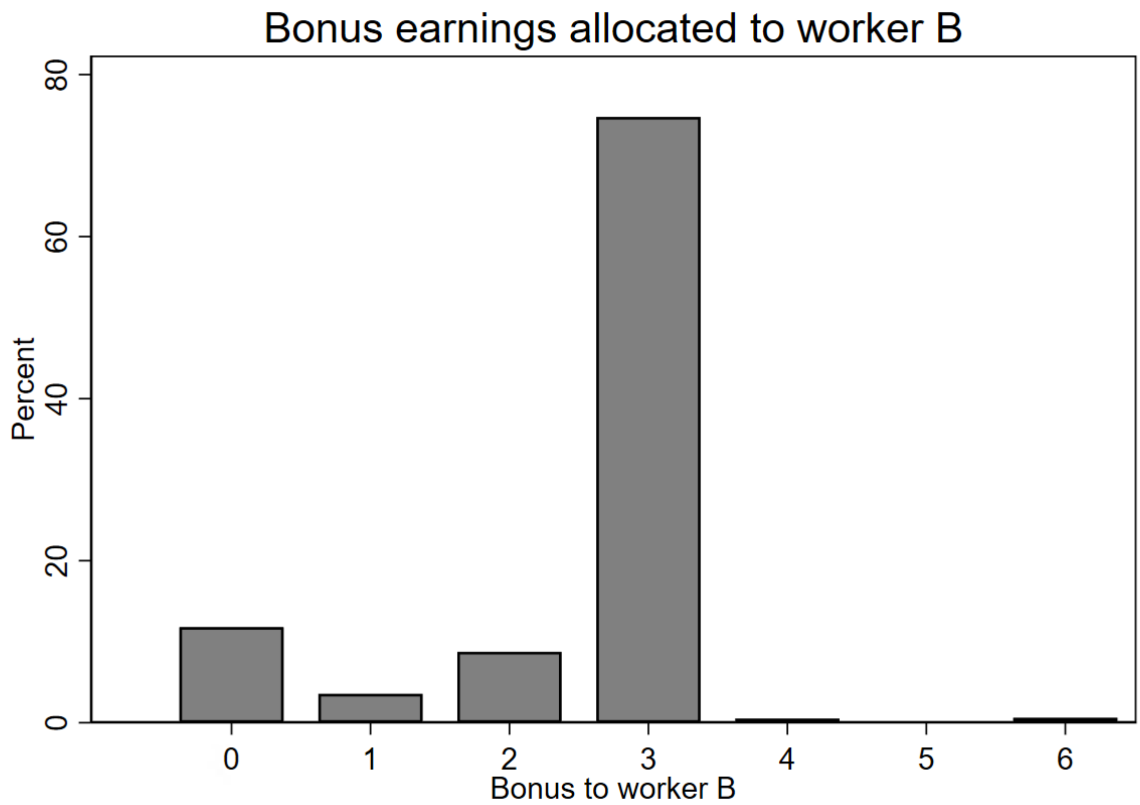


Figure 2.4: Distribution of bonus given to worker B in our data. 75% of participants split the bonus equally between workers

the direction of causality is unclear. One interpretation is that people believe others will think and do the same as themselves, an effect commonly referred to as the false-consensus effec (Engelmann and Strobel, 2000; Faro and Rottenstreich, 2006; Krueger and Clement, 1994; Mullen et al., 1985; Ross et al., 1977; Roth and Voskort, 2014).

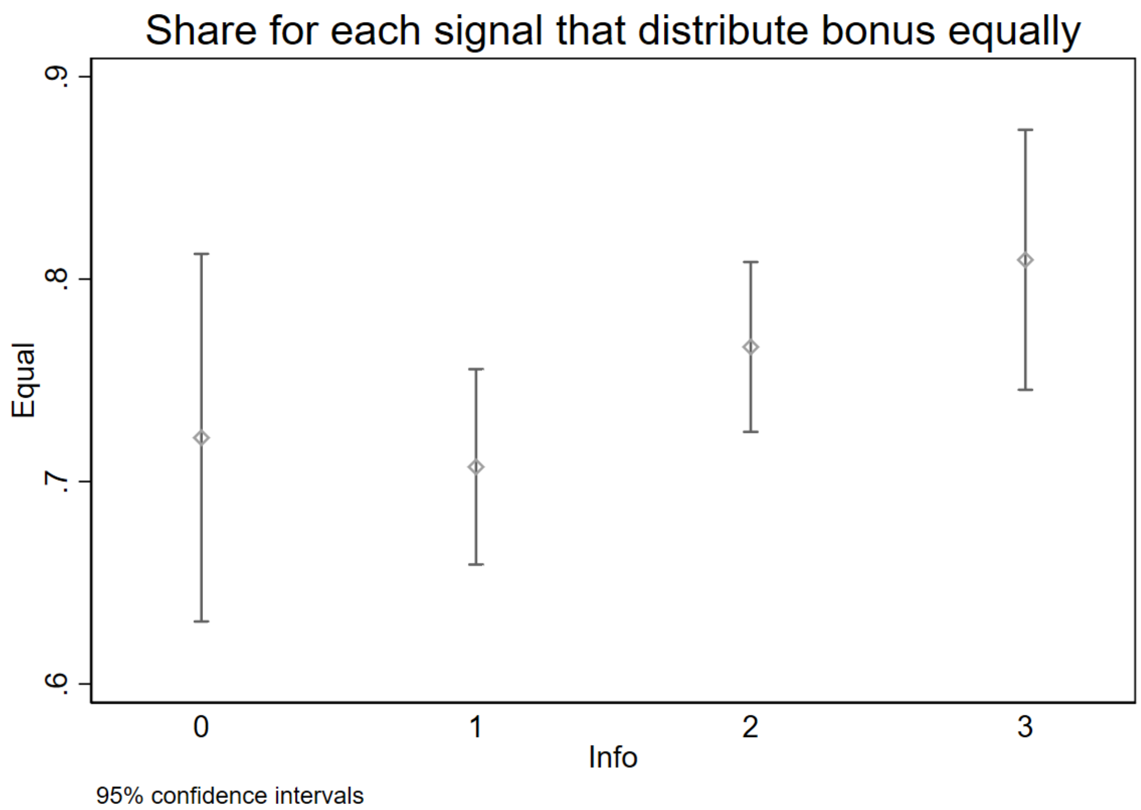


Figure 2.5: Share, by information received, that choose to equalize bonus earnings between workers. The x-axis shows the number of egalitarians in the reference group. The y-axis shows the share of spectators that divide the bonus equally between the pair of workers.

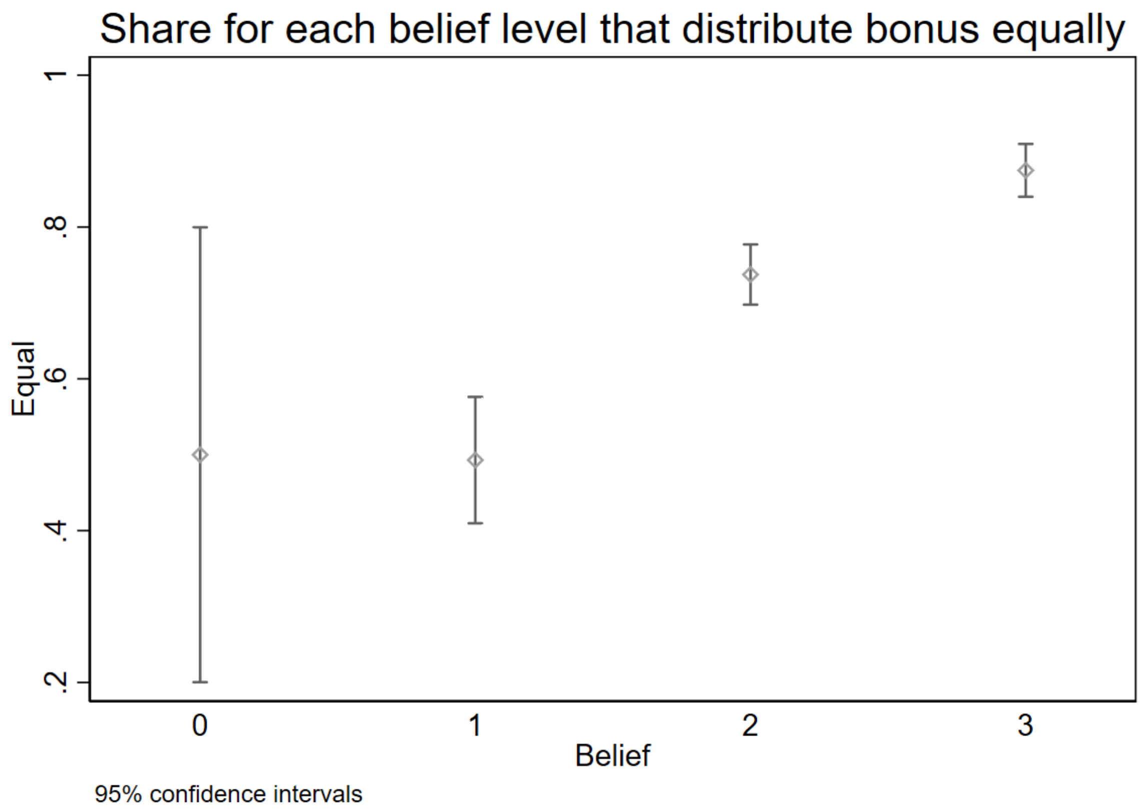


Figure 2.6: Share, by prior beliefs, that choose to equalize bonus earnings between workers. The x-axis indicates the prior belief about how many in the reference group that equalize bonus earnings between workers. The y-axis shows the share of spectators that divide the bonus equally between the pair of workers.

2.4.2 Main analysis

This section presents the findings from the regressions described in the pre-analysis plan as the regressions of primary interest, as specified in section 2.3. To test whether a social norm of equal distribution can be self-reinforcing, we run a linear regression on Equation 2.1, our main analysis. The results from this analysis are reported in Table 2.2. Column 1 in the table shows that increasing the number of payoff-equalizing members in the reference group by one causes a 3.9 percentage point increase in the estimated probability that the spectator in our experiment equalize earnings ($p = 0.016$). Thus, going from 0 to 3 members of the reference group who equalize bonus payments to the two workers is predicted to increase the probability of equalizing earnings by 11.6 percentage points. The effect is robust to controlling for age, sex, education level, and political affiliation, as shown in column 2. We also observe that females are 11 percentage points more likely to divide equally than males,which is consistent with earlier literature (Almås et al., 2020; Cappelen et al., 2022c). Splitting equally also increases with age and education. Political conservatives are slightly less likely to split equally, by 7.3 percentage points ($p = 0.025$)

Table 2.3 presents the non-parametric regression specified in equation 2.2. This regression estimates a positive effect of increasing the number of members of the reference groups who chose to equalize earnings, but the effects are not statistically significant, possibly due to power issues (in the pre-analysis plan, we planned for 80% power specifically for the regression in Table 2.2). In Table 2.3, the baseline is receiving a signal of 0, i.e., that none of the members in the reference group divided the bonus earnings equally between the pair of workers. The other signals are estimated as separate treatment variables. Although the relative size of the coefficients are as predicted, none are statistically significant.

Table 2.2: Main regression — Linear effect

	(1)	(2)
Info	0.0388^{**}	0.0356^{**}
	(0.0159)	(0.0158)
Age		0.00222^{***}
		(0.000854)
Female		0.111^{***}
		(0.0278)
College		0.0531^{*}
		(0.0300)
Conservative		-0.0733^{**}
		(0.0350)
Constant	0.686^{***}	0.513^{***}
	(0.0301)	(0.0541)
Observations	983	982

Note: The outcome variable is an indicator variable equal to one if the spectator equalizes bonus earnings between the two workers. Columns 1 and 2 are without and with background characteristics. "Info" takes discrete values in the interval [0, 3], depending on the number of people in the reference group that chose an equal distribution. "Age" is in years, "Female" is an indicator for being female. "College" is an indicator for having at least a two-year college degree. "Conservative" is an indicator for stating 4 or 5 on a scale from 1 to 5, where 1 is "very politically liberal" and 5 is "very politically conservative". Robust standard errors in parentheses. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

Table 2.3: Main regression — Non-parametric

	(1)	(2)
Info=1	-0.0144	-0.0321
	(0.0518)	(0.0519)
Info=2	0.0448	0.0296
	(0.0504)	(0.0507)
Info=3	0.0879	0.0699
	(0.0560)	(0.0561)
Constant	0.722^{***}	0.559^{***}
	(0.0456)	(0.0635)
Controls	No	Yes
Observations	983	982

Note: The outcome variable is an indicator variable equal to one if the spectator equalizes bonus earnings between the two workers. Columns 1 and 2 are without and with background characteristics. The "Info=X" variables are indicator variables for the number of people in the reference group that chose an equal distribution. Controls include "Age", "Female", "College" and "Conservative". "Age" is in years, "Female" is an indicator for being female. "College" is an indicator for having at least a two-year college degree. "Conservative" is an indicator for stating 4 or 5 on a scale from 1 to 5, where 1 is "very politically liberal" and 5 is "very politically conservative".
Robust standard errors in parentheses. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

2.4.3 Heterogeneity analysis

This section presents heterogeneity analyses that were specified in the pre-analysis plan. We consider these analyses exploratory, because statistical power was calculated for the main analyses, and not for studying heterogeneity.

Table 2.4 presents a regression in which we interact the information signal with prior beliefs, specified in equation 2.3. Column 1 shows regression results without control variables. Here, the coefficient for "Info" is the estimated percentage point effect of increasing the number of egalitarians in the reference group by one unit when "Belief" $= 0$. The estimated coefficient for "Belief" is 0.154 ($p < 0.01$). The coefficient predicts that increasing a spectator's prior belief about the number of egalitarians in the reference group by one increases the probability of distributing the bonus equally by 15.4 percentage points. "Info $\times$ Belief" is the interaction term between "Info" and "Belief". When controlling for, and interacting with, prior beliefs, the coefficient for the info variable is statistically indistinguishable from zero. We find no significant interaction effect between beliefs about redistribution in the ACT sample and the information signal.

In Table 2.5, we interact the treatment variable "Info" with each of the indicator variables for individual background characteristics, as shown in the column headers, as specified in Equation 2.4. From Table 2.5, we see that the effect observed in the main analysis is predominantly driven by political liberals and by individuals with less than two years of college education. There is a positive predicted effect of adding payoff-equalizing members to the reference group for participants who lack a college education: 8.2 percentage points in a linear regression ($p < 0.01$) — and a negative interaction effect of -6.8 percentage points ($p = 0.052$). There is no difference between liberals and conservatives in the estimated likelihood of equalizing bonus payments. We find that

Table 2.4: Info interacted with beliefs

	(1)	(2)
Info	0.0113	0.0210
	(0.0546)	(0.0548)
Belief	0.154***	0.152***
	(0.0409)	(0.0409)
Info × Belief	0.00925	0.00373
	(0.0219)	(0.0220)
Constant	0.361***	0.240**
	(0.0992)	(0.107)
Controls	No	Yes
Observations	983	982

Note: The outcome variable is the estimated probability of the spectator equalizing bonus earnings between the two workers. Columns 2 and 1 are with and without background characteristics. Controls include "Age", "Female", "College" and "Conservative". "Age" is in years, "Female" is an indicator for being female. "College" is an indicator for having at least a two-year college degree. "Conservative" is an indicator for stating 4 or 5 on a scale from 1 to 5, where 1 is "very politically liberal" and 5 is "very politically conservative".
Robust standard errors in parentheses. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

liberals are affected by the signal: the effect is a 4.7 percentage-point ($p < 0.01$) increase in the probability of equalizing, for each extra member in the reference group who equalizes payments. Conservatives, however, seem to be unaffected by the actions of others.

Table 2.5: "Info" interacted with background characteristics

	Older	Female	College	Conservative
Info	0.0338	0.0367	0.0822***	0.0471***
	(0.0267)	(0.0239)	(0.0292)	(0.0177)
Group × Info	0.00438	-0.000592	-0.0671*	-0.0408
	(0.0329)	(0.0313)	(0.0345)	(0.0378)
Group	0.0525	0.111*	0.162**	-0.00737
	(0.0620)	(0.0593)	(0.0666)	(0.0684)
Treatment effect on group	0.0381**	0.0361*	0.0151	0.00622
	(0.0193)	(0.0203)	(0.0186)	(0.0333)
Observations	983	983	983	983

Note: The table presents the treatment effects for different groups as well as interactions, estimated as in Equation 2.4. The outcome variable is the estimated probability that the spectator will equalize bonus earnings between the two workers. "Info" is the number of payoff-equalizing members in the reference group. The regression is run separately with each of the column labels "Older", "Female", "College" and "Conservative" being entered as x_k in the equation. Info is the estimated probability that the spectator will equalize bonus earnings between the two workers. "Group" is one if the participant is in the group indicated by the column header. "Group × Info" is the interaction between the variables "Group" and "Info". "Treatment effect on group" is the linear combination "Info + Group × Info". "Older" is equal to 1 if the participant is older than the median age in the US. "Female" is an indicator for being female. "College" is an indicator for having at least a two-year college degree. "Conservative" is an indicator for stating 4 or 5 on a scale from 1 to 5, where 1 is "very politically liberal" and 5 is "very politically conservative".
Robust standard errors in parentheses. $^{*} : p < 0.1,^{**} : p < 0.05,^{***} : p < 0.01.$

2.4.4 Exploratory analysis

In addition to the pre-registered analysis, we run a set of exploratory analyses primarily aimed at understanding the mechanisms driving the main effect. Figure 2.5 indicates that the probability of equalizing earnings only increases for spectators who receive information that the majority, either 2 or 3 members, of the reference group divided the bonus equally between workers, compared to those who receive a signal that a minority (0 or 1 member) in the reference group divided the bonus equally. Therefore, we test whether being told that the majority in the reference group chose to equalize earnings influences the spectator's decision. We do this by defining an alternative treatment variable, "Majority". The variable "Majority" is an indicator for being told that the majority in the reference group chose an equal split between the two workers. We refer to this as receiving a majority signal.

Columns 1 and 3 in Table 2.6 present the results from running the main regression, specified in Equation 2.1, with the alternative treatment variable "Majority" instead of the variable "Info" used earlier. In columns 2 and 4, we also include the variable "Majority belief", as well as the interaction between "Majority" and "Majority belief". "Majority belief" is an indicator for stating the initial belief that the majority (2 or 3) in the reference group divided the bonus equally. We find a statistically significant, and economically meaningful, effect of receiving information that the majority in the reference group chose to equalize earnings. The predicted effect of receiving a majority signal is 6.8 percentage points, compared to receiving a minority signal ($p < 0.05$ in an OLS regression, with and without control variables). The coefficient for "Majority belief" is around 30 percentage points, and highly significant, both with and without controlling for background characteristics. This shows a high correlation between what spectators do and what they believe others have done.

Table 2.7 presents results from a regression where the treatment variable "Majority" is interacted with each of the indicator variables for individual background characteristics, as shown by the column headers. In Table 2.7, we see a strong interaction effect between spectators' reported education level and the treatment variable; the effect of being exposed to a reference group where splitting bonus earnings equally is the majority choice is 17.2 percentage points among subjects without a college education ($p < 0.01$). The coefficient for the interaction between "Majority" and having at least two years of college education is -15.7 percentage points ($p < 0.01$). This suggests that spectators with less than a two-year college degree are more likely to be swayed by the majority opinion in the reference group, compared to those with at least two years of college education.

Table 2.6: Main regression with "Majority" as an explanatory variable

	(1)	(2)	(3)	(4)
Majority	0.0678^{**}	0.104	0.0657^{**}	0.0907
	(0.0280)	(0.0799)	(0.0277)	(0.0804)
Majority belief		0.326^{***}		0.294^{***}
		(0.0623)		(0.0624)
Majority × Maj. belief		-0.0471		-0.0335
		(0.0849)		(0.0850)
Constant	0.710^{***}	0.438^{***}	0.533^{***}	0.329^{***}
	(0.0216)	(0.0582)	(0.0508)	(0.0697)
Controls	No	No	Yes	Yes
Observations	983	983	982	982

Note: The outcome variable is the estimated probability that the spectator will equalize bonus earnings between the two workers. "Majority" is an indicator for being told that the majority in the reference group chose to divide the bonus equally. "Majority belief" is an indicator for having the prior belief that the majority in the reference group divided the bonus equally. Controls include "Age", "Female", "College", and "Conservative". "Age" is in years, "Female" is an indicator for being female. "College" is an indicator for having at least a two-year college degree. "Conservative" is an indicator for stating 4 or 5 on a scale from 1 to 5, where 1 is "very politically liberal" and 5 is "very politically conservative". Robust standard errors in parentheses. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

In light of the results in Tables 2.6 and 2.7, one interpretation is that our main result (see Table 2.2) is driven by a “tipping point” effect, where spectators who could potentially be affected by treatment only change their behavior when the number of people in the reference group exceeds some threshold (Centola et al., 2018). However, another possibility is that the effect is driven by spectators who receive a large shock to their initial beliefs, while there is no such effect for spectators who receive a smaller shock to their initial beliefs. To investigate this mechanism, we run a regression of the spectator’s decision on a “shock” variable that measures the distance between the signal and prior belief. The results are presented in Table 2.8. We run the regression separately for each level of prior beliefs, presented in each of the four columns. The "Shock" coefficients suggest that positive shocks to initial beliefs predict a higher probability of equalizing earnings in the current setting, although the results are noisy. Only the coefficient for the group who initially believed that 3 out of 3 in their group

Table 2.7: "Majority" interacted with background characteristics

	Older	Female	College	Conservative
Majority	0.0889*	0.0748*	0.172***	0.0793**
	(0.0457)	(0.0429)	(0.0501)	(0.0308)
Group × Majority	-0.0373	-0.0160	-0.156***	-0.0538
	(0.0573)	(0.0556)	(0.0599)	(0.0692)
Group	0.0814*	0.120***	0.138***	-0.0430
	(0.0443)	(0.0429)	(0.0469)	(0.0505)
Treatment effect on group	0.0515	0.0588*	0.0162	0.0255
	(0.0346)	(0.0354)	(0.0330)	(0.0619)
Observations	983	983	983	983

Note: The table presents the treatment effects for different groups as well as interactions, estimated as in Equation 2.4. The outcome variable is the estimated probability that the spectator will equalize bonus earnings between the two workers. "Majority" is an indicator for being told that the majority in the reference group chose to divide the bonus equally. The regression is run separately with each of the column labels "Older", "Female", "College", and "Conservative" being entered as x_k in the equation. Info is the estimated probability that the spectator will equalize bonus earnings between the two workers. "Group" is one if the participant is in the group indicated by the column header. "Treatment effect on group" is the linear combination "Info + Group × Info". "Older" is equal to 1 if the participant is older than the median age in the US. "Female" is an indicator for being female. "College" is an indicator for having at least a two-year college degree. "Conservative" is an indicator for stating 4 or 5 on a scale from 1 to 5, where 1 is "very politically liberal" and 5 is "very politically conservative".
Robust standard errors in parentheses. $^{*}: p < 0.1,^{**}: p < 0.05,^{***}: p < 0.01$.

divided equally is statistically significant at the 5 percent level (Table 8, column 4).

Table 2.8 suggests that the main effect is driven by the subgroup of participants who have a strong initial belief that the majority chose to equalize income between stakeholders. There are two possible ways that this subgroup could drive our main effect. First, the group with initial beliefs equal to 3 have the potential for a large negative shock to their initial belief, and larger shocks could be associated with a larger treatment effect. Second, this group could drive the main effect for reasons unrelated to the size of the information shock. To distinguish between these two alternatives, we perform an analysis in which we limit the sample to the subgroup of spectators with initial beliefs equal to 3. Then, we further limit the sample, using a stepwise procedure, by looking at the effect of the treatment variable on the probability of

Table 2.8: Effect of shock and belief on share that divide equally

	Belief=0	Belief=1	Belief=2	Belief=3
Shock	-0.106	0.0593	0.0154	0.0464^{**}
	(0.166)	(0.0500)	(0.0234)	(0.0200)
Constant	0.652^{**}	0.462^{***}	0.744^{***}	0.938^{***}
	(0.288)	(0.0495)	(0.0220)	(0.0278)
Observations	14	142	476	351

Note: The outcome variable is the estimated probability that the spectator will equalize bonus earnings between the two workers. Columns 2 and 1 are with and without background characteristics. "Shock" is defined as the difference between the information received and prior belief. "Belief" is the participant's guess about how many in the reference group chose an equal distribution. "Age" is in years, "Female" is an indicator for being female. "Conservative" is an indicator for stating 4 or 5 on a scale from 1 to 5, where 1 is "very politically liberal" and 5 is "very politically conservative". "College" is an indicator for having at least a two-year college degree.
Robust standard errors in parentheses. $^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$.

distributing equally for smaller and smaller belief shocks. If the main effect is driven by the spectators who receive larger shocks to their initial beliefs, we should see that the predicted treatment effect becomes smaller and smaller when we restrict the sample further, starting with the full sample.

The first column in Table 2.9 presents the estimations of the treatment effect of the initial "Info" variable on the likelihood of dividing equally for the full sample with initial beliefs equal to 3. Next, in column 2, we restrict the sample to only participants in this subsample who receive a belief shock less than or equal to two units, using the same regression. In column 3, we restrict the sample to participants who receive a belief shock less than or equal to one unit. In column 4, the coefficient cannot be estimated, as the variables "Info" and "Shock" are perfectly collinear for this subsample. We observe the largest treatment effects in column 1 and column 2. In column 3 (for shocks less than or equal to one unit) there is just a small predicted effect of 2.26 percentage points which is not statistically significant. This is consistent with the explanation that the subsample with initial beliefs equal to 3 drive the main effect because they receive the largest shocks to their initial beliefs.

Table 2.9: Linear effect with restricted sample

	(Full sample)	(-2≥Shock≥2)	(-1≥Shock≥1)	(Shock=0)
Info	0.0464**	0.0646**	0.0226	0
	(0.0200)	(0.0255)	(0.0414)	(.)
Constant	0.798***	0.760***	0.863***	0.931***
	(0.0415)	(0.0549)	(0.0990)	(0.0336)
Observations	351	318	200	58

Note: The sample is restricted to spectators who believe that all in members in their reference group divided the bonus equally. The outcome variable is the estimated probability that the spectator will equalize bonus earnings between the two workers. "Info" takes discrete values in the interval [0, 3], depending on the number of people in the reference group who chose an equal distribution. The sample is restricted stepwise based on the difference between the initial belief and the information received. Robust standard errors in parentheses. $^{*}: p < 0.1,^{**}: p < 0.05,^{***}: p < 0.01$.

2.5 Conclusion

This paper studies whether social norms concerning the redistribution of income can be self-reinforcing in an experimental setting. This is done by exogenously varying information about the number of members of a reference group who chose to divide money equally between two workers in a previous experiment. We are primarily interested in how the signal about the share of egalitarians influences the probability of equalizing payoffs between two stakeholders. We find a small but statistically significant effect of the number of payoff-equalizing group members of a reference group on the estimated probability of equalizing payoffs. Initial beliefs about equalizing behavior in the reference group are strongly correlated with the decision to implement equality in the current setting.

We show suggestive evidence that the main effect could be driven by the subgroup of participants who receive the largest shocks to their initial beliefs. One possible explanation for this, is that participants are torn between their own fairness ideal and the desire to conform with the perceived social norm. When spectators receive a signal which indicates that their initial beliefs are inaccurate, they modify their behavior in the direction of the perceived social norm. This effect is stronger if the information is very far from their initial beliefs, than if the information is just a slight modification of their initial beliefs. Another possible explanation is that participants initially plan to act according to the social norm they believe to apply in the current setting, but when they receive information running contrary to their initial beliefs, they think that another norm applies. Therefore, they modify their behavior to conform with this norm. For instance, a spectator with initial beliefs equal to 3, could have a strong belief that a norm of egalitarianism applies, and plan to redistribute equally. When this spectator receives a signal which indicates that just 1 out of 3 in their reference group chose to equalize earning, they may think that the norm is to not redistribute, and then adjust their behavior in the direction of this norm.

One proposed explanation in the literature on why people adhere to social norms is that people care about their social image - that is, how their actions will be perceived by others (Andreoni and Bernheim, 2009; Bernheim, 1994; Bursztyn and Jensen, 2017). In the setting studied in the present paper, people are not observed by others, but only observe the actions of others. Nevertheless, we find that redistributive decisions are affected by signals about social norms in a reference group. One possible interpretation of our findings is that people may care about their self-image - how they perceive themselves - in addition to their social image. The salience of self-image may be amplified by awareness of being situated in a social setting. This can increase the importance of self-image. It may also cause participants to envisage how their actions would be perceived if the reference group were to observe their action.

We contribute to two main parts of the literature. First, the paper adds to the literature on social pressure and social norms. We show that people are swayed by the actions of the majority even in a completely anonymous setting, where the only cue about the reference sample is that they are "A large sample of Americans". Thus, our paper contributes by showing that, even under minimal conditions, a social norm of equal distribution can be self-reinforcing, by making individuals aware of the actions of their peers. Second, the paper contributes to the broad literature on moral preferences. In a setting almost identical to one of the treatments in Almås et al. (2020), we replicate the finding that people are generally unwilling to accept inequalities that are exclusively due to luck.

We also add to the evidence that people's actions are affected by social norms in various domains. The findings are consistent with a descriptive norm effect, by showing that people are affected by the actions of a reference group even when the only thing the spectators know about the reference group is that they are Americans. Our findings inform the growing literature on social norms and beliefs. For instance, Bursztyn et al. (2020a,b) show that people tend to forgive others who publicly express xenophobic views as long as this conforms with a social norm. Although from a different domain, our results fit with the general idea that people may be sensitive to cues that their actions are socially acceptable.

Our study cannot fully isolate why information about social norms matters in anonymous settings. We believe that future research should endeavor to shed further light on this issue. One possibility is that the spectators care about what they think the workers perceive as fair, and that the choices of the spectators made in a previous experiment function as signals that inform the current spectators about the fairness views of the stakeholders. It would be interesting to conduct an experiment where information about the norm held by workers influences the spectator. That is left to future research.

2.A Additional figures

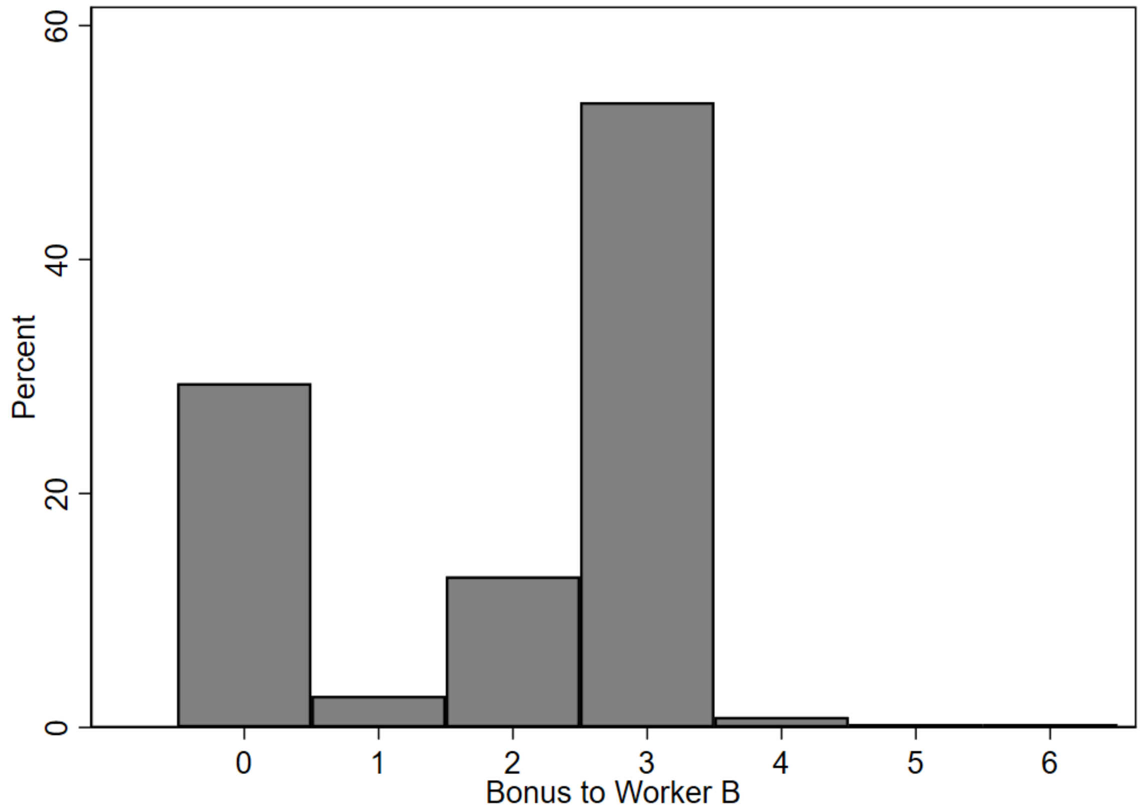


Figure 2.7: This figure shows the distribution of bonus earnings allocated to worker B in Almås et al. (2020).

2.B Worker Instructions

Welcome!

Please note that your participation will be registered on the following Amazon Mechanical Turk worker ID:

[*Auto generated worker ID*]

The worker ID was retrieved automatically when you clicked on the link that brought you here. This step is necessary for assigning payments to the right account and to ensure that you only participate in this study once.

[*New page*]

General instructions

Welcome to this research project! We very much appreciate your participation.

Your payment for taking the HIT will be sent to you shortly after the completion of this HIT. You get a fixed participation fee of 1 USD and may, depending on the actions you and others take, earn additional money.

The study consists of two parts, and you will be given instructions on your screen before every single part of the survey. Please always make sure to read the instructions carefully before you continue.

Participation

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate.

Confidentiality

Your worker ID will only be used for payment purposes, and will not be stored with the data collected. Anonymized data from the survey might be shared in open science repositories.

Verification

At the end of this survey, you will be given a **completion code**. You will need to copy this code to the survey code field on the AMT web page that directed you here at the beginning.

Questions about the Research

If you have questions regarding this study, you may contact

thechoicelab@nhh.no

[*New page*]

I have read and understood the above consent form and desire to participate in this study.

Yes/No

[*New page*]

The first part of the experiment is a production phase where you are given an assignment to work on. Go on to the next page to receive instructions for this assignment.

[*New page*]

Assignment:

In this assignment you are asked to work on ten sentence unscrambling tasks. There is no right or wrong answer, but we ask that you make your best effort and write a sentence on each task before moving on to the next.

Description of the assignment:

You will be shown five English words and are asked to form a sentence or an expression using four of these words. This means that each sentence or expression must only contain four words.

For example, if the words given to you are **"sky, blue, is, the, old"**, then you can construct the sentence:

the sky is blue

Write the sentence or expression that you form into the blank space using your keyboard.

When you have read and understood the instructions press » to start the assignment.

[*New page*]

Question 1

CUP THE OFF LIGHTS TURN

[*Open textbox*]

[*New page*]

Question 2

FEET TREE HAVE ELEPHANTS BIG

[*Open textbox*]

[*New page*]

Question 3

OFF PAYS HARD SCHOOL WORK

[*Open textbox*]

[*New page*]

Question 4

GIRL A BOY THOMAS IS

[*Open textbox*]

[*New page*]

Question 5

DISCIPLINED MAN FLOWER THE WAS

[*Open textbox*]

[*New page*]

Question 6

PEACEFUL FELT BIRD SHE VERY

[*Open textbox*]

[*New page*]

Question 7

EVENING NICE A SAD SUMMER

[*Open textbox*]

[*New page*]

Question 8

LOOKED CHILD WINTER THE DEJECTED

[*Open textbox*]

[*New page*]

Question 9

AUDIENCE WAS DOG ECSTATIC THE

[*Open textbox*]

[*New page*]

Question 10

WENT THE SHE SCHOOL TO

[*Open textbox*]

[*New page*]

You have completed the assignment.

Your payment for the assignment is decided by a lottery, where you are paired with one other participant. One of you are randomly chosen to receive a bonus of 6 USD, and the other gets no bonus. However, a third person, in a different study to be conducted in a few days time, will get the opportunity to redistribute earnings.

Click >> to reveal the result of the lottery

[*New page*]

Each worker is shown either of the two messages, each with probability one-half:

Either

You have been drawn to receive a 6 USD bonus payment and the other person in your pair gets 0.

or

The other person in your pair has been drawn to receive a 6 USD bonus payment and you get nothing.

[*New page*]

Within three weeks, we will pay you once your final earnings have been decided. Finally, if you have any comments or feedback to this research please write them in the box below. This is helpful to our research.

[*Open textbox*]

2.C Spectator instructions

Introduction

Welcome to this research project! We very much appreciate your participation.

You will be given instructions on your screen before every single part of the survey. Please always make sure to read the instructions carefully before you continue. At the end of this survey you will be redirected to Prolific to have your submission registered.

Participation

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate.

Payment

Your payment for taking the survey will be sent to you shortly after the completion of this survey. You get a fixed participation fee of 1,1 USD and may, depending on the actions you take, earn additional money.

Confidentiality

Your worker ID will only be used for payment purposes, and will not be stored with the data collected. Anonymized data from the survey might be shared in open science repositories.

Questions about the Research

If you have questions regarding this study, you may contact thechoicelab@nhh.no

[*New page*]

I have read and understood the above consent form and desire to participate in this study.

Yes/No

[*New page*]

Your Prolific ID: Please note that this field should auto-fill with the correct ID

[*Automatically generated Prolific ID*]

[*New page*]

Before proceeding to the main questions, you must answer a simple question about cities. Here, you must simply select 'London'. This is an attention check.

Based on the text above, which city have you been asked to enter?

London

Paris

Frankfurt

Madrid

Vienna

[*The order of the alternatives is randomized*]

[*New page*]

Instructions

In contrast to traditional survey questions that are about hypothetical situations, we now ask you to make a choice that has consequences for a real life situation. A few days ago, two individuals, let us call them worker A and worker B, were recruited via an international online labor platform to conduct an assignment. They were each offered a participation compensation of 1 USD regardless of what they were paid for the assignment.

After completing the assignment, they were told that their earnings from the assignment would be determined by a lottery. The worker winning the lottery would earn 6 USD for the assignment and the other worker would earn nothing for the assignment. They were informed about the outcome of the lottery. However, they were also told that a third person would be informed about the assignment and the outcome of the lottery, and would be given the opportunity to redistribute the earnings and thus determine how much they were paid for the assignment.

With some probability, you are the third person and we now want you to choose whether to redistribute the earnings for the assignment between worker A and worker B. Your decision is completely anonymous. The workers will receive the payment that you choose for the assignment within a few days, but will not receive any further information.

[*New page*]

Before making your choice, you will be informed about the choices of a group of real people who have faced the same decision. In a previous study, a large sample of Americans answered the same question that we need you to decide.

The people in the mentioned study chose from the following alternatives, the very same that we will ask you to choose from:

Not redistribute:

Worker A gets 6 USD and worker B gets 0

Redistribute:

Worker A gets 5 USD and worker B gets 1

Worker A gets 4 USD and worker B gets 2

Worker A gets 3 USD and worker B gets 3

Worker A gets 2 USD and worker B gets 4

Worker A gets 1 USD and worker B gets 5

Worker A gets 0 USD and worker B gets 6

We will shortly show you the decisions made by a randomly drawn group of three individuals from the sample in the earlier study. In this group, how many of the participants do you think chose to distribute the money equally, i.e. 3 USD to worker A and 3 USD to worker B? If your answer is correct you will earn an extra 10 cents.

- 0
- 1
- 2
- 3

[*New page*]

In your reference group [none] of the participants chose to divide the bonus equally.

*This page includes the word **none** with 10 percent probability, **one third** with 35 percent probability, **two thirds** with 40 percent probability, and **all** with 15 percent probability.*

[*New page*]

Now, we need you to make a decision for a real pair of workers.

Worker A won the lottery and earned 6 USD for the assignment, thus worker B earned nothing for the assignment.

Please state which of the following alternatives you choose:

I do not redistribute: Worker A is paid 6 USD and worker B is paid 0 USD

I redistribute: Worker A is paid 5 USD and worker B is paid 1 USD.

I redistribute: Worker A is paid 4 USD and worker B is paid 2 USD.

I redistribute: Worker A is paid 3 USD and worker B is paid 3 USD.

I redistribute: Worker A is paid 2 USD and worker B is paid 4 USD.

I redistribute: Worker A is paid 1 USD and worker B is paid 5 USD.

I redistribute: Worker A is paid 0 USD and worker B is paid 6 USD.

[*New page*]

You have completed the first part of the survey. We would now like to ask you some more questions before we conclude this survey.

[*New page*]

What is the highest level of education you have completed?

- Less than High School
- High School / GED
- Some College
- 2-year College Degree
- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

[*New page*]

Would you describe yourself as politically on the "left" (eg. a liberal) or on the "right" (eg. a conservative)?

1 - Very liberal	2	Neutral	4	5 - Very conservative

[*New page*]

Finally, if you have any comments or suggestions related to this study please write them down in the field below. Your feedback is important to improve our research.

[*Open textbox*]

[*New page*]

Thank you for taking part in this study. Please click the button below to be redirected to Prolific and register your submission.

Chapter 3

Fairness and attribute range: An experimental study of range-based context effects in fairness decisions

Kjetil Røiseland Madland *

Abstract

This paper reports from a novel, large-scale online experiment that tests for range-based context effects on fairness decisions. The experiment is motivated by the focusing model of Kőszegi and Szeidl (2012), and the relative thinking model of Bushong et al. (2021). The two models make opposite predictions about the reaction to an expansion of the utility range in choice dimensions. In the experiment reported here, decision-makers take the position of impartial spectators and make distributive decisions for stakeholders. Through a variety of treatments, distribution alternatives that increase the range of

*FAIR, Department of Economics, NHH Norwegian School of Economics
I would like to thank Kjetil Bjorvatn, Alexander W Cappelen, Mathias Ekström, Samuel Hirshman, Botond Kőszegi, Krishna Srinivasan, Pablo Ignacio Soto-Mota, Eirik André Strømland, Oda Sund, Erik Ø. Sørensen, Bertil Tungodden, Weijia Wang, Alexander Willén, and Florian Zimmermann for helpful comments and suggestions. I also owe huge thanks to Sebastian Fest for facilitating the data collection. Further, I want to thank FAIR, the Department of Economics at NHH, and the Research Council of Norway for the funding that has enabled this research. The study was approved by the Institutional Review Board committee at NHH Norwegian School of Economics, and was pre-registered with the Registry for Randomized Controlled Trials operated by the American Economic Association: RCT ID AEARCTR-0007057.

outcomes for one stakeholder are added to the choice set. I find no evidence of range-based context effects on fairness decisions. Compared to the broad literature on context effects on consumer choices, my findings suggest that fairness preferences are more stable, and not as easily manipulated by the choice context.

3.1 Introduction

Are fairness preferences stable and well-defined, or are they malleable with respect to the choice-set composition? Fairness judgements and distributive decision-making are ubiquitous in politics, parenting, and social interactions in general. Understanding what agents consider to be fair and unfair, and which processes govern fairness judgements, is crucial in order to understand and predict a wide range of social interactions. However, we still only have limited knowledge about how stable or malleable fairness preferences are, and what agents focus on when they make fairness decisions. Choice-set effects can have implications for how we study voting behavior, political economy more generally, and charitable giving, among other topics. This paper studies the effects of focusing, attention, and relative thinking on people's fairness considerations. This is done in a spectator experiment motivated by two competing theories of focusing and relative thinking.

Specifically, the paper studies the following research question: Are distributive decisions affected by the range of possible outcomes in the available choice set? I set up a controlled experiment where participants take the role of impartial spectators, who decide how money is to be distributed between anonymous workers. By adding alternatives that widen the range of possible outcomes in the choice set for one of the workers, I study range-based context effects in a fairness setting. We have seen from the extensive literature that choice-set effects are prevalent in consumer choices.[2] In this

[2]See, e.g., Azar (2011); Bordalo et al. (2013); Desvousges et al. (1993); Ekström (2021); Frederick

literature, adding alternatives to a choice set is often found to alter the relative ranking of the other alternatives in the choice set. In particular, it is observed that adding an alternative C, which dominates A but not B, increases the frequency with which A is chosen over B. This is known as the asymmetric-dominance effect, or attraction effect.

While evidence from the consumer literature offers valuable insight into general processes and the workings of human perception and decision-making, too little is known about decisions that mainly involve fairness. Fairness decisions may have similarities with consumption decisions, but are arguably different in important ways. For example, fairness decisions often evoke strong moral feelings that are not typically prominent in consumption decisions. The relevant choice dimensions in a fairness decision may also be harder to compare than material dimensions. Previous studies have found evidence of context effects on decisions concerning moral choices (Cherry et al., 2002; Desvousges et al., 1993; List, 2007). List (2007), for example, finds that dictators are less likely to transfer money when the choice set includes taking. However, this manipulation changes the framing of the game and potentially the perceived entitlement to the money. In the present paper, the range of outcomes is expanded, while keeping the framing of the distribution and entitlement to the money constant. The paper is also the first, to my knowledge, to study context effects in a choice setting where an impartial spectator distributes money between other people.

The aim of the paper is to explore one specific avenue through which fairness decisions may be affected by the composition of the choice set, namely manipulating the range of the choice dimensions. Here, a choice dimension is understood as an attribute, like the income of a specific individual, equality, efficiency, etc. The terms *dimension* and *attribute* are used interchangeably throughout the rest of the paper. The range of an attribute is formally defined in section 3.2. Colloquially speaking, the range

et al. (2014); Heath and Chatterjee (1995); Huber and Puto (1983); Huber et al. (1982); Kahneman (2003); Ratneshwar et al. (1987); Seshadri et al. (2019); Simonson (1989); Simonson and Tversky (1993); Somerville (2020); Tversky (1972, 1977); Tversky and Kahneman (1981); Wedell and Pettibone (1996).

of an attribute is the difference in utility for the decision-maker between the value of that attribute in the choice set or context that yield the highest and lowest utility, respectively.

Choice-set effects are widely studied both in economics and in psychology. Even though there are many explanations and hypotheses about mechanisms in the psychology literature, it is only quite recently that generally applicable models of economic choice, predicting choice-set effects under certain conditions, have emerged (See, e.g., Azar (2007); Bordalo et al. (2012, 2013); Bushong et al. (2021); Cunningham (2013); Kőszegi and Szeidl (2012)). The experiment reported in this paper is motivated by two specific models of range-based context effects: the *focusing model* of Kőszegi and Szeidl (2012) and the *relative thinking model* of Bushong et al. (2021).

The focusing model predicts that agents focus more on, and thus over-weight, dimensions in which their choices differ more, i.e., dimensions that have a wider range of outcomes. Thus, the relative decision weights of different choice dimensions are affected by the relative range of potential outcomes in the same dimensions. Earlier literature has documented that agents in some settings over-weight choice dimensions that have more variation (see Anderson (1981); Schkade and Kahneman (1998); Wedell (1991)). However, the phenomenon has only recently been formalized for generally applicable choice models. Some studies have tested the focusing model specifically. Dertwinkel-Kalt et al. (2022) test the focusing model in a setting of inter-temporal optimization. They find that workers commit to too much overtime work spread over time, in order to reap a benefit that is concentrated in time. Their findings are consistent with the inter-temporal concentration bias predicted by the focusing model. Nunnari and Zápal (2017) adapt the focusing model to political economy. They show that, if voters have an unbalanced focus, political parties may not try to win the median voter, but instead cater disproportionately to the preferences of more extreme groups.

The relative thinking model builds on the same theoretical framework as the focusing model. However, in the relative thinking model, expanding the range in one dimension affects how agents perceive changes along this dimension. Both models agree that utility for the decision-maker is separable across dimensions, and that the range of a dimension k is the difference in utility obtainable from dimension k between the alternatives in the choice set that are best and worst on dimension k. In the relative thinking model, however, "a fixed difference looms smaller when compared to something large than when compared to something small" (Bushong et al., 2021). Thus, if the range of possible outcomes is expanded in one choice dimension, agents become less sensitive to changes in this dimension. This is consistent with evidence and examples from Savage (1954); Soltani et al. (2012); Thaler (1999); Tversky and Kahneman (1981).

The focusing model and the relative thinking model offer readily available modifications of a standard separable utility model to describe economic choice. Importantly, the two models share the assumption that choice alternatives, even unchosen ones, which expand the range of utility obtainable from a choice dimension, might affect agents' revealed preference relations between other alternatives in the choice set. However, the focusing model and the relative thinking model make different assumptions about the effect of expanding the range of a choice dimension. Therefore, they make opposite predictions concerning the direction of the effect of expanding the outcome range in a single choice dimension.[3]

To provide an illustration of range-based context effects on fairness decisions, consider the following example: Mark is a member of parliament. Mark is advocating for a new, nationally funded opera house in his constituency. The opera house provides large benefits for the people in Mark's constituency, while the costs are shared by all

[3]The focusing model and the relative thinking model represent what the psychology literature has dubbed *dimensional-weight models* and *value-shift models*, respectively (see Wedell (1991) for a discussion).

taxpayers in the country. Thus, the benefits are concentrated to relatively few people, while the costs are dispersed between many. Imagine that Mark does not think that his proposition will be approved in the parliament. Let us assume that the other members of parliament act as benevolent planners, with no self-interest. According to the focusing model, Mark will have a better chance of having his proposition approved if he proposes a scaled-up version of the opera house project, perhaps including a park surrounding it. Making the project larger increases the benefits for the few, and the costs for the many. If the per-person benefit to the people in Mark's constituency increases more than the per-taxpayer cost, the range of potential outcomes increases more for the users of the park and opera house than for taxpayers. In that case, the focusing model predicts that members of the parliament will focus more on, and thus over-weight, the benefits when the project is scaled up. On the other hand, because the costs are more dispersed, the cost side receives less focus as the project is scaled up.

The relative thinking model predicts that increasing the size of the project will have the opposite effect. According to this model, because the range of potential outcomes increases more for the users of the planned opera house than for the taxpayers, any given difference for the users will appear smaller than a similar difference for taxpayers. Thus, this model suggests that Mark ought to scale *down* the project rather than scale it up, if he wants to get the project approved. According to the relative thinking model, scaling down the project will increase its likelihood of being approved. Common to both the above scenarios is that the two models make opposite predictions about who will be favored when the ranges of potential outcomes are differentially changed for a group of individuals.

The present paper studies the effects of range-based choice-set effects in a stylized, incentivized, online experiment. To isolate the range effect on fairness, I use a spectator design, a key feature of which is that it allows the researcher to study people's moral preferences directly in situations without self-interest (Aguiar et al., 2013; Almås et al., 2020; Cappelen et al., 2007, 2013, 2022c; Konow, 2009; Konow et al., 2020; Lane, 2022; Mollerstrom et al., 2015).[4]

The paper relates to several strands of the behavioral literature: First, it contributes to the literature on range-based context effects (Bushong et al., 2021; Dertwinkel-Kalt et al., 2022; Gabaix, 2014; Kőszegi and Szeidl, 2012; Nunnari and Zápal, 2017; Somerville, 2020), by studying complete rankings of alternatives from a menu. I test the effects of expanding the range for different stakeholders in a distributive setting in a controlled experiment. This relates to Somerville (2020), who studies consumption behavior. Using data from an incentivized lab experiment, he structurally estimates a model by means of a horse race between the focusing model, the relative thinking model, and salience theory (Bordalo et al., 2012, 2013). His results are consistent with the relative thinking model, and inconsistent with the others. Salience theory predicts many of the same effects as both the relative thinking model and the focusing model. In many situations, it makes predictions that are in line with either the focusing model or the relative thinking model, depending on the situation. A key difference between salience theory and the two other models is that, in salience theory, choice-set effects come from deviations from a reference point, rather than from differences in attribute range, which is the focus of the present paper. The present paper is the first to test predictions from the focusing model and the relative thinking model in a distributive setting using a spectator experiment. The paper is also among the first to study choice-set effects on fairness decisions more generally.

Second, the paper also contributes to the broader literature on context effects in eco-

[4]A recent study shows that own payoff can itself be an important factor in the formation of beliefs about the sources of inequality and poverty (Somville et al., 2020). In the much-studied dictator game and public goods game, fairness ideals are traded off against one's own payoff. Endogeneity of fairness ideals with respect to one's own endowment, as suggested by Somville et al. (2020), may pose problems in games where the decision-maker is also a stakeholder. These potential problems are avoided by using a spectator design.

nomic decision-making (Bordalo et al. (2013); Desvousges et al. (1993); Ekström (2021); Frederick et al. (2014); Heath and Chatterjee (1995); Huber and Puto (1983); Huber et al. (1982); Kahneman (2003); Mazar et al. (2014); Ratneshwar et al. (1987); Simonson (1989); Simonson and Tversky (1993); Somerville (2020); Tversky (1972, 1977); Tversky and Kahneman (1981); Wedell and Pettibone (1996)), by showing how the relative ranking of distribution alternatives is affected by the introduction of a new distribution alternative in the choice set. In the experiment reported in this paper, the complete ranking of alternatives is incentivized through consequential choices for others. This allows for clean identification of the effect of adding alternative C on the relative ranking of alternatives A and B.

Finally, the paper contributes to the broad literature on fairness preferences and the moral mind in general (see, e.g., Alesina and Angeletos (2005); Almås et al. (2020, 2010, 2011); Andreoni and Miller (2002); Bolton and Ockenfels (2000); Cappelen et al. (2007, 2013, 2019, 2022c); Charness and Rabin (2002); Durante et al. (2014); Engelmann and Strobel (2004); Fehr and Schmidt (1999); Fehr and Gächter (2000); Konow (2000, 2009); Konow et al. (2009); Rabin (1993)) by studying how moral preferences respond to changes in choice-set structure. Understanding what people consider to be fair, and how fairness preferences might be influenced by the choice context, is important to predicting a wide range of social interactions.

The rest of the paper is organized as follows: Section 3.2 outlines the theoretical framework of the focusing model and the relative thinking model, and applies them to fairness decisions. Section 3.3 describes the experimental design and the participants. Section 3.4 discusses the research strategy and identification, as well as the predictions made from theory. Section 3.5 is dedicated to the empirical analysis and discussion of the results. Lastly, section 3.6 concludes.

3.2 Theoretical framework

This section outlines the theoretical framework for the focusing model (Kőszegi and Szeidl, 2012) and the relative thinking model (Bushong et al., 2021), and shows how the two models apply to fairness decisions. First, I present the general formal framework, before outlining the differences between the models and discussing how they generate different predictions of behavior.

The spectator's welfare is described by a separable welfare function

$$U(c) = \sum_{k=1}^{K} u_k(c_k), \tag{3.1}$$

where c is a $k-$dimensional vector of dimensions or attributes, chosen from the finite choice set $C \subset \mathbb{R}^K$. Thus, any choice outcome can be represented by a vector $c = c_1, ..., c_K$, where each dimension k is seen as a choice attribute. $u_k(c_k)$ is the utility or welfare obtained from dimension k.

If agents have an unbalanced focus or are relative thinkers, they do not maximize welfare. Rather, they maximize focus-weighted utility:

$$\tilde{U}(c, C) = \sum_{k=1}^{K} g_k \cdot u_k(c_k), \tag{3.2}$$

where $\tilde{U}(c, C)$ is the focus-weighted utility of the k-dimensional consumption vector c chosen from the choice set $C \subset \mathbb{R}^K$.

The focus-weight is a function of the range (i.e., the span) of outcomes in the agent's consideration set. For simplicity, assume that the consideration set is equal to the available choice set.[5] Formally, the weight of dimension k, g_k, is given by $g_k =$

[5]See, e.g., Bordalo et al. (2020); Kőszegi and Szeidl (2012) for discussions of how the consideration

$g(\Delta_k(C))$, where

$$\Delta_k(C) = \max_{c' \in C} u_k(c'_k) - \min_{c' \in C} u_k(c'_k) \tag{3.3}$$

That is, the weight of dimension k is a function of the range of utility obtainable from dimension k in the choice set.

In the focusing model, the function $g(\Delta)$ is assumed to be strictly increasing in Δ, i.e., that choice dimensions with more variation attract disproportionately more focus from the decision-maker. In the relative thinking model, the function $g(\Delta)$ is assumed to be decreasing in Δ. The relative thinking model predicts that a given change will appear smaller in a choice dimension with a wide range of outcomes than in a dimension with a narrow range. Both models share the fundamental assumption that focus or relative thinking does not affect welfare directly, but that it does affect the choices that agents make.

To illustrate, consider a situation where an impartial spectator is to choose a distribution of money to a group of stakeholders, from a choice set C. Kőszegi and Szeidl (2012) propose that different individuals' utilities are the relevant choice dimensions in social choices.[6] Assume that the spectator is benevolent, they care about the income of each stakeholder $k \in [1, K]$, and their utility from any given distribution is represented by the separable utility function in Equation 3.1, where the K stakeholders are the relevant choice dimensions. If the spectator is prone to unbalanced focusing, or is a relative thinker, they maximize the focus-weighted utility function in Equation 3.2.

Now imagine adding to the choice set C an alternative (i.e., a vector of individual payoffs) $c = c_1, ..., c_k, ..., c_K$, which increases the range of outcomes for stakeholder k, as defined by Equation 3.3, while keeping the ranges for the other stakeholders unchanged. This has opposite effects in the two models:

set may differ from the available choice set.

[6]It is easy to imagine that fairness and efficiency may enter as separate dimensions. This is discussed in section 3.3.

- In the focusing model, the increased range makes the spectator focus *more* on the outcome for stakeholder k and assign a *larger* decision weight g_k to stakeholder k.
- In the relative thinking model, the increased range makes the spectator focus *less* on the outcome for stakeholder k and assign a *lower* decision weight g_k to stakeholder k.

Although the theoretical framework is the same in both models, the two models aim to capture different cognitive processes. In the focusing model, it is focus, or the lack thereof, which drives the model. The idea is that when agents optimize over multiple dimensions, dimensions with little variability in the choice set tend to be ignored or under-weighted. As a result of limited attention or processing power, focus is unbalanced between choice dimensions, depending on their respective ranges.

The relative thinking model aims to capture a tendency for comparisons to be made not only in absolute terms, but also in relative terms. Bushong et al. (2021) argue that decision processes are prone to the same biases as visual perceptions. Thinking in relative terms makes a fixed amount loom smaller when compared to a relatively wide range than when compared to a narrow range, just as a line on a piece of paper may seem longer or shorter depending on the length of surrounding lines. Bushong et al. (2021) draw a distinction between proportional thinking and range-based thinking. While the former judges changes relative to a reference point, the latter, they argue, makes all changes along a dimension seem smaller when the range is greater. I proceed to outline how the models work in a distributive setting, first with two individuals, and then with more than two individuals.

3.2.1 Two-dimensional distributive choices

For simplicity's sake, consider a situation in which a spectator divides money between two stakeholders: 1 and 2. The spectator chooses between two distributions: Distribution $X = [x_1, x_2]$ and distribution $Y = [y_1, y_2]$. In Equation 3.3, the range in the stakeholder-1 dimension in this choice set is given by

$$\Delta_1(x_1, y_1) = \max[u_1(x_1), u_1(y_1)] - \min[u_1(x_1), u_1(y_1)]$$

Imagine adding to the choice set alternative $Z = [z_1, z_2]$ such that either $u_1(z_1) > \max[u_1(x_1), u_1(y_1)]$ or $u_1(z_1) < \min[u_1(x_1), u_1(y_1)]$, and such that $\min[u_2(x_2), u_2(y_2)] < u_2(z_2) < \max[u_2(x_2), u_2(y_2)]$. In this case, we say that distribution Z expands the range in the stakeholder-1 dimension and not in the stakeholder-2 dimension. Thus, $\Delta_1(x_1, y_1) < \Delta_1(x_1, y_1, z_1)$ and $\Delta_2(x_2, y_2) = \Delta_2(x_2, y_2, z_2)$. In the focusing model, the addition of distribution Z to the choice set makes the spectator assign more weight to stakeholder 1 than to stakeholder 2, because her focus is disproportionately drawn to the outcome for stakeholder 1. In the relative thinking model, the increased range for stakeholder 1 makes the spectator less sensitive to changes for stakeholder 1, and to thereby under-weight the outcome for stakeholder 1 in the decision-making.

3.2.2 Multi-dimensional distributive choices

The influence of unbalanced focus, and of relative thinking, may depend on factors such as the complexity of the situation or how familiar the decision-maker is with the choice environment. The focusing model aims to capture inattention or the cost of cognitive effort. Because focus is limited, it is allocated to choice dimensions that are salient and seem impactful. Thus, it is reasonable to expect a stronger focusing effect when the choice situation becomes more complex, i.e., when more dimensions are competing for the decision-maker's focus.

Consider now a distribution decision with N stakeholders. Define distributions $X = [x_1, ..., x_N]$, $Y = [y_1, ..., y_N]$ and $Z = [z_1, ..., z_N]$. As in the example with two stakeholders, assume that alternative Z expands the range for stakeholder 1, and not for the other stakeholders. As N increases, it is easy to imagine that the focusing-effect of adding Z to the choice set becomes stronger. Intuitively, keeping track of all dimensions of a choice problem becomes increasingly harder, and requires more focus as the number of choice dimensions increases. The focusing model could easily be extended by including the number of dimensions, K, in the weighting function $g(\Delta_k(C))$.

The relative thinking model captures a different mechanism. In this model, it is the size of a change *within* a choice dimension that is evaluated in relative terms. In other words, the decision weight in the focusing model is decided between choice dimensions, while the decision weights in the relative thinking model are decided within choice dimensions. Thus, if agents are prone to making both the "relative-thinking mistake" and the "unbalanced-focusing mistake", Bushong et al. (2021) argue that the number of choice dimensions could be a catalyst for the latter. Any (absence of a) treatment effect in the two-dimensional distributive setting could potentially hide two opposing effects. When the number of choice dimensions increases, we can expect the focusing effect to become more pronounced.

3.3 Experimental design and participants

In order to test for range-based context effects on distributive choices, I ran an incentivized online experiment. In this experiment, participants made distributive decisions as impartial spectators for stakeholders (hereinafter workers). The experiment included four different treatments, which all expanded the range of outcomes along one dimension of a distributive choice. The treatments were randomized between subjects.

3.3.1 Participants and sampling

Workers and spectators were recruited through Amazon's Mechanical Turk (MTurk), an online labor platform.[7] I recruited a total of 1397 spectators and 220 workers. Each spectator was matched to a group of workers with 5 percent probability. If matched, this spectator would dictate the distribution of money to the group of workers. All participants gave written consent to participate in the study and answered background questions about age, gender, education, and political views. The data reported here were collected in February 2021.

Table 3.1 presents sample characteristics for the spectator sample. Age varies from 18 to 83 years, with an average of 40.5 and a median of 37. The share of females in the sample is 50%. One-third of the sample identified as politically conservative, stating 4 or 5 on a five-point scale from "very liberal" to "very conservative". 76% of the sample have at least a two-year college degree. This is higher than the US national average, which for the adult population was 43% (Census Bureau, 2021).

3.3.2 Experimental design and treatment

This subsection gives a detailed description of the experimental design, and the different choice conditions and treatments. The spectators are told that a group of workers have carried out some unspecified tasks, and that each worker in the group did the

[7]Participants, both workers and spectators, were redirected from MTurk to a Qualtrics survey. To avoid any language issues, only participants with an IP address in the USA, and with at least a 98 percent approval rating in MTurk, were allowed to participate. Spectators who participated in the pilot study conducted in May 2020 were blocked from participating.

Table 3.1: Sample characteristics

	Sample	US population
Age median	40.45	38.3
Female share	0.50	0.51
College share	0.76	0.43
Conservative share	0.34	0.36
Observations	1,397	

Note: "Age median" is the median age in years. "Female share" is the proportion of females. "College share" is the share with at least a two year college degree. Age, gender, and education data for the US population are from the US Census Bureau (2021). "Conservative share" in Study 1 and Study 2 samples indicates 4 or 5 on a five-point scale for the question "Would you describe yourself as politically on the 'left' (e.g., a liberal) or on the 'right' (e.g., a conservative)?", where 1 is very liberal and 5 is very conservative. "Conservative share" in the US population is from Gallup, asking "How would you describe your political views – [very conservative, conservative, moderate, liberal or very liberal]?", including the answers "conservative" and "very conservative" (Gallup, 2022).

same job and performed equally well. The spectators have no further information about the workers. The spectators rank, from best to worst, alternative distributions of bonus payments to the group of workers. The spectators have to report a strict ranking of all available alternatives. Spectators in the control groups rank two alternatives, while spectators in a treatment group rank three alternatives. I use the probabilistically incentivized ranking scheme from Bushong et al. (2021). For the spectators who are drawn to have their decisions implemented, a random pair of alternatives is drawn, from which the one ranked as most preferred is implemented to decide bonus payments to the matched workers. Because all spectators provide a complete ranking, I am able to study the preference relation between alternatives 1 and 2, even in cases where alternative 3 is preferred over both.

In the first choice condition, the aim is to cleanly identify the effect of expanding the range of outcomes for one worker, while keeping the range for the other(s) unchanged. The experimental design needs to accommodate two key features: (i) The design must be robust to choice of dimensions. Kőszegi and Szeidl (2012) specify individuals as the relevant choice dimensions in distributive decisions. However, some spectators may view equality and efficiency as separate choice dimensions. Therefore, the experimental design should not depend on the exclusion of these or others as separate choice dimensions. The experiment is set up to test the focusing model and the relative thinking model with individual payoffs as the relevant choice dimensions. However, if some spectators consider other dimensions to be the relevant choice dimensions, this should not contaminate the results from the experiment. (ii) The experiment should be designed to maximize the likelihood of detecting an effect, i.e., it should move as many spectators as possible. To achieve this, spectators must be (close to) indifferent between alternative 1 and alternative 2 in the choice set.

Spectators in the control group are asked to rank, from most preferred to least preferred, the following two alternatives:

1) $7 to worker A and $5 to worker B

2) $5 to worker A and $7 to worker B

Because spectators have no information to distinguish between worker A and worker B, I expect all spectators to be indifferent between distributions 1 and 2. Spectators in the treatment group rank the same two distributions as those in the control group, in addition to a third one:

1) $7 to worker A and $5 to worker B

2) $5 to worker A and $7 to worker B

3) $7 to worker A and $0 to worker B

The addition of distribution 3 expands the range of possible outcomes in the worker-B dimension. Because the range is expanded downwards, I refer to this as the *negative* treatment. In the control group, the range in the worker-B dimension is

$$\Delta_B(C) = \max[u_B(7), u_B(5)] \\ - \min[u_B(7), u_B(5)] = u_B(7) - u_B(5)$$

In the treatment group, the range in the worker-B dimension is expanded to

$$\Delta_B(C) = \max[u_B(7), u_B(5), u_B(0)] \\ - \min[u_B(7), u_B(5), u_B(0)] = u_B(7) - u_B(0)$$

This experimental design satisfies points (i) and (ii) above. Expanding the range of equality or efficiency cannot affect the relative ranking of distributions 1 and 2, because distributions 1 and 2 have the same level of equality and efficiency. If equality and efficiency are viewed as separate choice dimensions by some spectators, they may allocate a higher or lower decision weight to equality or efficiency. Because the levels of equality and efficiency are the same in distribution 1 and distribution 2, equality and efficiency as separate choice dimensions cannot explain a treatment effect in this setting. I refer to this as the distributions being *symmetric*. This is an important feature of the treatment.

Because the only difference between the distributions is whether A or B get the better outcome, I assume that all spectators are indifferent between distribution 1 and distribution 2, thus satisfying point (ii). Distribution 3 expands the range for both equality and total efficiency. Given that all spectators are indifferent between distribution 1 and distribution 2, even a small focusing or relative-thinking effect should be enough to shift the choices of many spectators. The focusing model predicts that the inclusion of distribution 3 leads agents to focus more on, and thus over-weight, the worker-B dimension. Thus, spectators who see distributions 1, 2, and 3 are more likely to favor worker B over worker A, i.e., to rank distribution 2 over distribution 1, compared to the control group.

Hypothesis 1 — Focusing effect: Expanding the range for worker B makes spectators *more* likely to favor worker B.

The relative thinking model predicts that expanding the range of possible outcomes for worker B makes spectators less sensitive to changes along the worker-B dimension. That is, the difference between worker B getting $7 and $5 seems smaller in the treatment group than in the control group. The result, according to the relative thinking model, is that spectators who see alternatives 1, 2, and 3 are more likely to favor worker A over worker B, compared to those in the control group. This is the exact opposite of the prediction from the focusing model. In the simple setting described above, the relative-thinking model predicts the well-known asymmetric-dominance effect (Huber and Puto, 1983; Huber et al., 1982; Kamenica, 2008; Mellers and Cooke, 1994; Noguchi and Stewart, 2014; Pettibone and Wedell, 2000; Simonson, 1989; Simonson and Tversky, 1992), while the focusing model predicts the opposite.

Hypothesis 2 — Relative thinking effect: Expanding the range for worker B makes spectators *less* likely to favor worker B.

To further test the predictions of the models, I include a choice condition in which the treatment expands the range of outcomes upwards for worker A, rather than downwards for worker B. In this choice condition, the spectators in the control group rank the the same distributions as in the control group above:

1) $7 to worker A and $5 to worker B

2) $5 to worker A and $7 to worker B

In the treatment group, the spectators rank the following distributions:

1) \$7 to worker A and \$5 to worker B

2) \$5 to worker A and \$7 to worker B

3) \$12 to worker A and \$7 to worker B

Alternative 3 expands the range of outcomes for worker A. Because the range is expanded upwards, I refer to this as the *positive* treatment. In the positive treatment, the range of outcomes in the worker-A dimension is

$$\Delta_A(C) = \max[u_A(7), u_A(5), u_A(12)] \\ - \min[u_A(7), u_A(5), u_A(12)] = u_A(12) - u_A(5)$$

The main reason for including the positive treatment is that the predictions from the focusing model and the relative thinking model are reversed relative to the negative treatment. Distribution 3 in the positive treatment expands the range for worker A. The focusing model predicts that the presence of distribution 3 *increases* the share of spectators that favor worker A, by ranking distribution 1 as more preferred than distribution 2. The relative thinking model predicts that the presence of distribution 3 *decreases* the share of spectators who favor worker A. A second reason for including this extension, is that it allows me to study whether spectators react differently to increasing the range of outcomes with a dominating alternative, relative to a dominated alternative. In the positive treatment group, distribution 3 dominates distributions 1 and 2, whereas distribution 3 in the negative treatment is dominated by both of the other distributions in the choice set. The positive treatment design tests whether spectators react differently to dominated and dominating alternatives. I refer to the positive and negative treatments, as well as their common control group, as the *simple* choice conditions.

3.3.3 More complex distributive choices

To study the role of choice-set complexity, I add two choice conditions; one treatment group and one control group. In these conditions, the spectators make distributive decisions for six workers rather than only two. I refer to them as the *complex* choice conditions. The focusing model captures the allocation of limited focus between competing choice dimensions. Therefore, the focusing effect may become stronger when the number of choice dimensions increases. This could happen if the number of choice dimensions K enters into the weighting function $g(\Delta_k(C))$ in a way that interacts positively with the focusing effect.

In the complex choice conditions, spectators rank menus of distributions that involve six workers. The distribution decisions are otherwise similar to those in the negative treatment and the corresponding control condition.[8] Spectators in the complex control group face the following set of distributions:

1) $7 to worker A, $5 to worker B, and $7 to worker C, $5 to worker D, $7 to worker E, and $5 to worker F

2) $5 to worker A, $7 to worker B, and $5 to worker C, $7 to worker D, $5 to worker E, and $7 to worker F

The spectators in the complex treatment group face the same two alternatives as in the complex control group, in addition to one alternative that expands the range of possible outcomes for one of the workers, but not for the others:

1) $7 to worker A, $5 to worker B, and $7 to worker C, $5 to worker D, $7 to worker E, and $5 to worker F

[8] To maximize the likelihood of observing an effect, the range is expanded for the first worker to be mentioned, rather than the last one. This design choice implies that the treatment in the complex condition is not directly comparable to the negative treatment, where worker B's range is expanded.

2) $5 to worker A, $7 to worker B, and $5 to worker C, $7 to worker D, $5 to worker E, and $7 to worker F

3) $0 to worker A, $5 to worker B, and $7 to worker C, $5 to worker D, $7 to worker E, and $5 to worker F

The focusing model predicts that the introduction of $0 to worker A as a potential outcome will lead to a higher share of spectators ranking alternative 1 over alternative 2 relative to the control group without this distribution. The relative-thinking model predicts the opposite, for the same reason as described above. If the focusing effect becomes stronger when the number of choice dimensions increases, we should see an interaction effect between expanding the range and being in the complex choice condition.

Hypothesis 3 — Complexity: The focusing effect is relatively stronger when the choice setting is more complex.

3.3.4 Testing context effects with non-symmetric alternatives

As a final extension, I set up two distributive choice conditions with alternatives that are not equal as regards equality and efficiency. In the rest of the paper, I refer to these conditions as the *non-symmetric* conditions. Because distribution 1 and distribution 2 in the simple conditions are symmetric, it might also be salient to the decision-makers that the two distributions are equally good, and that they should be indifferent between them. In the non-symmetric conditions, spectators in the control group choose between two distributions, where one is more equal and the other is more efficient. In the treatment group, the choice set includes a distribution that expands the range of possible outcomes for worker A (and for efficiency), but not for worker B (and not for equality). Using the same procedures as described above, a control group of spectators face the following choice alternatives:

1) $8 to worker A and $4 to worker B

2) $6 to worker A and $5 to worker B

Because distribution 1 and distribution 2 are not symmetric, spectators are not necessarily indifferent between them. If more spectators than in the simple condition have a strict preference relation between 1 and 2 in this condition, it will take a stronger focusing effect or relative thinking effect to affect the ranking between 1 and 2. In the choice condition described here, both distribution 1 and distribution 2 favor worker A in terms of the relative bonus payment. However, distribution 1 favors worker A to a larger degree than distribution 2 does. To keep in line with the other choice conditions, I shall refer to ranking distribution 1 as more preferred than distribution 2 simply as favoring worker A. The spectators in the treatment group face the following alternatives:

1) $8 to worker A and $4 to worker B

2) $6 to worker A and $5 to worker B

3) $3 to worker A and $4 to worker B

Distribution 3 expands the range in the worker-A dimension, while leaving the range in the worker-B dimension unchanged. Thus, the focusing model predicts that spectators in the treatment group will be more likely to favor worker A, compared to those in the control group. The relative-thinking model predicts the opposite. Distribution 3 also expands the range of efficiency in the choice set, but not the range of equality.[9] If

[9]However, the average level of equality in the choice set decreases. If spectators' utility depends on a context-specific reference point for equality, this could in itself lead to the same prediction as in the relative-thinking model in this scenario. See, e.g., Bordalo et al. (2012, 2013); Cunningham (2013);

spectators see efficiency as a relevant choice dimension, the focusing model predicts that those who are given distribution 3 as an alternative should be more likely to favor worker A, compared to those who only see distributions 1 and 2. The relative thinking model predicts the opposite.

Table 3.2 provides an overview of the different choice conditions. The positive and negative treatments share the same control group. The main hypothesis in the paper

Table 3.2: Treatment design

	Negative	Positive	Complex	Non-symmetric
Choice setin control	1. (7, 5) 2. (5, 7)	1. (7, 5) 2. (5, 7)	1. (7, 5, 7, 5, 7, 5) 2. (5, 7, 5, 7, 5, 7)	1. (8, 4) 2. (6, 5)
Add	3. (7, **0**)	3. (**12**, 7)	3. (**0**, 5, 7, 5, 7, 5)	3. (**3**, 4)
Dim. expanded	B	A	A	A

Note: The table shows the alternative distributions (numbered) in the different control and treatment groups. The bottom row indicates for which worker distribution 3 expands the range, relative to the control group.

is the same in all the choice conditions: That there is a difference, between treatment and control, in the share of spectators who favor worker A over worker B, i.e., who rank distribution 1 as more preferred than distribution 2. The focusing model and the relative thinking model both predict a difference, but in opposite directions.

3.4 Empirical strategy and analysis

This section outlines the empirical strategy and regression specification.

I randomly assigned 1397 spectators to one of the seven choice conditions shown in Table 3.2.[10] The treatments are randomized between subjects, so that each spectator is

Kahneman and Tversky (1979); Kőszegi and Rabin (2006) for models of reference-dependent utility.

[10] The pre-analysis plan stated that subjects would be assigned equally to all seven treatment and

either in a control group or in a treatment group, and makes one distributive decision.

Here, I specify the OLS regression used to test the main hypothesis. The outcome variable is an indicator for favoring worker A, i.e., ranking alternative 1 as more preferred than alternative 2. There are two treatments, each indicated with a treatment variable: The first, "Expand A", is an indicator variable that is equal to 1 if the spectator is in a treatment group where the added alternative expands the range of outcomes in the worker-A dimension, and 0 otherwise. This variable is equal to 1 for spectators in the positive treatment group, the complex treatment group with six workers, and for the non-symmetric treatment group. The second treatment variable, "Expand B", is an indicator variable that is equal to 1 for spectators for whom the range is expanded in the worker-B dimension, and 0 otherwise. This variable is equal to 1 for spectators in the negative treatment group and 0 for all others.

In the main specification (Equation 3.4), the analysis is done separately for each of the four treatments compared to their respective control groups, both with and without control variables for background characteristics. When run separately, only the treatment variable relevant to the specific choice condition is included, as specified in the following equation:

$$C_i = \alpha + \beta T_i + \boldsymbol{\gamma} \boldsymbol{X}_i + \epsilon_i, \tag{3.4}$$

where T_i is either A or B, depending on the treatment group. C_i is an indicator for favoring worker A over worker B. α is a constant term. β is the estimated treatment effect of treatment T_i, which indicates that the range of outcomes is expanded in the worker-A(-B) dimension. Both models predict that "Expand A" and "Expand B" should

control groups with the same probability. Because of a mistake in the implementation, the randomization was instead done at decision-group level. This affects the number of participants in each treatment group. Specifically, treatment 1, treatment 2, and their common control group, have fewer spectators than intended, while the other groups have more.

have opposite directions, as shown by Hypotheses 1 and 2. γ is a vector of coefficients for background variables $\boldsymbol{X}_i$, which includes age, gender, education, and political views.

I also pool all treatment conditions in a regression that includes both of the treatment variables "Expand A" and "Expand B". This is done under the assumption that the individual workers' incomes are the relevant choice dimensions, and that expanding the range for one worker is a context-general treatment. In the pooled regression, I include indicator variables for the complex group and the non-symmetric group, to control for level differences. I also interact the treatment group indicators with the treatment variable relevant to the treatment group. This is done in order to control for differential treatment effects across the different choice conditions. According to Hypothesis 3, we should see a positive interaction between expanding the range for worker A and being in the complex group. The pooled regression is specified as follows:

$$C_i = \alpha + \beta_A T_{Ai} + \beta_B T_{Bi} + \kappa_D D_i + \kappa_N N_i + \theta_D (T_{Ai} \times D_i) + \theta_N (T_{Ai} \times N_i) + \boldsymbol{\gamma} \boldsymbol{X}_i + \epsilon_i \quad (3.5)$$

C_i is 1 if worker A is favored over worker B. α is a constant and the βs are the treatment coefficients for the T_i terms. In this regression, the baseline is the simple control group. β_B and β_A are the estimated treatment effects for the negative and positive treatments, respectively. D_i is an indicator for being in the complex condition, and N_i is an indicator for being in the non-symmetric condition. θ_D and θ_N are coefficients for the interaction between T_{Ai} and D_i, and T_{Ai} and N_i, respectively. The total estimated effect for spectators in the complex treatment, relative to the baseline, is $\kappa_D + \beta_A + \theta_D$. The total estimated effect for spectators in the non-symmetric treatment, relative to the baseline, is $\kappa_N + \beta_A + \theta_N$. The regression specified in Equation 3.5 is run with and without controlling for background characteristics.

The treatment variables are interacted with background characteristics in order to study heterogeneity. Because of power issues, this is considered as exploratory work, and the results should be interpreted with caution. Each treatment variable is interacted with each of the background indicators "Female", "College", and "Conservative" in turn. The regression is run separately for each subgroup. The regression for the negative treatment is specified by the following equation:

$$C_i = \alpha + \beta_B T_{Bi} + \delta_B (T_{Bi} \times Group) + \boldsymbol{\gamma} \boldsymbol{X}_i + \epsilon_i, \tag{3.6}$$

where *Group* is an indicator for either "Female", "College", or "Conservative". β_B is the treatment coefficient, and δ_B is the group-relevant coefficient for the interaction between the treatment and the background variable. ϵ_i is a mean-zero error term.

For the positive treatment, the complex treatment, and the non-symmetric treatment, the regression is specified by the following equation:

$$C_i = \alpha + \beta_A T_{Ai} + \delta_A (T_{Ai} \times Group) + \boldsymbol{\gamma} \boldsymbol{X}_i + \epsilon_i, \tag{3.7}$$

where β_A is the treatment coefficient, and δ_A is the group-relevant coefficient for the interaction between the treatment and the background variable. ϵ_i is a mean-zero error term.

3.5 Results

Here, I provide descriptive statistics and the main analysis of the spectator behavior.

Figure 3.1 shows, for each choice condition, the share of spectators who favor worker A over worker B, i.e., rank distribution 1 as more preferred than distribution 2, in the treatment group versus the control group. In all four panels, the control group is on the left, and the treated group is on the right. The bars show the 95% confidence intervals. In the simple control group, and in the complex control group, spectators are assumed to be indifferent between distribution 1 and distribution 2. Hence, we should expect that spectators randomize the rankings, so that we observe close to a 50/50 split between favoring workers A and B. In the simple control group, 57 percent of the spectators favored worker A over worker B, as shown in the upper panels in Figure 3.1. This is not statistically significantly different from the expected 50 percent $(p = 0, 076)$. In the complex control group, 61 percent of spectators ranked distribution 1 as more preferred than distribution 2, as seen in the bottom left panel. This share is different from 50 percent with $p = 0.0006$. The deviation from the 50/50 seen in the complex control group could be interpreted as something resembling a focusing effect. If the spectators are overwhelmed by the many choice dimensions, they might simply focus disproportionately on the first dimension, because it is more salient. This type of order-effect is not predicted by either of the models.

The lower right panel shows the plots for the non-symmetric group. In this choice condition, where distributions 1 and 2 have different levels of inequality, 76 percent of the spectators in the control group ranked distribution 2, the distribution with the highest level of equality, and lowest level of efficiency, as most preferred. This shows that the spectators are quite unwilling to implement more inequality in order to gain efficiency.

In the negative and positive treatment groups, 50 and 51 percent, respectively, of spectators favor worker A, compared to 57 percent in the control group. In the complex condition, 56 percent of the treated spectators favor worker A, compared to 61 percent in the control group. In the non-symmetric treatment group, 25 percent do the same, compared to 24 percent of those in the control group. In all choice conditions, the

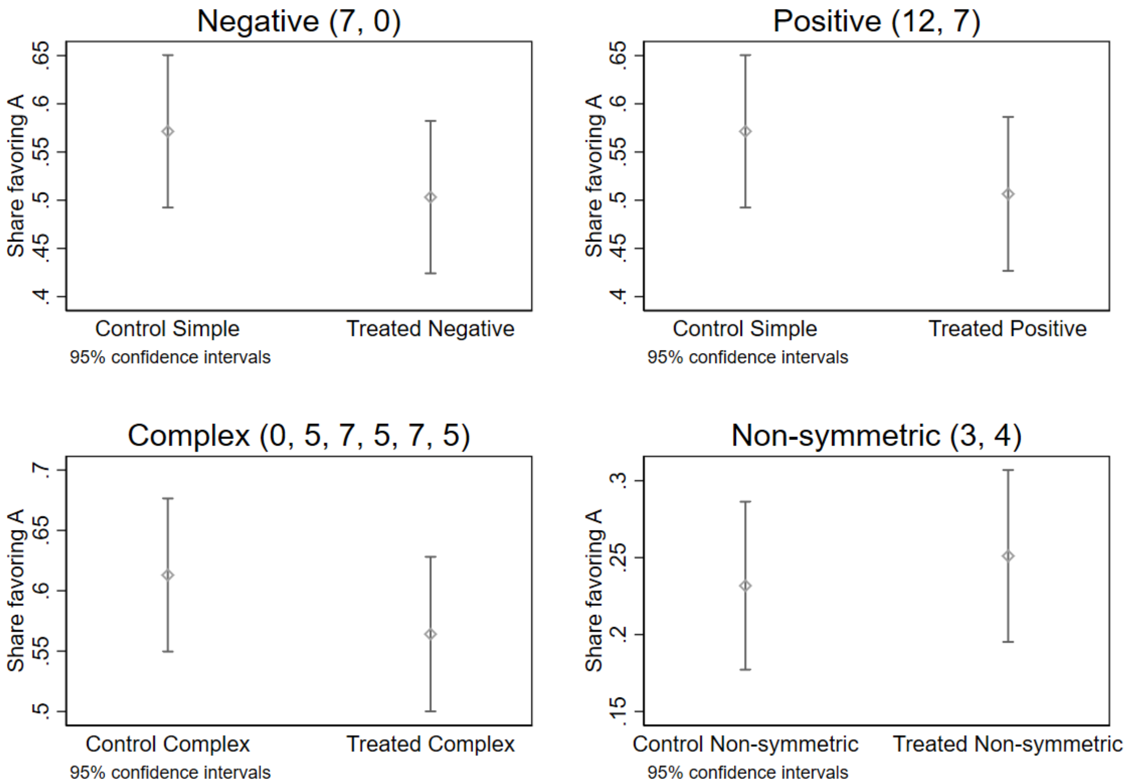


Figure 3.1: Each panel represents one choice condition pair, and shows the difference between the treatment and control group in the share of spectators who favor worker A over worker B, i.e., rank alternative 1 as more preferred than alternative 2.

differences between treatment and control are small.

In Table 3.3, I provide the regression analysis corresponding to Figure 3.1. I regress the outcome variable, an indicator for favoring worker A, on the treatment variable and the background characteristics. The regression is specified in Equation 3.4. Table 3.3 presents the coefficients for the treatment variables separately for all treatment conditions, compared to their respective control groups. The upper panel of the table includes background characteristics as control variables. The bottom panel is without controls. The first column shows the treatment effect for the negative treatment group, where distribution 3 is a dominated alternative. The second column shows the treatment effect from the positive treatment, where distribution 3 is a dominating alternative. Columns three and four are the complex group and the non-symmetric group, respectively. In the complex treatment and the non-symmetric treatment, alternative 3 is a dominated alternative.

Table 3.3: Effect on share that favor A, by treatment condition

	Negative	Positive	Complex	Non-symmetric
With controls				
Expand B	-0.0682			
	(0.0566)			
Expand A		-0.0649	-0.0489	0.0193
		(0.0569)	(0.0457)	(0.0396)
Without controls				
Expand B	-0.0620			
	(0.0567)			
Expand A		-0.0663	-0.0520	0.0370
		(0.0572)	(0.0461)	(0.0392)
Observations	311	308	464	468

Note: The table shows the treatment effect in all four treatment conditions separately. The outcome variable is an indicator for favoring worker A, i.e., ranking alternative 1 as more preferred than alternative 2. "Expand A" is a treatment variable indicating that the spectator is in a treatment group where distribution 3 expands the range of outcomes in the worker-A dimension. "Expand B" is a treatment variable indicating that the spectator is in a treatment group where distribution 3 expands the range of outcomes in the worker-B dimension. "Complex" indicates whether the spectator is in the group where the decision involves six workers rather than two. "Non-symmetric" is an indicator for being in the final decision group, where distributions 1 and 2 have different levels of equality. The top panel is with controls while the bottom panel is without controls.
Robust standard errors in parentheses. ($^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$).

The focusing model predicts that spectators focus more on, and are more prone to favoring, the worker for whom the range is expanded. In the current setting, this corresponds to a negative effect in the first column, and positive effects in the other three columns. The relative thinking model predicts the opposite: a positive effect in the first column, and a negative effect in the other three columns. As seen from Table 3.3, the effects are not consistently in the direction predicted by either of the models. None of the effects are significantly different from zero.

Table 3.4 presents regression results from the regression specified in equation 3.5, with all choice conditions pooled. The coefficients show the percentage-point increase in the estimated probability of favoring worker A when a given explanatory variable increases by one unit. The variables "Complex group" and "Non-symmetric" are indicators for the respective choice conditions. Their coefficients measure level differences compared to the simple control group as baseline. "Expand A × Complex" and "Expand A × Non-symmetric" are the corresponding interaction terms with the treatment.

Table 3.4: All treatments pooled

	(1) Favor A	(2) Favor A
Expand A	-0.0649 (0.0568)	-0.0712 (0.0569)
Expand B	-0.0682 (0.0566)	-0.0692 (0.0566)
Complex	0.0416 (0.0513)	0.0436 (0.0515)
Non-symmetric	-0.340^{***} (0.0486)	-0.348^{***} (0.0486)
Expand A × Complex	0.0160 (0.0730)	0.0141 (0.0730)
Expand A × Non-symmetric	0.0842 (0.0693)	0.0953 (0.0693)
Constant	0.571^{***} (0.0400)	0.622^{***} (0.0661)
Controls	No	Yes
Observations	1397	1397

Note: The outcome variable is an indicator for favoring worker A, i.e., ranking alternative 1 as more preferred than alternative 2. "Expand A" is a treatment variable indicating that the spectator is in a treatment group where distribution 3 expands the range of outcomes in the worker-A dimension. "Expand B" is a treatment variable indicating that the spectator is in a treatment group where distribution 3 expands the range of outcomes in the worker-B dimension. "Complex" indicates whether the spectator is in the group where the decision involves six workers rather than two. "Non-symmetric" is an indicator for being in the final decision group, where distributions 1 and 2 have different levels of equality. "Expand A × Complex" is the interaction between being in the complex treatment and having the range expanded in the worker A-dimension. "Expand A × Non-symmetric" is the interaction between being in the non-symmetric treatment and having the range expanded in the worker A-dimension. Column 1 is without control variables for background characteristics, and column 2 is with controls.
Robust standard errors in parentheses. ($^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$).

In Table 3.4, the constant term is the share of spectators in the simple control group that favor worker A. The coefficients for "Expand A" and "Expand B" show the treatment effects in the negative and positive treatments, respectively. Both models predict that these two coefficients should go in opposite directions, but they are both negative in Table 3.4. The estimated treatment coefficients are not statistically significantly different from zero. Spectators in the complex group favor A slightly more, but the coefficient is not significant. The coefficient for "Non-symmetric" is negative and significant. This shows that there is unwillingness to accept inequality in the non-symmetric choice setting. The coefficient for the interaction term "Expand A $\times$ Complex" is positive, as predicted by Hypothesis 3. However, the coefficient is close to zero, and far from statistically significant.

The regression analyses presented here provides no evidence in support of either the focusing model or the relative thinking model for distributive decisions. The focusing model predicts a positive coefficient for the variable "Expand A", and a negative coefficient for "Expand B". The relative-thinking model predicts the opposite. I find no statistically significant effects in either direction. Tables 3.3 and 3.4 both tell the same story: Expanding the range of possible outcomes for one of the workers has no effect on which of the workers gets the more favorable outcome. A null-finding can potentially hide opposing treatment effects that cancel out across different demographic groups. I therefore present the heterogeneity analysis specified in section 3.4.

To study heterogeneity, I interact the treatment variables "Expand A" and "Expand B" with each of the demographic group indicators "Female", "College" and "Conservative" in turn. The regressions specified in Equation 3.6 and Equation 3.7 are run separately for each of the treatment conditions. Table 3.5, shows results from the regression specified in Equation 3.6. Table 3.6, Table 3.7, and Table 3.8 show results from the regression specified in Equation 3.7. In all four tables, each column represents a separate regression, where the treatment variable is interacted with the background variable indicated by the column header. The tables also show the linear combinations of the coefficients for the treatment variables and their respective interaction terms. The linear combinations are reported as "Treatment effect on group".

Table 3.5: Heterogeneity — Negative treatment (7, 0)

	Female	College	Conservative
Expand B	-0.00123	-0.0587	-0.0574
	(0.0814)	(0.125)	(0.0695)
Expand B $\times$ Group	-0.119	-0.00414	-0.0147
	(0.114)	(0.141)	(0.121)
Group	0.0972	-0.00288	0.139^{*}
	(0.0799)	(0.0954)	(0.0826)
Treatment effect on group	-0.121	-0.0629	-0.0720
	(0.0794)	(0.0640)	(0.0987)
Observations	311	311	311

Note: The outcome variable is an indicator for favoring worker A, i.e., ranking alternative 1 as more preferred than alternative 2. "Expand B" is a treatment variable indicating that the spectator is in the treatment group where the third alternative expands the range of outcomes in the worker-B dimension. "Expand B $\times$ Group" is the interaction between "Expand B" and "Group". "Group" is an indicator for being in the group defined by the column-header background variable. "Treatment effect on group" is the linear combination "Expand B" + "Expand B $\times$ Group".
Robust standard errors in parentheses. ($^{*}: p < 0.1$,$^{**}: p < 0.05$,$^{***}: p < 0.01$).

In Table 3.5, the focusing model predicts negative coefficients for "Expand B" and for "Treatment effect on group" in all three columns. The relative thinking model predicts the opposite. The estimated coefficients are in the direction predicted by the focusing model, but are not statistically significant.

In Table 3.6, the focusing model predicts positive coefficients for "Expand A" and for "Treatment effect on group" in all three columns. The relative thinking model predicts the opposite. The coefficients are in the direction predicted by the relative thinking model, but are not statistically significant.

In Table 3.7, the focusing model predicts positive coefficients for "Expand A" and

Table 3.6: Heterogeneity — Positive treatment (12, 7)

	Female	College	Conservative
Expand A	−0.0188	−0.0302	−0.0518
	(0.0800)	(0.119)	(0.0708)
Expand A × Group	−0.0984	−0.0464	−0.0436
	−0.0984	(0.135)	(0.119)
Group	0.100	−0.00724	0.142*
	(0.0800)	(0.0953)	(0.0824)
Treatment effect on group	−0.117	−0.0765	−0.0954
	(0.0817)	(0.0651)	(0.0965)
Observations	308	308	308

Note: The outcome variable is an indicator for favoring worker A, i.e., ranking alternative 1 as more preferred than alternative 2. "Expand A" is a treatment variable indicating that the spectator is in the treatment condition where the third alternative expands the range of outcomes in the worker-A dimension. "Expand A × Group" is the interaction between "Expand A" and "Group". "Group" is an indicator for being in the group defined by the column-header background variable. "Treatment effect on group" is the linear combination "Expand A" + "Expand A × Group".
Robust standard errors in parentheses. ($^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$).

for "Treatment effect on group" in all three columns. The relative thinking model predicts the opposite. All coefficients, except for the treatment coefficient for political conservatives, are in the direction predicted by the relative thinking model, but are not statistically significant. The treatment coefficient for political conservatives is close to zero, and not statistically significant.

In Table 3.8, the focusing model predicts positive coefficients for "Expand A" and for "Treatment effect on group" in all three columns. The relative thinking model predicts the opposite. The estimated treatment effect is positive for males, those with at least a two-year college degree, and for political conservatives. The estimated treatment effect is negative for females, those with less than two years of college education, and for political liberals and moderates. There is a positive interaction effect between the treatment and being politically conservative, and a positive estimated treatment effect on the group. Because the heterogeneity analyses are under-powered, the results must be interpreted with great caution. However, the heterogeneity analyses show, quite

Table 3.7: Heterogeneity — Complex treatment

	Female	College	Conservative
Expand A	-0.0346	-0.0628	-0.0811
	(0.0682)	(0.0889)	(0.0572)
Expand A × Group	-0.0325	0.0150	0.0837
	(0.0928)	(0.104)	(0.0962)
Group	0.0275	0.0614	-0.00601
	(0.0657)	(0.0720)	(0.0693)
Treatment effect on group	-0.0671	-0.0478	0.00264
	(0.0628)	(0.0542)	(0.0775)
Observations	464	464	464

Note: The outcome variable is an indicator for favoring worker A, i.e., ranking alternative 1 as more preferred than alternative 2. "Expand A" is a treatment variable indicating that the spectator is in the treatment condition where the third alternative expands the range of outcomes in the worker-A dimension. "Expand A × Group" is the interaction between "Expand A" and "Group". "Group" is an indicator for being in the group defined by the column-header background variable. "Treatment effect on group" is the linear combination "Expand A" + "Expand A × Group".
Robust standard errors in parentheses. ($^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$).

robustly across subgroups in society, that there are no significant range-based context effects in the settings presented here.

Table 3.8: Heterogeneity — Non-symmetric treatment

	Female	College	Conservative
Expand A	0.0802	−0.0570	−0.0152
	(0.0573)	(0.0715)	(0.0469)
Expand A × Group	−0.0937	0.121	0.154*
	(0.0773)	(0.0844)	(0.0841)
Group	−0.0586	0.0219	−0.0320
	(0.0547)	(0.0624)	(0.0585)
Treatment effect on group	−0.0135	0.0642	0.138**
	(0.0520)	(0.0458)	(0.0700)
Observations	468	468	468

Note: The outcome variable is an indicator for favoring worker A, i.e., ranking alternative 1 as more preferred than alternative 2. "Expand A" is a treatment variable indicating that the spectator is in the treatment condition where the third alternative expands the range of outcomes in the worker-A dimension. "Expand A × Group" is the interaction between "Expand A" and "Group". "Group" is an indicator for being in the group defined by the column-header background variable. "Treatment effect on group" is the linear combination "Expand A" + "Expand A × Group".
Robust standard errors in parentheses. ($^{*}: p < 0.1, ^{**}: p < 0.05, ^{***}: p < 0.01$).

3.6 Discussion and conclusion

This paper finds no evidence for range-based context effects in fairness decisions, which is contrary to predictions from the focusing model and the relative thinking model, and to empirical findings in consumer research. My findings suggest that distributive preferences may be more stable than preferences for consumption goods.

One explanation for why my findings contradict those in the consumer literature, is the scale of different choice dimensions. The choice dimensions we encounter in consumer behavior, for example price and some measure of quality, or different quality dimensions, are often on different scales. This can make different choice dimensions hard to compare. For example, knowing how much better a TV with OLED technology is compared to one with QLED can be hard to determine. Thus, it may be hard to determine one's willingness to pay for the upgrade. This is true for many consumption decisions. In consumption decisions, the alternatives in the choice-set may provide some information about the relative values of the choice dimensions.

The choice dimensions in the experiment presented in this paper are easily comparable. When choosing between paying money to worker A and to worker B, the two dimensions are measured on the same scale, and are thus easy to compare. This may contribute to explaining the null-findings in the present study. If the worker dimensions appear indistinguishable, spectators might instead adopt other measures as the relevant choice dimensions, such as equality and total efficiency. In choice settings where the involved workers are completely anonymous, each individual's payoff might not be considered as a separate choice dimension. The experimental design presented here cannot identify what the spectators consider to be the relevant choice dimensions in fairness decisions. In both the focusing model and the relative thinking model, the relevant choice dimensions are defined exogenously. Identifying what the decision makers consider as the relevant choice dimensions is an interesting avenue for future research. It would help us to better understand what affects distributive preferences, and guide the application of choice models for consumer behavior to understand distributive behavior.

In the non-symmetric treatment, the range of efficiency in the choice set is expanded, while the inequality dimension is kept constant. If equality and efficiency are the relevant dimensions, the focusing model predicts that spectators in the non-symmetric treatment group should focus more on efficiency and less on equality. This means that we should have observed more spectators favoring worker A in this treatment group compared to the control group. While the point estimate does go in the direction predicted by the focusing model, the effect is small and far from statistically significant. The relative thinking model predicts the opposite, which is not supported by the data either.

As pointed out by Bushong et al. (2021), the effect captured by the focusing model is likely to become more relevant as the number of choice dimensions increases, because the model is driven by limited attention or computational capacity. In low-dimensional problems, cognitive capacity may not be a relevant restriction to appropriately weight the different choice dimensions. If the number of choice dimensions were an important factor in terms of explaining how focus is allocated between choice dimensions, we would expect to see a stronger focusing effect in the complex treatment than in the other treatments. The coefficients are not directly comparable between treatments, but there does not seem to be a strong focusing effect in any of the treatments.

Taken together, my results suggest that people are either not as susceptible to choice-set effects when making fairness decisions as when making consumption decisions or the choice-set effects only occur when choice dimensions are both clearly defined and not easily comparable. More research is needed in order to conclude. A lot still remains to be studied in order to fully understand how fairness decisions are made, and what decision-makers focus on in distributive decisions. Determining what the relevant choice dimensions are in different choice settings appears to be an important avenue for future research.

3.A Spectator instructions

Welcome!

Please note that your participation will be registered on the following Amazon Mechanical Turk worker ID:

[Automatically generated worker ID]

The worker ID was retrieved automatically when you clicked on the link that brought you here. This step is necessary for assigning payments to the right account and to ensure that you only participate in this study once.

[New page]

Introduction

Welcome to this research project! We very much appreciate your participation.

This is a study about the economics of decision-making. Several research institutions have provided funds for this research.

Payment

Your payment for taking the HIT will be sent to you shortly after the completion of this HIT.

Procedures

The study consists of two parts and you will be given instructions on your screen before every single part of the survey. Please always make sure to read the instructions carefully before you continue.

Participation

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate entirely without jeopardy to future participation in other studies conducted by us.

Confidentiality

All data obtained from you will be kept confidential and will only be reported in aggregate format (by reporting only combined results and never reporting individual ones). All questionnaires will be concealed, and no one other than the primary investigator will have access to them.

Verification

At the end of this survey, you will be given a completion code. You will need to copy this code to the survey code field on the AMT web page that directed you here at the beginning.

Questions about the Research

If you have questions regarding this study, you may contact

thechoicelab@nhh.no

[New page]

I have read and understood the above consent form and desire to participate in this study.

- Yes
- No

You will now be asked to rank a list of alternative distributions of money between two anonymous workers, A and B, who have completed the same task and performed equally well. In addition to a participation fee, the workers are told that they may receive a bonus. They know that this bonus is decided by someone else, but they have no further information.

With a certain probability your decision will decide the bonus for these two workers. This is done by randomly drawing two alternative distributions from the list that you have ranked and implement the one which you rank as most preferred between the two alternatives.

Keep in mind that your decision affects real people, so it is important that you take this seriously.

[New page]

Spectators are randomly allocated into one out of seven decision groups. The first three have the following instructions:

You will now be asked to rank a list of alternative distributions of money between two anonymous workers, A and B, who have completed the same task and performed equally well. In addition to a participation fee, the workers are told that they may receive a bonus. They know that this bonus is decided by someone else, but they have no further information.

With a certain probability your decision will decide the bonus for these two workers. This is done by randomly drawing two alternative distributions from the list that you have ranked and implement the one which you rank as most preferred between the two alternatives.

Keep in mind that your decision affects real people, so it is important that you take this seriously.

1) Please rank the following distributions of money to worker A and worker B from 1 (most preferred) to 2 (least preferred)

- $7 to worker A and $5 to worker B
- $5 to worker A and $7 to worker B

2) Please rank the following distributions of money to worker A and worker B from 1 (most preferred) to 3 (least preferred)

- $7 to worker A and $5 to worker B
- $5 to worker A and $7 to worker B

- $7 to worker A and $0 to worker B

3) Please rank the following distributions of money to worker A and worker B from 1 (most preferred) to 3 (least preferred)

- $7 to worker A and $5 to worker B
- $5 to worker A and $7 to worker B
- $12 to worker A and $7 to worker B

Spectators in the complex decision groups see the following instructions:

You will now be asked to rank a list of alternative distributions of money between six anonymous workers, A, B, C, D, E, and F, who have completed the same task and performed equally well. In addition to a participation fee, the workers are told that they may receive a bonus. They know that this bonus is decided by someone else, but they have no further information.

With a certain probability your decision will decide the bonus for these six workers. This is done by randomly drawing two alternative distributions from the list that you have ranked and implement the one which you rank as most preferred between the two alternatives.

Keep in mind that your decision affects real people, so it is important that you take this seriously.

4) Please rank the following distributions of money to worker A, B, C, D, E, and F from 1 (most preferred) to 2 (least preferred)

- $7 for worker A, $5 for worker B, $7 for worker C, $5 for worker D, $7 for worker E, and $5 for worker F
- $5 for worker A, $7 for worker B, $5 for worker C, $7 for worker D, $5 for worker E, and $7 for worker F

5) Please rank the following distributions of money to worker A, B, C, D, E, and F from 1 (most preferred) to 3 (least preferred)

- $7 for worker A, $5 for worker B, $7 for worker C, $5 for worker D, $7 for worker E, and $5 for worker F
- $5 for worker A, $7 for worker B, $5 for worker C, $7 for worker D, $5 for worker E, and $7 for worker F
- $0 for worker A, $5 for worker B, $7 for worker C, $5 for worker D, $7 for worker E, and $5 for worker F

Spectators in the strict group see the following instructions:

You will now be asked to rank a list of alternative distributions of money between two anonymous workers, A and B, who have completed the same task and performed equally well. In addition to a participation fee, the workers are told that they may receive a bonus. They know that this bonus is decided by someone else, but they have no further information.

With a certain probability your decision will decide the bonus for these two workers. This is done by randomly drawing two alternative distributions from the list that you have ranked and implement the one which you rank as most preferred between the two alternatives.

Keep in mind that your decision affects real people, so it is important that you take this seriously.

6) Please rank the following distributions of money to worker A and worker B from 1 (most preferred) to 2 (least preferred).

- $6 for worker A and $5 for worker B
- $8 for worker A and $4 for worker B

7) Please rank the following distributions of money to worker A and worker B from 1 (most preferred) to 3 (least preferred).

- $6 for worker A and $5 for worker B
- $8 for worker A and $4 for worker B
- $3 for worker A and $4 for worker B

Finally, the spectators answer some background questions:

You have completed the first part of the survey. We would now like to ask you five more questions before we conclude this survey.

What is your gender?

- Male
- Female

How old are you?

[Open textbox for number]

What is the highest level of education you have completed?

- Less than High School
- High School / GED
- Some College
- 2-year College Degree
- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

Would you describe yourself as politically on the "left" (i.e. a liberal) or on the "right" (i.e. a conservative)?

1 - Very liberal	2	Neutral	4	5 - Very conservative

Finally, if you have any comments or suggestions related to this study please write them down in the field below. Your feedback is very important to improve our research.

[Open textbox]

[New page]

We thank you for your time spent taking this survey.

Your response has been recorded.

References

Aguiar, Fernando, Alice Becker, and Luis Miller, "Whose impartiality? An experimental study of veiled stakeholders, involved spectators and detached observers," *Economics & Philosophy*, 2013, *29* (2), 155–174.

Akerlof, George A and Rachel E Kranton, "Economics and identity," *The quarterly journal of economics*, 2000, *115* (3), 715–753.

_ and _ , "Identity and the Economics of Organizations," *Journal of Economic perspectives*, 2005, *19* (1), 9–32.

Alesina, Alberto and George-Marios Angeletos, "Fairness and redistribution," *American economic review*, 2005, *95* (4), 960–980.

_ and Paola Giuliano, "Preferences for redistribution," in "Handbook of social economics," Vol. 1, Elsevier, 2011, pp. 93–131.

_ , Stefanie Stantcheva, and Edoardo Teso, "Intergenerational mobility and preferences for redistribution," *American Economic Review*, 2018, *108* (2), 521–54.

Almås, Ingvild, Alexander W Cappelen, and Bertil Tungodden, "Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians?," *Journal of Political Economy*, 2020, *128* (5), 1753–1788.

_ , _ , Erik Ø Sørensen, and Bertil Tungodden, "Fairness and the development of inequality acceptance," *Science*, 2010, *328* (5982), 1176–1178.

__ , __ , **Jo Thori Lind, Erik Ø Sørensen, and Bertil Tungodden**, "Measuring unfair (in) equality," *Journal of Public Economics*, 2011, *95* (7-8), 488–499.

__ , __ , **Kjell G Salvanes, Erik Ø Sørensen, and Bertil Tungodden**, "Fairness and family background," *Politics, Philosophy & Economics*, 2017, *16* (2), 117–131.

Alvarez-Cuadrado, Francisco, Goncalo Monteiro, and Stephen J Turnovsky, "Habit formation, catching up with the Joneses, and economic growth," *Journal of economic growth*, 2004, *9* (1), 47–80.

Anderson, Norman H, "Foundations of information integration theory," 1981.

Andreoni, James and B Douglas Bernheim, "Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects," *Econometrica*, 2009, *77* (5), 1607–1636.

__ **and John Miller**, "Giving according to GARP: An experimental test of the consistency of preferences for altruism," *Econometrica*, 2002, *70* (2), 737–753.

Atkinson, Anthony B, Thomas Piketty, and Emmanuel Saez, "Top incomes in the long run of history," *Journal of economic literature*, 2011, *49* (1), 3–71.

Austen-Smith, David and Roland G Fryer Jr, "An economic analysis of "acting white"," *The Quarterly Journal of Economics*, 2005, *120* (2), 551–583.

Azar, Ofer H, "Relative thinking theory," *The Journal of Socio-Economics*, 2007, *36* (1), 1–14.

__ , "Do consumers make too much effort to save on cheap items and too little to save on expensive items? Experimental results and implications for business strategy," *American Behavioral Scientist*, 2011, *55* (8), 1077–1098.

Becker, Gary S and Kevin M Murphy, "A theory of rational addiction," *Journal of political Economy*, 1988, *96* (4), 675–700.

Ben-Ner, Avner, Brian P McCall, Massoud Stephane, and Hua Wang, "Identity and in-group/out-group differentiation in work and giving behaviors: Experimental evidence," *Journal of Economic Behavior & Organization*, 2009, *72* (1), 153–170.

Bénabou, Roland and Jean Tirole, "Incentives and prosocial behavior," *American economic review*, 2006, *96* (5), 1652–1678.

Benabou, Roland and Jean Tirole, "Laws and norms," Technical Report, National Bureau of Economic Research 2011.

Benjamin, Dan, Aaron Bodoh-Creed, and Matthew Rabin, "Base-rate neglect: Foundations and implications," Technical Report, working paper 2019.

Benjamin, Daniel J, James J Choi, and A Joshua Strickland, "Social identity and preferences," *American Economic Review*, 2010, *100* (4), 1913–28.

Bernheim, B Douglas, "A theory of conformity," *Journal of political Economy*, 1994, *102* (5), 841–877.

Bicchieri, Cristina and Erte Xiao, "Do the right thing: but only if others do so," *Journal of Behavioral Decision Making*, 2009, *22* (2), 191–208.

_ , Eugen Dimant, Simon Gächter, and Daniele Nosenzo, "Social proximity and the erosion of norm compliance," *Available at SSRN*, 2021.

Bodoh-Creed, Aaron L, "Mood, memory, and the evaluation of asset prices," *Review of Finance*, 2020, *24* (1), 227–262.

Bolton, Gary E and Axel Ockenfels, "ERC: A theory of equity, reciprocity, and competition," *American Economic Review*, 2000, *90* (1), 166–193.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience theory of choice under risk," *The Quarterly Journal of Economics*, 2012, *127* (3), 1243–1285.

__ , __ , and __ , "Salience and consumer choice," *Journal of Political Economy*, 2013, *121* (5), 803–843.

__ , __ , and __ , "Memory, attention, and choice," *The Quarterly Journal of Economics*, 2020, *135* (3), 1399–1442.

Bott, Kristina M, Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden, "You've got mail: A randomized field experiment on tax evasion," *Management Science*, 2020, *66* (7), 2801–2819.

Bureau, United States Census, "Educational attainment," Web page 2021. Accessed 20th November 2022.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott, "Misperceived social norms: Women working outside the home in saudi arabia," *American economic review*, 2020, *110* (10), 2997–3029.

__ and Robert Jensen, "Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure," *Annual Review of Economics*, 2017, *9*, 131–153.

__ , Georgy Egorov, and Stefano Fiorin, "From extreme to mainstream: The erosion of social norms," *American economic review*, 2020, *110* (11), 3522–48.

Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein, "A model of relative thinking," *The Review of Economic Studies*, 2021, *88* (1), 162–191.

Camerer, Colin F, *Behavioral game theory: Experiments in strategic interaction*, Princeton university press, 2011.

Cappelen, Alexander W, Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden, "The pluralism of fairness ideals: An experimental approach," *American Economic Review*, 2007, *97* (3), 818–827.

__ , **James Konow, Erik Ø Sørensen, and Bertil Tungodden**, "Just luck: An experimental study of risk-taking and fairness," *American Economic Review*, 2013, *103* (4), 1398–1413.

__ , **Johanna Mollerstrom, Bjørn-Atle Reme, and Bertil Tungodden**, "A Meritocratic Origin of Egalitarian Behavior," 2019.

__ , __ , __ , **and** __ , "A meritocratic origin of egalitarian behaviour," *The Economic Journal*, 2022, *132* (646), 2101–2117.

__ , **Karl O Moene, Siv-Elisabeth Skjelbred, and Bertil Tungodden**, "The merit primacy effect," *The Economic Journal*, November 2022.

__ , **Sebastian Fest, Erik Ø Sørensen, and Bertil Tungodden**, "Choice and personal responsibility: What is a morally relevant choice?," *The Review of Economics and Statistics*, 2022, pp. 1–35.

Carpio, Lucia Del, "Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru Job Market Paper," 2013.

Carroll, Christopher D, Jody Overland, and David N Weil, "Saving and growth with habit formation," *American Economic Review*, 2000, *90* (3), 341–355.

Centola, Damon, Joshua Becker, Devon Brackbill, and Andrea Baronchelli, "Experimental evidence for tipping points in social convention," *Science*, 2018, *360* (6393), 1116–1119.

Chamberlin, Edward H, "An experimental imperfect market," *Journal of political economy*, 1948, *56* (2), 95–108.

Charness, Gary and Matthew Rabin, "Understanding social preferences with simple tests," *The Quarterly Journal of Economics*, 2002, *117* (3), 817–869.

— **and Yan Chen**, "Social identity, group behavior, and teams," *Annual Review of Economics*, 2020, pp. 691–713.

— **, Uri Gneezy, and Michael A Kuhn**, "Experimental methods: Between-subject and within-subject design," *Journal of Economic Behavior & Organization*, 2012, *81* (1), 1–8.

Chen, Yan and Sherry Xin Li, "Group identity and social preferences," *American Economic Review*, 2009, *99* (1), 431–57.

Cherry, Todd L, Peter Frykblom, and Jason F Shogren, "Hardnose the dictator," *American Economic Review*, 2002, *92* (4), 1218–1221.

Cialdini, Robert B, Raymond R Reno, and Carl A Kallgren, "A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places.," *Journal of personality and social psychology*, 1990, *58* (6), 1015.

Constantinides, George M, "Habit formation: A resolution of the equity premium puzzle," *Journal of political Economy*, 1990, *98* (3), 519–543.

Croson, Rachel and James Konow, "Social preferences and moral biases," *Journal of Economic Behavior & Organization*, 2009, *69* (3), 201–212.

Cunningham, Tom, "Comparisons and choice," *Unpublished manuscript*, 2013.

DellaVigna, Stefano, Attila Lindner, Balázs Reizer, and Johannes F Schmieder, "Reference-dependent job search: Evidence from Hungary," *The Quarterly Journal of Economics*, 2017, *132* (4), 1969–2018.

— **, John A List, and Ulrike Malmendier**, "Testing for altruism and social pressure in charitable giving," *The quarterly journal of economics*, 2012, *127* (1), 1–56.

— **,** — **,** — **, and Gautam Rao**, "Voting to tell others," *The Review of Economic Studies*, 2016, *84* (1), 143–181.

Dertwinkel-Kalt, Markus, Holger Gerhardt, Gerhard Riener, Frederik Schwerter, and Louis Strang, "Concentration bias in intertemporal choice," *The Review of Economic Studies*, 2022, *89* (3), 1314–1334.

Desvousges, William H, F Reed Johnson, Richard W Dunford, Sara P Hudson, K Nicole Wilson, and Kevin J Boyle, "Measuring natural resource damages with contingent valuation: tests of validity and reliability," in "Contributions to Economic Analysis," Vol. 220, Elsevier, 1993, pp. 91–164.

Durante, Ruben, Louis Putterman, and Joël Van der Weele, "Preferences for redistribution and perception of fairness: An experimental study," *Journal of the European Economic Association*, 2014, *12* (4), 1059–1086.

Dynan, Karen E, "Habit formation in consumer preferences: Evidence from panel data," *American Economic Review*, 2000, *90* (3), 391–406.

Ekström, Mathias, "The (un) compromise effect: How suggested alternatives can promote active choice," *Journal of Behavioral and Experimental Economics*, 2021, *90*, 101639.

Elster, Jon, "Social norms and economic theory," *Journal of economic perspectives*, 1989, *3* (4), 99–117.

Engelmann, Dirk and Martin Strobel, "The false consensus effect disappears if representative information and monetary incentives are given," *Experimental Economics*, 2000, *3* (3), 241–260.

_ **and** _ , "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments," *American economic review*, 2004, *94* (4), 857–869.

Enke, Benjamin, Frederik Schwerter, and Florian Zimmermann, "Associative memory and belief formation," Technical Report, National Bureau of Economic Research 2020.

Epstein, Ziv, Alexander Peysakhovich, and David G Rand, "The good, the bad, and the unflinchingly selfish: Cooperative decision-making can be predicted with high accuracy when using only three behavioral types," in "Proceedings of the 2016 ACM Conference on Economics and Computation" 2016, pp. 547–559.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global evidence on economic preferences," *The Quarterly Journal of Economics*, 2018, *133* (4), 1645–1692.

Fang, Hanming and Glenn C Loury, "" Dysfunctional Identities" Can Be Rational," *American Economic Review*, 2005, *95* (2), 104–111.

Faro, David and Yuval Rottenstreich, "Affect, empathy, and regressive mispredictions of others' preferences under risk," *Management Science*, 2006, *52* (4), 529–541.

Fehr, Ernst and Klaus M Schmidt, "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 1999, *114* (3), 817–868.

_ **and Simon Gächter**, "Fairness and retaliation: The economics of reciprocity," *Journal of economic perspectives*, 2000, *14* (3), 159–181.

Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler, "Testing enforcement strategies in the field: Threat, moral appeal and social information," *Journal of the European Economic Association*, 2013, *11* (3), 634–660.

Frederick, Shane, Leonard Lee, and Ernest Baskin, "The limits of attraction," *Journal of Marketing Research*, 2014, *51* (4), 487–507.

Funk, Patricia, "Social incentives and voter turnout: evidence from the Swiss mail ballot system," *Journal of the European economic association*, 2010, *8* (5), 1077–1103.

Gabaix, Xavier, "A sparsity-based model of bounded rationality," *The Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

Gallup, "U.S. Political Ideology Steady; Conservatives, Moderates Tie," Web page January 2022.

Gerber, Alan S, Donald P Green, and Christopher W Larimer, "Social pressure and voter turnout: Evidence from a large-scale field experiment," *American political Science review*, 2008, *102* (1), 33–48.

Giuliano, Paola and Antonio Spilimbergo, "Growing up in bad times: Macroeconomic Volatility and the Formation of Beliefs," Technical Report 2008.

Gunnthorsdottir, Anna, Daniel Houser, and Kevin McCabe, "Disposition, history and contributions in public goods experiments," *Journal of Economic Behavior & Organization*, 2007, *62* (2), 304–315.

Heath, Timothy B and Subimal Chatterjee, "Asymmetric decoy effects on lower-quality versus higher-quality brands: Meta-analytic and experimental evidence," *Journal of Consumer Research*, 1995, *22* (3), 268–284.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger et al., ""Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies," *Behavioral and brain sciences*, 2005, *28* (6), 795–815.

Huber, Joel and Christopher Puto, "Market boundaries and product choice: Illustrating attraction and substitution effects," *Journal of consumer research*, 1983, *10* (1), 31–44.

_, John W Payne, and Christopher Puto, "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," *Journal of consumer research*, 1982, *9* (1), 90–98.

Kahneman, Daniel, "Maps of bounded rationality: Psychology for behavioral economics," *American Economic Review*, 2003, *93* (5), 1449–1475.

__ **and Amos Tversky**, "On the psychology of prediction.," *Psychological review*, 1973, *80* (4), 237.

__ **and __** , "Prospect theory: An analysis of decision under risk," *Econometrica*, 1979, *47* (2), 363–391.

__ **and Dale T Miller**, "Norm theory: Comparing reality to its alternatives.," *Psychological review*, 1986, *93* (2), 136.

Kamenica, Emir, "Contextual inference in markets: On the informational content of product lines," *American Economic Review*, 2008, *98* (5), 2127–49.

Keizer, Kees, Siegwart Lindenberg, and Linda Steg, "The spreading of disorder," *science*, 2008, *322* (5908), 1681–1685.

Klor, Esteban F and Moses Shayo, "Social identity and preferences over redistribution," *Journal of Public Economics*, 2010, *94* (3-4), 269–278.

Konow, James, "Fair shares: Accountability and cognitive dissonance in allocation decisions," *American economic review*, 2000, *90* (4), 1072–1091.

__ , "Is fairness in the eye of the beholder? An impartial spectator analysis of justice," *Social Choice and Welfare*, 2009, *33* (1), 101–127.

__ **, Tatsuyoshi Saijo, and Kenju Akai**, "Equity versus equality: Spectators, stakeholders and groups," *Journal of Economic Psychology*, 2020, *77*, 102171.

__ **, __ , __ et al.**, "Morals and mores: Experimental evidence on equity and equality," Technical Report, Working paper, Loyola Marymount University 2009.

Kőszegi, Botond and Adam Szeidl, "A model of focusing in economic choice," *The Quarterly Journal of Economics*, 2012, *128* (1), 53–104.

__ **and Matthew Rabin**, "A model of reference-dependent preferences," *The Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

Koszegi, Botond, George Loewenstein, and Takeshi Murooka, "Fragile self-esteem," *Available at SSRN 3412006*, 2019.

Kranton, Rachel E, "Identity economics 2016: Where do social distinctions and norms come from?," *American Economic Review*, 2016, *106* (5), 405–09.

Krueger, Joachim and Russell W Clement, "The truly false consensus effect: an ineradicable and egocentric bias in social perception.," *Journal of personality and social psychology*, 1994, *67* (4), 596.

Krupka, Erin and Roberto A Weber, "The focusing and informational effects of norms on pro-social behavior," *Journal of Economic psychology*, 2009, *30* (3), 307–320.

Lane, Tom, "Intrinsic preferences for unhappy news," *Journal of Economic Behavior & Organization*, 2022, *202*, 119–130.

Lindbeck, Assar, Sten Nyberg, and Jörgen W Weibull, "Social norms and economic incentives in the welfare state," *The Quarterly Journal of Economics*, 1999, *114* (1), 1–35.

List, John A, "On the interpretation of giving in dictator games," *Journal of Political economy*, 2007, *115* (3), 482–493.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection bias in predicting future utility," *the Quarterly Journal of economics*, 2003, *118* (4), 1209–1248.

Mazar, Nina, Botond Koszegi, and Dan Ariely, "True context-dependent preferences? The causes of market-dependent valuations," *Journal of Behavioral Decision Making*, 2014, *27* (3), 200–208.

Mellers, Barbara A and Alan DJ Cooke, "Trade-offs depend on attribute range.," *Journal of Experimental Psychology: Human Perception and Performance*, 1994, *20* (5), 1055.

Messinis, George, "Habit formation and the theory of addiction," *Journal of Economic Surveys*, 1999, *13* (4), 417–442.

Modica, Emmanuelle, Sabine Laudage, and Michelle Harding, "Domestic Revenue Mobilisation: A new database on tax levels and structures in 80 countries," 2018.

Mollerstrom, Johanna, Bjørn-Atle Reme, and Erik Ø Sørensen, "Luck, choice and responsibility—An experimental study of fairness views," *Journal of Public Economics*, 2015, *131*, 33–40.

Mullainathan, Sendhil, "A memory-based model of bounded rationality," *The Quarterly Journal of Economics*, 2002, *117* (3), 735–774.

Mullen, Brian, Jennifer L Atkins, Debbie S Champion, Cecelia Edwards, Dana Hardy, John E Story, and Mary Vanderklok, "The false consensus effect: A meta-analysis of 115 hypothesis tests," *Journal of Experimental Social Psychology*, 1985, *21* (3), 262–283.

Noguchi, Takao and Neil Stewart, "In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions," *Cognition*, 2014, *132* (1), 44–56.

Nunnari, Salvatore and Jan Zápal, "A Model of Focusing in Political Choice," *CEPR Discussion Paper*, 2017.

Pennycook, Gordon and Valerie A Thompson, "Base-rate neglect," in "Cognitive illusions," Psychology Press, 2016, pp. 44–61.

Pettibone, Jonathan C and Douglas H Wedell, "Examining models of nondominated decoy effects across judgment and choice," *Organizational behavior and human decision processes*, 2000, *81* (2), 300–328.

Peysakhovich, Alexander and David G Rand, “Habits of virtue: Creating norms of cooperation and defection in the laboratory,” *Management Science*, 2016, *62* (3), 631–647.

__ **and** __ , “In-group favoritism caused by Pokémon Go and the use of machine learning for principled investigation of potential moderators,” *Available at SSRN 2908978*, 2017.

__ **, Martin A Nowak, and David G Rand**, “Humans display a ‘cooperative phenotype’that is domain general and temporally stable,” *Nature communications*, 2014, *5* (1), 1–8.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva, “Optimal taxation of top labor incomes: A tale of three elasticities,” *American economic journal: economic policy*, 2014, *6* (1), 230–71.

Pryor, Campbell, Amy Perfors, and Piers DL Howe, “Conformity to the descriptive norms of people with opposing political or social beliefs,” *PloS one*, 2019, *14* (7), e0219464.

Rabin, Matthew, “Incorporating fairness into game theory and economics,” *The American Economic Review*, 1993, pp. 1281–1302.

__ , “Inference by believers in the law of small numbers,” *The Quarterly Journal of Economics*, 2002, *117* (3), 775–816.

Ratneshwar, Srinivasan, Allan D Shocker, and David W Stewart, “Toward understanding the attraction effect: The implications of product stimulus meaningfulness and familiarity,” *Journal of Consumer Research*, 1987, *13* (4), 520–533.

Reigstad, Amanda G, Eirik A Strømland, and Gustav Tinghög, “Extending the cooperative phenotype: Assessing the stability of cooperation across countries,” *Frontiers in psychology*, 2017, *8*, 1990.

Ross, Lee, David Greene, and Pamela House, "The "false consensus effect": An egocentric bias in social perception and attribution processes," *Journal of experimental social psychology*, 1977, *13* (3), 279–301.

Roth, Benjamin and Andrea Voskort, "Stereotypes and false consensus: How financial professionals predict risk preferences," *Journal of Economic Behavior & Organization*, 2014, *107*, 553–565.

Savage, Leonard, *The Foundations of Statistics*, John Wiley and Sons, 1954.

Schkade, David A and Daniel Kahneman, "Does living in California make people happy? A focusing illusion in judgments of life satisfaction," *Psychological Science*, 1998, *9* (5), 340–346.

Schultz, P Wesley, Jessica M Nolan, Robert B Cialdini, Noah J Goldstein, and Vladas Griskevicius, "The constructive, destructive, and reconstructive power of social norms," *Psychological science*, 2007, *18* (5), 429–434.

Schwartzstein, Joshua, "Selective attention and learning," *Journal of the European Economic Association*, 2014, *12* (6), 1423–1452.

Seshadri, Arjun, Alex Peysakhovich, and Johan Ugander, "Discovering context effects from raw choice data," in "International Conference on Machine Learning" PMLR 2019, pp. 5660–5669.

Shayo, Moses and Asaf Zussman, "Judicial ingroup bias in the shadow of terrorism," *The Quarterly journal of economics*, 2011, *126* (3), 1447–1484.

Simonsohn, Uri and George Loewenstein, "Mistake: The Effect of Previously Encountered Prices on Current Housing Demand," *The Economic Journal*, 2006, *116* (508), 175–199.

Simonson, Itamar, "Choice based on reasons: The case of attraction and compromise effects," *Journal of consumer research*, 1989, *16* (2), 158–174.

— **and Amos Tversky**, "Choice in context: Tradeoff contrast and extremeness aversion," *Journal of marketing research*, 1992, *29* (3), 281–295.

— **and** — , "Context-dependent preferences," *Management science*, 1993, *39* (10), 1179–1189.

Smith, Alec, "Lagged beliefs and reference-dependent utility," *Journal of Economic Behavior & Organization*, 2019, *167*, 331–340.

Smith, Vernon L, "An experimental study of competitive market behavior," *Journal of political economy*, 1962, *70* (2), 111–137.

— , "Effect of market organization on competitive equilibrium," *The Quarterly Journal of Economics*, 1964, *78* (2), 181–201.

— , "Experimental auction markets and the Walrasian hypothesis," *Journal of Political Economy*, 1965, *73* (4), 387–393.

Soltani, Alireza, Benedetto De Martino, and Colin Camerer, "A range-normalization model of context-dependent choice: a new model and evidence," *PLoS computational biology*, 2012, *8* (7), e1002607.

Somerville, Jason, "Choice-Set-Dependent Preferences in Consumer Choice: An Experimental Test," *Available at SSRN 3590240*, 2020.

Somville, Vincent, Asbjørn Andersen, Simon Franklin, Tigabu Getahun, Andreas Kotsadam, and Espen Villanger, "Does wealth reduce support for redistribution? evidence from an ethiopian housing lottery," *NHH Dept. of Economics Discussion Paper*, 2020, (18).

Sugden, Robert, "Reference-dependent subjective expected utility," *Journal of economic theory*, 2003, *111* (2), 172–191.

Tajfel, Henri, John C Turner, William G Austin, and Stephen Worchel, "An integrative theory of intergroup conflict," *Organizational identity: A reader*, 1979, *56* (65), 9780203505984–16.

__ **, Michael G Billig, Robert P Bundy, and Claude Flament**, "Social categorization and intergroup behaviour," *European journal of social psychology*, 1971, *1* (2), 149–178.

Thaler, Richard H, "Mental accounting matters," *Journal of Behavioral decision making*, 1999, *12* (3), 183–206.

Tversky, Amos, "Elimination by aspects: A theory of choice.," *Psychological review*, 1972, *79* (4), 281.

__ , "Features of similarity.," *Psychological review*, 1977, *84* (4), 327.

__ **and Daniel Kahneman**, "Belief in the law of small numbers.," *Psychological bulletin*, 1971, *76* (2), 105.

__ **and** __ , "The framing of decisions and the psychology of choice," *Science*, 1981, *211* (4481), 453–458.

Wachter, Jessica A and Michael Jacob Kahana, "A retrieved-context theory of financial decisions," Technical Report, National Bureau of Economic Research 2019.

Wedell, Douglas H, "Distinguishing among models of contextually induced preference reversals.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1991, *17* (4), 767.

__ **and Jonathan C Pettibone**, "Using judgments to understand decoy effects in choice," *Organizational Behavior and Human Decision Processes*, 1996, *67* (3), 326–344.

Wenzel, Michael, "An analysis of norm processes in tax compliance," *Journal of economic psychology*, 2004, *25* (2), 213–228.

— , "Misperceptions of social norms about tax compliance: From theory to intervention," *Journal of Economic Psychology*, 2005, *26* (6), 862–883.