Essays on Distributive Preferences

Weber, Nina

Awarding institution:
King's College London

Essays on Distributive Preferences

July 2023

36,923 words

A thesis presented for the degree of

Doctor of Philosophy

by

Nina Weber

King's College London

Department of Political Economy

Supervisors: Shaun Hargreaves Heap and Konstantinos Matakos

Abstract

This dissertation addresses three questions in relation to distributive preferences. First, I test, together with Shaun Hargreaves Heap and Konstantinos Matakos, whether distributive choices made in experimental settings, in fact, reveal underlying social preferences. We find that while social preferences explain distributive choices to an extent, following descriptive social norms is a significantly better explanatory factor for these choices, irrespective of the institutional mechanisms used to elicit preferences and perceived social norms. Second, I test with cross-country survey data and a survey experiment whether personal experience with social mobility affects perceptions of procedural fairness in society and, in turn, distributive preferences. The results indicate a divide between people who experienced upward mobility as opposed to downward mobility – experiencing downward mobility increases support for redistribution while experiencing upward mobility does not affect redistributive preferences. Third, I test in an interactive online experiment how distributive choices are affected by potential positive externalities of risky decisions. Many personally risky decisions, such as innovation and entrepreneurship, have the potential to increase overall welfare by creating positive externalities for society. Rewarding such prosocial risk-taking may therefore be an important strategy in addressing societal challenges, but individuals' distributive preferences are a fundamental constraint for policymakers who wish to do so. The results of my experimental study indicate that although individuals have a preference for rewarding prosocial risk-taking, they display outcome bias in doing so.

Acknowledgements

There are many people whose support I am incredibly grateful for. My supervisors, Shaun Hargreaves Heap and Konstantinos Matakos, have not just been co-authors on one of my chapters, but are also the reason why I was able to pursue a PhD in the first place. Their encouragement and help to secure funding was the reason why I was able to quit my job and go back to university. Throughout the past years, starting already during my bachelor's degree, they have both supported me with every idea and plan that I wanted to pursue. Shaun has always had tremendous patience when reading the first, second, or fifth version of one of my chapters and has unfailingly provided me with useful and detailed feedback every time. I am fully aware that this kind of support is not common in academia and I know that I would not have been able to learn as much as I have without them.

There have also been many people at the Department of Political Economy and beyond whose advice, feedback, and general support I am very grateful for: Christel Koop, Damien Bol, Daniel Carvajal, David Hope, Elisa Cavatorta, Frauke Stehr, Heidi Thysen, Julian Limberg, Karen Jeffrey, Kai Barron, Max Lobeck, Peter John, Sam Hirshman, my examiners Sigrid Suetens and Simon Gächter, and Stormzy, for giving me the energy to spend so many days and nights with Stata and R over the years. I am also grateful for the 9.16 gang, Jeevun Sandher, and Peter Paine, who have made many long days at Bush House more enjoyable, especially through our table tennis afternoons.

Being able to spend three months in Bergen at the FAIR institute during my PhD was a dream I had for a long time. I am very grateful for Bertil Tungodden and Alexander Cappelen who made this possible and gave me so much of their time to develop my ideas further, in particular my last chapter.

Without the financial support of the Economic and Social Research Council, who provided me with a full scholarship through the London Interdisciplinary Social Science Doctoral Training Partnership, I would not have been able to complete my PhD. I am very grateful for the academic freedom and time the ESRC support gave me to pursue my research.

Finally, I want to thank my family, Toja, and Hannu, for always reminding me that there is more to life than research and for their continuous trust in my abilities, even when I was not so confident in them myself.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the central insights of behavioural economics is that people have social, and not just selfish, preferences. People give to charities (Andreoni and Payne 2013; Ottoni-Wilhelm et al. 2017), voluntarily contribute to public goods (Ledyard 1995; Fehr et al. 2002; Chaudhuri 2011), and vote for redistributive policies they do not directly benefit from (Alesina and Giuliano 2011; Enke 2019; Bregman 2020). One of the most direct measures of social preferences in economic decision-making are distributive choices. This thesis will focus on three questions in relation to distributive preferences in experimental and survey data.

Studies measuring distributive choices go back to simple dictator games in which one subject must decide how to divide an endowment between herself and another subject (Güth et al. 1982). This basic experimental design is commonly used to study social preferences as it presumably eliminates any potential other motivations from the subject's choice set. The subject making the choice is hereby the dictator as the other subject can neither accept nor reject the proposed division.

Strict rational choice theory would predict in this setting that the dictator keeps the entire endowment for herself, as this choice maximises her own monetary payoff. This is however not what most experiments find. In a meta-analysis of dictator games, Engel (2011) finds that, in fact, 63.89% of dictators give at least some of their endowment to the other subject. On average, dictators give 28.35% of their endowment away. Unlike in the ultimatum game, where the receiver can reject or accept the division proposed by the first mover, the dictator game eliminates any potential self-interested motivation from the choice of sharing. The results of this experiment have therefore been interpreted as strong evidence for the existence of social preferences for equality.

As pointed out by Franzen and Pointner (2012) and others, even the dictator game however does not entirely eliminate selfish preferences. Dictators may try to appear fair in front of the recipient or experimenters for social desirability, reputational or reciprocal motivations. They implement a randomized response technique (Warner 1965) whereby both, the experimenter and the receiver, cannot know which dictators made which specific choice. They find that under this condition, dictators give substantially less to the recipients and a much larger share of dictators keeps the entire endowment for themselves. More interestingly though, about 40% of dictators still gave some of their endowment away. Even if reputational and reciprocal motivations are therefore close to eliminated, many people still express social preferences.

Given the importance of distributive choices for understanding social preferences, a growing literature in experimental economics and beyond aims to understand the nature of these choices and what shapes them (e.g. Fehr and Schmidt 1999; Corneo and Grüner 2002; Falk et al. 2003; Alesina and Angeletos 2005; Klor and Shayo 2010; Cappelen et al. 2010; Alesina and Giuliano 2011; Luttmer and Singhal 2011; Cappelen et al. 2013a,b; Durante et al. 2014; Kuziemko et al. 2015; Cappelen et al. 2016; Alesina et al. 2018; Stantcheva 2021; Alesina et al. 2023). Factors which have been found to influence distributive preferences fairly consistently are, among others, gender – with women giving significantly more than men (Engel 2011), age – older people giving significantly more (Engel 2011), economics education – with those trained in economics giving less (e.g. Fehr et al. 2006), and, as previously discussed, anonymity – with less giving when choices are kept anonymous (e.g. Andreoni and Bernheim 2009; Engel 2011; Franzen and Pointner 2012).

A particularly important factor identified in the recent literature are procedural fairness perceptions (Cherry et al. 2002; Cherry and Shogren 2008; Cappelen et al. 2010; Krawczyk 2010; Brock et al. 2013; Cappelen et al. 2013a; Mollerstrom et al. 2015; Lefgren et al. 2016; Gee et al. 2017; Rey-Biel et al. 2018; Almås et al. 2020). If outcomes have been reached through procedures which are perceived as fair by the individual who judges them, less money tends to be redistributed. Generally, outcomes tend to be deemed fair in these studies if people had agency in the procedure to reach these outcomes.

In this thesis, I focus on three questions in relation to distributive preferences which have so far received comparably little attention.

In the first paper, together with Shaun Hargreaves Heap and Konstantinos Matakos, I test whether distributive choices are, in fact, motivated by underlying social preferences, or whether perceived social norms can better predict the distributive choices people make. In a series of online experiments, we ask subjects to make dictator-like distribution decisions for groups of five. Prior to asking subjects to make their decision, we elicit social preferences and perceived injunctive and descriptive social norms for each subject. We then test whether the elicited social preferences are consistent with the choices made and whether social norms can provide a plausible alternative motivation for the decision-making of our subjects. We also run three separate treatments: an impartial spectator, a veil of ignorance, and a non-veil of ignorance treatment, to test whether the elicitation procedure affects the preference- or norm-following of our subjects.

Our main finding is that while social preferences predict some of the decisions our subjects make, descriptive social norms are a significantly better predictor of distributive choices. Interestingly, the elicitation mechanism used does not affect the preference- or norm-following of our subjects. In a series of follow-up experiments, we probe the validity of our results and test potential explanations for our findings. Here, we find that subjects who are more confident in their social preference are also more likely to follow it in the distribution decision and are less likely to follow the descriptive social norm instead.

We further find evidence for this type of norm-following in the distribution task to translate to a different decision problem; specifically, a one-shot hypothetical public goods game. Here, subjects who followed the descriptive social norm in the distribution decision also contribute more to the public good on average.

Our results are particularly important for the use of criteria like the Pareto criterion in welfare evaluations. If distributive choices do not always reflect underlying social preferences but instead perceived social norms, then evaluating the welfare improvement of policies faces a significant challenge.

In the second paper, I address a puzzle observed in the literature on social mobility and distributive preferences. While perceptions of social mobility fairly robustly predict preferences for redistribution, experience of social mobility does not. I test a potential explanation for this puzzle; specifically, the self-serving bias. I first use cross-country survey data from the International Social Survey Programme (ISSP) and the Luxembourg Income Study (LIS) to document how the intergenerational socio-economic mobility of respondents correlates with their perceptions of societal mobility and their distributive preferences. I find that there is a divide between people who experienced upward mobility as opposed to downward mobility – experiencing downward mobility increases support for redistribution while experiencing upward mobility does not affect redistributive preferences. This divide is in line with the self-serving bias: Those with negative mobility experiences *blame the system* and extrapolate from their experience onto society, which increases their demand for redistribution. Conversely, those who experienced positive mobility *accept the system* and do not extrapolate from their experience onto society, leading to no change in support for redistribution.

In a subsequent information provision survey experiment I test the causality of this pattern by providing respondents with a shock to their personal mobility experience. I find that receiving a negative mobility shock decreases perceptions of societal mobility and increases support for redistributive policies while a positive mobility shock has no effect on these variables.

The findings of this paper suggest a potential alternative explanation for the Great Gatsby Curve: As overall absolute mobility decreases (increases), ceteris paribus, demand for redistribution also decreases (increases).

In the third paper, I test how potential positive externalities of risk-taking behaviour affect the distributive choices of third-party spectators by developing a theoretical framework and testing its theoretical predictions in an interactive online experiment. Many personally risky decisions, such as innovation and entrepreneurship, have the potential to increase overall welfare by creating positive externalities for society. Rewarding such prosocial risk-taking may therefore be an important strategy in addressing societal challenges like, for example, the climate emergency, by promoting innovation that has positive externalities for the environment. A fundamental constraint for policy makers in rewarding such behaviour are, however, individuals' distributive preferences. So far, the role of potential externalities has however not received a lot of attention in the literature on distributive preferences.

The results of the experiment indicate that individuals, on average, have a preference for rewarding prosocial risk-taking, but they display outcome bias in doing so. Specifically, individuals reward prosocial risk-taking primarily if the positive externalities ultimately realise and do not compensate unlucky prosocial risk-takers. If outcomes are, however, unknown, prosocial risk-taking is substantially rewarded. These findings have potential implications for the stability of policies aimed at rewarding prosocial risk-taking: Policies which are supported ex ante might lose support ex post, once outcomes are realised.

The findings of this paper also suggest that outcome bias poses an important challenge to the role of procedural fairness in explaining distributive choices: Although risk-takers cannot affect the likelihood of whether positive externalities realise ex post in the experimental design, they are nonetheless held accountable for it through the distributive choices made by the spectators.

Bibliography

Alberto Alesina and George-Marios Angeletos. Fairness and redistribution. *American Economic Review*, 95(4):960–980, 2005.

Alberto Alesina and Paola Giuliano. Preferences for redistribution. In *Handbook of Social Economics*, volume 1, pages 93–131. Elsevier, 2011.

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

Alberto Alesina, Armando Miano, and Stefanie Stantcheva. Immigration and redistribution. *The Review of Economic Studies*, 90(1):1–39, 2023.

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

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

James Andreoni and A Abigail Payne. Charitable giving. In *Handbook of Public Economics*, volume 5, pages 1–50. Elsevier, 2013.

R Bregman. *Humankind: A hopeful history.* London: Bloomsbury, 2020.

J Michelle Brock, Andreas Lange, and Erkut Y Ozbay. Dictating the risk: Experimental evidence on giving in risky environments. *American Economic Review*, 103(1):415–37, 2013.

Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden. Responsibility for what? fairness and individual responsibility. *European Economic Review*, 54(3):429–441, 2010.

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

Alexander W Cappelen, Karl O Moene, Erik Ø Sørensen, and Bertil Tungodden. Needs versus entitlements—an international fairness experiment. *Journal of the European Economic Association*, 11(3):574–598, 2013b.

Alexander W Cappelen, Sebastian Fest, Erik Ø Sørensen, and Bertil Tungodden. Choice and personal responsibility: What is a morally relevant choice? *The Review of Economics and Statistics*, pages 1–35, 2016.

Ananish Chaudhuri. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, 14(1):47–83, 2011.

Todd L Cherry and Jason F Shogren. Self-interest, sympathy and the origin of endowments. *Economics Letters*, 101(1):69–72, 2008.

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

Giacomo Corneo and Hans Peter Grüner. Individual preferences for political redistribution. *Journal of Public Economics*, 83(1):83–107, 2002.

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

Christoph Engel. Dictator games: A meta study. *Experimental Economics*, 14:583–610, 2011.

Benjamin Enke. Kinship, cooperation, and the evolution of moral systems. *The Quarterly Journal of Economics*, 134(2):953–1019, 2019.

Armin Falk, Ernst Fehr, and Urs Fischbacher. On the nature of fair behavior. *Economic Inquiry*, 41(1):20–26, 2003.

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

Ernst Fehr, Urs Fischbacher, and Simon Gächter. Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, 13:1–25, 2002.

Ernst Fehr, Michael Naef, and Klaus M Schmidt. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment. *American Economic Review*, 96(5):1912–1917, 2006.

Axel Franzen and Sonja Pointner. Anonymity in the dictator game revisited. *Journal of Economic Behavior & Organization*, 81(1):74–81, 2012.

Laura K Gee, Marco Migueis, and Sahar Parsa. Redistributive choices and increasing income inequality: experimental evidence for income as a signal of deservingness. *Experimental Economics*, 20(4):894–923, 2017.

Werner Güth, Rolf Schmittberger, and Bernd Schwarze. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4):367–388, 1982.

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

Michał Krawczyk. A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. *Journal of Public Economics*, 94(1-2):131–141, 2010.

Ilyana Kuziemko, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva. How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, 105(4):1478–1508, 2015.

O Ledyard. Public goods: some experimental results. *Handbook of Experimental Economics*, 1, 1995.

Lars J Lefgren, David P Sims, and Olga B Stoddard. Effort, luck, and voting for redistribution. *Journal of Public Economics*, 143:89–97, 2016.

Erzo FP Luttmer and Monica Singhal. Culture, context, and the taste for redistribution. *American Economic Journal: Economic Policy*, 3(1):157–79, 2011.

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

Mark Ottoni-Wilhelm, Lise Vesterlund, and Huan Xie. Why do people give? testing pure and impure altruism. *American Economic Review*, 107(11):3617–3633, 2017.

Pedro Rey-Biel, Roman Sheremeta, and Neslihan Uler. *When income depends on performance and luck: The effects of culture and information on giving.* Emerald Publishing Limited, 2018.

Stefanie Stantcheva. Perceptions and preferences for redistribution. Technical report, National Bureau of Economic Research, 2021.

Stanley L Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965.

Chapter 2

Do Distributive Choices reveal Social Preferences? The Role of Social Norms

with Shaun Hargreaves Heap and Konstantinos Matakos

I Introduction

Experimental Economics has provided ample evidence that people frequently choose to reduce their own income to reduce inequality within groups or to help others (see Charness and Rabin 2002; Fehr et al. 2006; Bolton and Ockenfels 2006; Cappelen et al. 2013a). Such behaviour is perfectly compatible with standard models of rational choice if it reflects an underlying social preference that motivates behaviour (Andreoni and Miller 2002). Specifically, when people give up part of their own income to reduce inequality within a group, this behaviour might, for example, be motivated by a social preference for equality. In this paper, we test whether such distributive choices, in fact, reflect underlying social preferences.

Under the assumptions of revealed preference theory, preferences can always be inferred by the choices people make (Samuelson 1938). If a person chooses to reduce inequality within a group, this person has, therefore, revealed a preference for equality. A behaviour which seemingly contradicts this assumption is, however, that individuals' choices are often inconsistent across different decision problems when it is assumed that social preferences revealed in one decision apply to another (Andreoni and Miller 2002; Levitt and List 2007; Blanco et al. 2011). A potential explanation for the inconsistencies in revealed social preferences across decision problems could be that social preferences are context-dependent and that different preferences dominate in different decision problems. This explanation is however somewhat tautological.

An alternative explanation would simply be that distributive choices do not actually reveal social preferences, or, at least, not exclusively. Instead, they may reflect perceived social norms. Given that social norms vary significantly across different decision problems (List 2007; Kimbrough and Vostroknutov 2016), this explanation could shed light on the inconsistency of people's distributive choices in the experimental literature. We, therefore, focus specifically on social norms as an alternative to social preferences in explaining distributive choices.

In our pre-registered experiment[1], our subjects make four decisions. Prior to making the decisions, subjects are told how a particular distribution of income in a group of people arose. In their first decision, they are asked to select a principle of justice, from a set of four, that they think should govern distribution for that group. This first decision identifies the personal principle of a subject. The idea of this elicitation is that if a person indeed assesses a distributive situation through a social preference, then these preferences will be underpinned by a personally held principle of justice. We justify this key assumption because this connection is similarly made by economists when they categorise revealed social preferences in distribution experiments (e.g. see Charness and Rabin 2002; Fehr et al. 2006; Bolton and Ockenfels 2006; Cappelen et al. 2013a).

In the second decision, we then elicit what people regard as the appropriate behaviour in the given circumstance - the *injunctive* norm (Cialdini et al. 1990) – through a Krupka and Weber (2013) elicitation. There are two ways in which an injunctive norm might be related

[1] The pre-analysis plan was registered at Harvard Dataverse and can be accessed here as well as in appendix IX.5. The experiment was granted ethical clearance from the Research Ethics Committee at King's College London (reference number MRS-18/19-9148).

to whether or not someone is guided by a social preference. First, if a person has doubts about the correctness of their personal view about what should be done, they might turn to the injunctive social norm to reduce those doubts. Equally, in social identity theory, people acquire a sense of identity by behaving in accordance with their group's injunctive social norm (e.g. see Tajfel et al. 1979; Akerlof and Kranton 2000). The injunctive norm might therefore affect the personal principle of our subjects. The second possibility is that people follow the injunctive norm in their distribution choice directly due to conformism. This second possibility would contradict more directly the assumption that choices reflect a person's preferences. Explaining conformism within the revealed preference model is also self-contradictory as a person would need to have a preference to not follow their own preference, yet follow this first preference.

The third decision is our main outcome variable and it is incentivized. Subjects are asked to choose a distribution of income for the group. They are given four possible distributions, each corresponding to one of the four principles of justice identified in the first two decisions.

In the final decision, we elicit what people regard as the usual behaviour - the *descriptive* norm (see Cialdini et al. 1990) – again using a Krupka and Weber (2013) elicitation. Following a descriptive norm in the distribution choice would provide an even stronger case against the revealed preference model than following an injunctive norm. While the previously discussed concern about conformism remains when following descriptive norms, they are different from injunctive norms in that they are concerned with what *is* as opposed to what *ought* to be. They can therefore not as easily be reframed as a different version of personal principle following. Someone who follows a descriptive norm might nonetheless be motivated by what *ought* to be done. This argument is made by Adam Smith in his *Theory of Moral Sentiments* (1759) where he suggests that when selfish interests conflict with what a person believes *ought* to be done, a person cannot be relied upon to judge how best to act. In such instances, people need an external standard to guide their behaviour and this standard can be provided by the behaviour of others. Despite a motivation to make a morally correct choice, this version of descriptive norm following is a distinct challenge to the social preference model.

Our subjects make these four decisions in one of three different treatment groups. Our treatments vary the elicitation mechanism used to identify subjects' social preferences and norms. Specifically, we ask subjects to make their decisions as either impartial spectators, whereby they have no personal stake in the decision, behind a veil of ignorance, where they have a personal stake but do not know their position in the distribution, or as a stakeholder who knows her likely position.

We then observe subjects' consistency between their personal principle, perceived social norms, and their actual distribution choice. We further test whether any deviations between personal principles and distribution choices can be explained by a desire to follow an injunctive or descriptive social norm.

Our results suggest that descriptive norm following is a better predictor of distributive choices than personal principles. We find that people's chosen distributions differ significantly from their chosen principles, both at the aggregate and at the individual level. We further find that this discrepancy can be explained by descriptive norm-following for a significant share of our subjects. Somewhat surprisingly, but perhaps less so given our main finding, the elicitation mechanism we use to identify people's preferences does not affect the level of preference-following, nor does it significantly affect the actual distributive choices made.

To probe the validity of our findings we conduct several robustness checks. First, we run a second experiment that inverts the order between the actual distribution decision (the third decision above) and the decision that reveals a person's perception of the descriptive norm (the fourth decision referred to above). We do this to test for the possibility that the actual distribution choices influence perceptions of the descriptive social norms. This further experiment also allows us to explore the origins of such descriptive norm-guided behaviour. The second experiment again reveals the primacy of descriptive social norms and it reinforces the social norm account by yielding some plausible insights into why people follow such social norms. In particular, when subjects are confident in their choice of personal principle, they are more likely to follow it in the distribution decision and an individual's strong social identification helps build such confidence. Lower levels of confidence, in contrast, are more likely to lead to selfish or descriptive norm-following behaviour. These additional findings are broadly consistent with Adam Smith's account of norm following in the *Theory of Moral Sentiments* (1759). They also suggest that because such distributive decisions are rare, some might not have strong social preferences, indicated by low confidence in one's stated principle, and therefore follow what they perceive to be the descriptive social norm instead.

In our second and third robustness checks, we examine with further surveys two additional possibilities that might have contributed to the weak evidence in favour of social preferences in our experiment. One concerns the possibility that people are guided by more than one personal principle of justice and, in such circumstances, their secondary personal principle might explain the drift to maximin outcomes in the data. The other concerns the possibility that our subjects may not be able to associate a principle of justice with a particular distribution outcome and so could be unable to apply their personal principle when making the distribution decision. Again, we conclude the original result favouring descriptive norm following is robust to these considerations.[2]

Given that our main results suggest distributive choices are better explained by descriptive norms rather than social preferences, we test, in a subsequent hypothetical one-shot public goods game, whether this type of descriptive norm-following in the distribution task translates into norm-following, specifically conditional cooperation, in this different context. Our results suggest that this is indeed the case and that descriptive norm-following in the distribution task is, on average, associated with larger contributions in the public goods game. This result suggests that the type of descriptive norm-following we identify in the distribution task also translates into at least this particular other experimental context.

We are not the first to examine whether social preferences or social norms better explain prosocial behaviour (e.g. see Ellingsen et al. 2012; Gächter et al. 2013; Guala et al. 2013). The evidence from these earlier studies typically comes from trust and public goods games, although not exclusively, and is mixed in its conclusions. Our study is distinct from these earlier studies in two important ways.

First, in trust and public goods games, social norms can function as coordination devices

[2] We also subject these results to various robustness checks regarding the wording of the principles, see appendix sections IX.1.3 and IX.3.7

when there are multiple Nash equilibria. Social norms are therefore not a challenge to the social preference model in this context but a complement, as norm-following helps to coordinate on an equilibrium that best satisfies preferences. For norm-following to be a challenge to the social preference model, it must be distinct from preference-following. We focus on non-interactive distribution games to ensure that there is no scope for norms to act as coordination devices. Our version of norm-following is therefore potentially more challenging to the social preference model than that used in earlier studies.

Second, our test is more direct. In earlier studies, social preferences are usually identified through a particular theory of what social preferences are expected to be present in the given setting. With a particular theory in mind, these studies then test whether social norms or the given theory can better organise the data. Our approach is more direct and requires fewer background assumptions, because we ask our subjects to identify directly what, in effect, if they were motivated by a social preference, would be its character. This approach would be more difficult to do in trust and public goods games than in a simple distribution game because there are a larger number and variety of potential social or moral motives that could be at play.

While it is sufficient for positive economics to assume that choices always reveal preferences, it is not for welfare economics and, especially, for the use of criteria like the Pareto criterion. To determine whether a distributive choice is welfare-enhancing, it is important to know whether choices actually reflect underlying preferences. This is the case because welfare improvements are judged by preference satisfaction if the Pareto criterion is applied and these judgements can only apply to a world where people are actually preference satisfiers. Suppose this is not the case and distributive choices reflect something other than social preferences. In that case, one cannot be certain that a policy implemented based on such a choice, in fact, increases welfare.

Additionally, whether choices actually reflect social preferences also matters for the stability of implemented policies and, more broadly, for constitutional politics, which establishes the rules or principles that should guide action in a society (e.g. see Rawls 1971). For the stability of any policy or constitutional arrangement, it is important that consistency between preferences over principles and actual decisions over outcomes exists. If individuals however regularly agree to allocative principles, the practical implementation of which they oppose, then it may create a build-in source of opposition to the constitution or to distributive policies.

The outline of this paper is as follows. In section II, we will begin by outlining the relevant literature and theory. Section III describes the experimental design and section IV lists the hypotheses and the empirical strategy. Section V reports the main results and section VI reports the results of a follow-up experiment aimed at testing the underlying motivations for norm-following in the distributive task. Section VII reports further results of a series of robustness checks and section VIII concludes.

II Literature Review

As previously outlined, a large literature in experimental economics is concerned with measuring distributive preferences. Most commonly, distributive preferences are studied with dictator games, whereby one subject has to decide how much of their endowment to give to another. Even once all reputational and reciprocal motivations are close to eliminated, about 40% of subjects still give some of their endowment away (Franzen and Pointner 2012).

What dictator game results illustrate is that both, selfish and social preferences, appear to influence distributive decisions in redistributive tasks. A growing literature is therefore concerned with identifying the social preferences motivating people's choices. This literature tends to consist of experiments in which subjects make distributive choices, whereby each choice is classified as being motivated by one or more specific social preference. Subsequently, the set of distributive choices made by subjects is analysed by testing which social preferences can best organise the data (e.g. Fehr et al. 2006; Bolton and Ockenfels 2006; Cappelen et al. 2010, 2013a).

Based on this literature, four main social preferences commonly found to motivate distributive choices can be identified:

II.1 Inequality aversion

Considerable experimental evidence has found that an aversion to inequality can help understand distributive decisions made by subjects (e.g. Fehr and Schmidt 1999; Charness and Rabin 2002; Fehr et al. 2006; Bolton and Ockenfels 2006). Inequality aversion hereby refers to a dislike for inequitable outcomes, meaning that a distribution of material payoffs that is equal is preferred to a distribution of material payoffs that is unequal (Fehr and Schmidt 1999). We represent this principle with the following statement in the experiment:

Inequalities should be minimized.

II.2 Maximin preferences

A somewhat weaker form of inequality aversion is a preference to maximise the income of the least well-off group in society. Such a social preference was first proposed by Rawls (1971) as the logical redistributive principle to be chosen under a veil of ignorance once equal freedoms are ensured. Experimental evidence on whether people indeed follow such a social preference in their choices exists (e.g. Charness and Rabin 2002; Engelmann and Strobel 2004), but is not as strong as the evidence for inequality aversion (e.g. see Fehr et al. 2006). In the experiment, we refer to this principle with the following statement:

Inequalities are only justifiable if they improve the position of the least well-off group in society.

II.3 Utilitarianism/Efficiency

Utilitarianism suggests that societies should aim to maximise happiness; or, in Bentham's words, produce the 'greatest happiness for the greatest number' (Mulgan 2014). In the absence of specific knowledge about how income translates into happiness for different people, this becomes a preference for arrangements that produce the highest average income level. This is in line with Harsanyi's derivation of utilitarianism from the same veil of ignorance procedure as Rawls, when individuals are expected utility maximisers as opposed to being expected to follow the maximin principle (Harsanyi 1980). He argues that a risk-neutral utility maximiser would aim to maximise average income behind the veil of ignorance. The distribution which maximises the average income in society would also be associated with exhausting all potential Pareto improvements (when allowing for compensation schemes) and so reflects a concern for efficiency.

There is robust experimental evidence that subjects indeed follow such a concern for efficiency or utilitarianism and care about maximising the average income of distributions. Such findings tend to be particularly strong for economics students (e.g. Engelmann and Strobel 2004; Fehr et al. 2006). The corresponding statement for this principle in our experiment is the following:

Income should be distributed to maximize the average income in society.

II.4 Meritocracy

Recent experimental studies on distributive choices have focused increasingly on another potential social preference: a preference for meritocracy. These studies differ somewhat from the previously discussed ones in that they incorporate another dimension into the experimental design – procedural justice. Of particular importance is thereby whether subjects interpret the source of the experimental income as being luck or effort. On average, people tend to be willing to redistribute a larger share of income if it is earned through luck as opposed to effort (Cappelen et al. 2013b). There is now substantial evidence in the literature that meritocratic concerns can explain much of the variation in distributive preferences (e.g. Almås et al. 2010; Cappelen et al. 2010; Krawczyk 2010; Cappelen et al. 2013a; Lefgren et al. 2016). This principle is referred to with the following statement:

Individual income should be based exclusively on his/her ability and talents.[3]

We are not aware of any studies directly comparing the exact same set of principles in distributive decision games as we do; however, as mentioned previously, various studies focus on testing the explanatory power of a subset of these principles. What is common to all these studies is that ex post explanations for behaviour are used to test which one can best

[3]Meritocracy is usually defined as also including a person's effort provision (Cappelen et al. 2010), which we did not explicitly state in our definition but would include if we were to rerun the experiment.

organise the data. Such a test is only valid if one assumes revealed preference theory to hold. This means that one has to assume that the result of the action, e.g. less inequality, is also the motivation for it.

A further point to note on the existing literature on distributive preferences and the dictator game in particular is that real-world distributive tasks almost never involve only two people. Increasingly, studies therefore test preferences in large groups. These studies tend to assume a redistribution parameter so that any preference is expressed as a degree of general redistribution from poor individuals to rich individuals, thereby limiting the potential principles motivating subjects' distributive choices (e.g. Durante et al. 2014). To allow for more nuanced motivations, Fisman et al. (2021) test distributive preferences in large groups by asking subjects on mTurk to choose between two hypothetical societies of 7-9 individuals with different income distributions. They find that both, inequality averse and maximin preferences, can explain decision-making, as well as a desire to reduce the income of the person directly 'above' the subject. This second finding is in line with 'last-place aversion' as found in Kuziemko et al. (2015).

The secondary research question of our paper asks whether social norms can explain any deviation away from one's social preferences when making distributive choices. This question is not new (e.g. see Ellingsen et al. 2012; Gächter et al. 2013; Guala et al. 2013). As discussed previously, the evidence from these earlier studies typically comes from trust and public goods games, although not exclusively, and is mixed in its conclusions.

II.5 Elicitation mechanism

We test three elicitation mechanisms that are most commonly used in experimental studies for the revelation of social preferences (e.g. see Durante et al. (2014) who also use all three mechanisms): Impartial spectator, veil of ignorance, non-veil of ignorance. We do this because the degree to which selfish interests can enter the decision-making process varies across treatments and, therefore, might affect the level of principle- and norm-following.

II.5.1 Impartial spectator

This mechanism asks subjects to make a distributive decision for a group of people they are not a part of. Their decision, therefore, does not affect their own pay-off which ensures that the decision can only reveal social preferences, as selfish preferences are removed from the choice set (e.g. see Cappelen et al. 2013a).

II.5.2 Veil of Ignorance

The veil of ignorance was developed by Rawls (1971) and asks subjects to make a distributive decision for a group of people they belong to; however, without knowing which position they will take up in this group. They will therefore be affected by the choice they make and selfish preferences may affect the decision-making process. However, as subjects do not know which position they will be in once the distributive decision is being applied, the extent to which selfish preferences can influence choices is limited.

II.5.3 Non-veil of Ignorance

The third mechanism removes the veil of ignorance so that subjects know which position they will likely hold in the group prior to making the distributive decision. Thus, the decision subjects make in this treatment is a combination of social and selfish preferences. Arguably, this mechanism is closest to distributive decisions made in the real world.

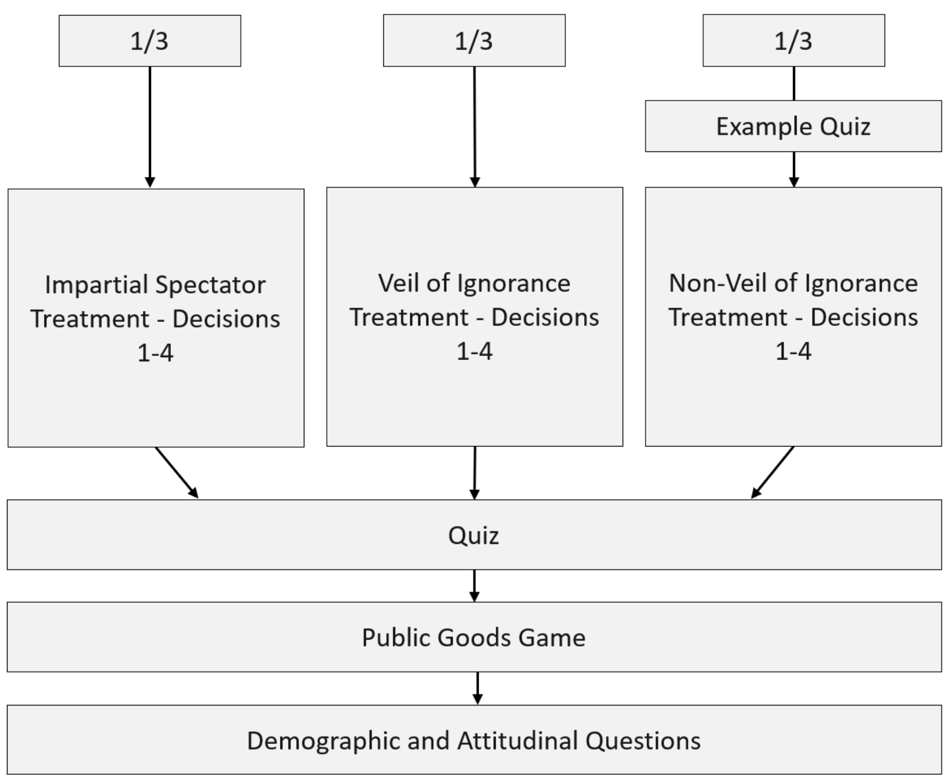
II.6 Norm-following and cooperation

In addition to our main experiment, we also test whether norm-following in the distribution decision is related to behaviour in a subsequent one-shot public goods game. One important behaviour observed in public goods games is conditional cooperation (Fischbacher et al. 2001). A conditional co-operator contributes when others contribute and reduces her contributions if others reduce theirs. Given that the largest group of subjects in public goods games appear to follow conditionally cooperative behaviour (Fischbacher et al. 2001), understanding what motivates conditional cooperation is relevant for an understanding of cooperative behaviour more broadly. A plausible interpretation of conditional cooperation would be that it is a version of descriptive norm-following (Fehr et al. 2002; Fehr and Fischbacher 2004; Vollan and Ostrom 2010). Subjects who want to follow the norm observe the behaviour of others and adjust their own contributions accordingly. There is, in fact, evidence that conditional cooperators who have a high propensity to contribute seek information about others' contributions - the descriptive norm - if given the option (Bigoni and Suetens 2012). Our experimental setting allows us to test whether this type of norm-following in the public goods game is related to norm-following in a different setting – our distribution task. If this is the case, it would suggest that conditional co-operators might be general norm-followers.

III Experimental Design

To test whether choices over outcomes in distributive settings reveal underlying social preferences, we leverage a within-subject design that requires people to choose a principle that they believe should govern distributive choices in a hypothetical group of five subjects, before

Figure 2.1: Experimental Design Overview



making an incentivised decision over allocations in this group.

The novel aspect of this experimental design is to identify people's preferences directly, by simply asking them which principle of distributive justice they believe should govern decision-making. The underlying assumption of our experimental design is that people are aware of the principle they apply in distributive decision-making. Figure 2.1 provides an overview of the experimental design.

Introduction

Subjects are told, by way of background, that a group of people are asked to do a quiz and their answers generate income. Their performance is ranked from the bottom 20% to the top 20% of performers and the average income by quintile of performers is given in table 2.1. This table was presented to all subjects in the introductory part of the experiment.

Table 2.1: Average Income per Quintile

Performance Level	*Average Income*
Bottom 20% of performers	£20
2nd 20%	£30
3rd 20%	£40
4th 20%	£70
5th 20%	£110

Decision 1: Choice of Principle

To identify subjects' social preference, we ask which of the four statements outlined in section II best describes how they believe income should be distributed in this group of five. The statements are presented in random order. As social preferences can be context-dependent and to hold assumptions about the distribution constant between subjects, it is important to identify people's social preference after the background situation is described and the status quo distribution is given.

Decision 2: Elicitation of injunctive Social Norm

To elicit subjects' perception of the injunctive social norm, we follow the experimental design

Figure 2.2: Injunctive Social Norm elicitation

All the participants in your group are now asked to select a statement. Each of you will be rewarded with a bonus payment of 50p if you select the statement chosen by most of the members of your group.

- Income should be distributed to maximize the average income in society.
- Individual income should be based exclusively on his/her ability and talents.
- Inequalities should be minimized.
- Inequalities are only justifiable if they improve the position of the least well-off group in society.

developed by Krupka and Weber (2013), who elicit social norms in incentivised coordination games. Our primary difference to the Krupka and Weber design is that we do not elicit norms with a separate subject pool but with the same subjects that participate in the main experiment.[4] We do this as we assume that what influences people's decisions is what they themselves perceive to be the norm, irrespective of whether there is a general agreement of what the social norm consists of. This is in line with d'Adda et al. (2016).

Figure 2.2 illustrates the question we ask all subjects to identify which principle they believe to be the injunctive norm within their group.[5]

Decision 3: Choice of Distribution

Subjects are now asked to make the main distributive decision of the experiment. They are informed that the income generated by the quiz can be redistributed in four possible ways and they are then asked to choose one of the options. Our experimental design ensures that

[4] In a robustness check experiment where we invert decision 3 and 4, we find a similar distribution of perceived social norms (in line with D'Adda et al.'s findings (2016)). In a further robustness check subject pool, however, we find significant differences in elicited social norms at the aggregate. The results can be found in appendices IX.2.1 and IX.3.3.

[5] In the impartial spectator treatment, the question we ask does not refer to "your group" but "the group". The exact wording of each part of the experiment can be found in appendix IX.5.

each possible principle matches directly onto a possible distribution option, allowing us to test whether the stated principle and the chosen distribution match. Such a test is only possible by restricting subjects' possible options to choose from a defined set of distributions. Table 2.2 provides an overview of each principle and its corresponding distribution.[6] These options were also randomised in order during the experiment.

Table 2.2: Distribution Options

	Average Income			
Performance Level	*Inequality Aversion*	*Maximin*	*Meritocracy*	*Utilitarianism*
Bottom 20%	£30	£40	£20	£20
2nd 20%	£60	£40	£30	£30
3rd 20%	£60	£50	£40	£50
4th 20%	£60	£60	£70	£70
5th 20%	£60	£80	£110	£110
Total	£270	£270	£270	£280

Notes: The exact wording and presentation of the distributions to respondents can be found in appendix IX.5.

By design, each of the four distribution options is only compatible with one of the stated principles. The inequality averse distribution minimises the average income differences between each person in the distribution. If one wants to maximise the income of the poorest person, only the maximin distribution can fulfil this principle. The meritocratic distribution is equivalent to the distribution based on quiz performance and the utilitarian distribution has the highest average and total income of all possible distributions.

Decision 4: Elicitation of descriptive Social Norm

As in decision 2, we now ask subjects to choose a distribution and we tell them that 'you will be rewarded with a bonus payment of 50p if you select the distribution chosen by most of the members of your group'. We are then able to compare the distribution chosen in decision 3 and 4 to test whether a perceived descriptive social norm guided subjects' distribution

[6] We also run a robustness check where we report the average income, rather than the total income, for each distribution option. The results can be found in appendix IX.3.2. The different wording as no effect on subjects' choices.

decision.

III.1 Treatments

In addition to our main research questions, we are interested in testing whether the method used to elicit preferences and distribution choices affects our findings. To do so we compare subjects' choices across three treatments, each using one of the three previously outlined elicitation mechanisms commonly used in distributive experiments: An impartial spectator treatment, a veil of ignorance treatment and a non-veil of ignorance treatment.

In the impartial spectator treatment (treatment 1), subjects were asked to make decisions 1-4 for a group of subjects they do not belong to.

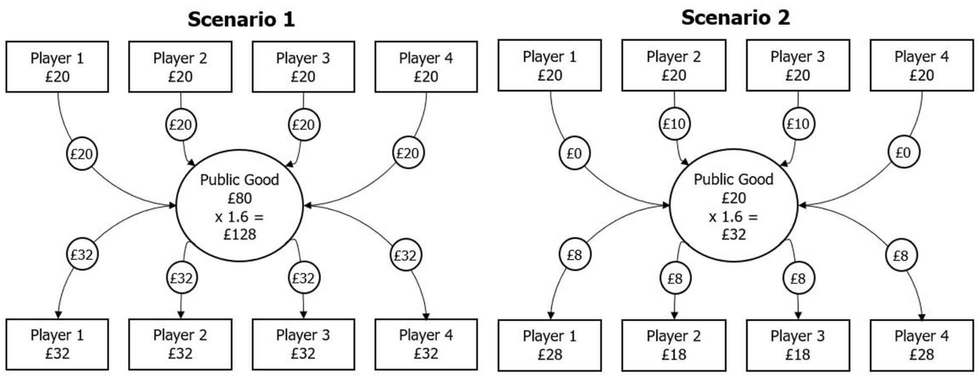
In the veil of ignorance treatment (treatment 2), subjects were told that they are part of the group, that they will take part in the quiz after making decisions 1-4 and that their performance, together with the selected distribution, will affect the bonus payment they receive. Subjects however did not know what the quiz consisted of when making decisions 1-4 and could therefore not predict their position in the distribution.

In the non veil of ignorance treatment (treatment 3), subjects were given the same information as in the veil of ignorance treatment but prior to making decisions 1-4 they were given two sample questions of the quiz. Based on their performance on these sample questions, we told subjects the quintile they would most likely be in for the actual quiz. When making decisions 1-4 subjects, therefore, knew their likely quintile position and so could, if they so wished, choose a distribution that would maximise the income of that position.

III.2 Public Goods Game

After making decisions 1-4 and completing the quiz, subjects were asked to take part in a one-shot hypothetical public goods game. A short explanation of the game was given in which subjects were told that the game consists of four players, each having an endowment of £20. The multiplier used was 1.6 and each subject was given two example scenarios prior to making the decision to illustrate the dynamics of the game. The example scenarios

Figure 2.3: Public Goods Game Example Scenarios



are outlined in figure 2.3. Given that we asked subjects to play a hypothetical one-shot public goods game, it is possible that higher contribution levels can be explained by subjects being conditional co-operators. That is, as conditional co-operators tend to start with high contribution levels in the first round and then decrease their contribution levels over time depending on other players' choices (Fischbacher et al. 2001). To account for this possibility, we use the strategy method design by Fischbacher et al. (2001) to create a control variable for conditional cooperation in one of our robustness checks. Previous research has found that subjects who can be classified as conditional cooperators with the strategy method also behave like conditional cooperators in repeated games (Fischbacher et al. 2012). The results of this robustness check can be found in appendix section IX.3.4.

IV Hypotheses and Empirical Strategy

To answer the primary research question, we compare subjects' chosen principle (decision 1) and chosen distribution (decision 3) for consistency.[7] If a subject chooses, for example, the maximin principle and the maximin distribution, she is coded as a social preference-follower. If her principle and distribution do not match, she is not. We, therefore, formulate our first hypothesis:

[7] We deviate slightly from our pre-analysis plan in this section and in some of the reporting of our results based on feedback we received since submitting the original pre-analysis plan in 2019. None of these changes substantially change the main motivation or hypotheses of the paper.

Hypothesis 1: A person's personal principle (their social preference) predicts their distribution choice.

To answer the secondary research question, which asks whether potential deviations away from social preferences in the distribution choice can be explained by norm-following, we test whether the variation in social preference following can be explained by norm following. Norm following is thereby defined as choosing a distribution in decision 3 that matches the perceived descriptive social norm elicited in decision 4, or the perceived injunctive norm elicited in decision 2. Our second hypothesis is then:

Hypothesis 2: Differences between subjects' chosen principle and distribution can be explained by descriptive and/or injunctive norm-following in the distribution choice.

As discussed previously, we expect there to be a significant difference in chosen principles and social preference following across our three treatments. In particular, we expect preference following to be higher in the impartial spectator treatment, which is generally assumed to be a clean test of social preferences as selfish considerations do not apply here (Cappelen et al. 2013a). We further expect selfish considerations to become more important in the non-veil treatment where the position in the distribution is known. This might reduce social preference following compared to treatments 1 and 2. This leads to our third hypothesis:

Hypothesis 3: Social preference following is lower in the non-veil of ignorance treatment compared to the other two treatments.

Lastly, we are interested in whether norm-following in the distribution decision helps predict behaviour in the subsequent one-shot public goods game. A potential channel through which norm-following and public goods game contributions might be linked is conditional cooperation. Conditional co-operators, on average, are likely to make higher contributions to the public good in a one-shot game than other types of players, for example, free riders. If conditional cooperation is a form of descriptive norm-following, and norm-following in the distribution task translates to norm-following in the public goods game, then we would expect contributions, on average, to be higher for subjects who are descriptive norm-followers. Based on the above we state our final hypothesis:

Hypothesis 4: There is a significant positive relationship between the level of norm-following in the distributive choice and subjects' contributions in the public goods game.

The main variable of interest in our estimation is the distributive choice subjects make in our experiments; namely, whether they choose inequality aversion, maximin, meritocracy or utilitarianism. We therefore estimate the following model:

$$D_{t,i} = \beta_0 + \beta_1 P_{t,i} + \beta_2 ND_{t,i} + \beta_3 NI_{t,i} + \lambda_{t,i} + \epsilon_{t,i} \tag{2.1}$$

Whereby D is the outcome measured as a dummy variable capturing the chosen distribution by subject i in treatment t, P is a dummy variable equal to 1 if the chosen social preference is equal to the chosen distribution, ND is a dummy variable equal to 1 if the elicited descriptive social norm is equal to the distribution choice, NI is a dummy variable equal to 1 if the elicited injunctive social norm is equal to the distribution choice, λ a vector of controls and ϵ is the error term.

As we further want to test whether any potential inconsistency between preferences and distribution choices can be explained by the role of perceived social norms, we further estimate the following binary logit model:

$$Y_{t,i} = \beta_0 + \beta_1 ND_{t,i} + \beta_2 NI_{t,i} + \lambda_{t,i} + \epsilon_{t,i} \tag{2.2}$$

Whereby Y is a binary outcome variable equal to 1 if subject i's choice of principle and outcome is consistent in treatment t, ND is a dummy variable equal to 1 if the elicited descriptive social norm is equal to the distribution choice, NI is a dummy variable equal to 1 if the elicited injunctive social norm is equal to the distribution choice, λ the vector of controls and ϵ the error term.

Lastly, to identify whether there is a relationship between norm-following in the distribution decision and public goods game contributions we will estimate the following additional model:

$$C_i = \beta_0 + \beta_1 PD_i + \beta_2 N_i + \lambda_i + \epsilon_i \tag{2.3}$$

Whereby C is the outcome variable measured as subject i's contribution to the one-shot public goods game, P is the chosen principle, D the chosen distribution, PD is a dummy variable equal to 1 if subject i followed her social preference in the distribution choice, N is a dummy variable equal to 1 if the elicited descriptive and/or injunctive social norm is equal to the distribution choice, λ the vector of controls and ϵ the error term.

Previous research has found risk-aversion to be an important factor in determining peoples' distributive preferences, particularly behind a veil of ignorance (Carlsson et al. 2005; Schildberg-Hörisch 2010). To account for this potentially biasing factor we include a control variable for risk-aversion in our model. We further control for participants' income as various studies have previously found a causal link between income-levels and preferences for redistribution (see e.g. Esarey et al. 2012; Owens and Pedulla 2014; Naumann et al. 2016). Previous research has also found participants' gender (Alesina and Angeletos 2005; Rehm 2005) and nationality to influence distributive preferences; particularly, whether participants live in the United States or Europe affects their preferred level of redistribution (Alesina and Glaeser 2004). In addition to these two factors we will also control for age and student status, as economics students have been found to be less inequality averse in experimental settings than the average population (Fehr et al. 2006), which may further bias our estimation.

Our online survey experiment was coded in Qualtrics and run with a subject pool of 2,408 people recruited via Prolific Academic. Our experiment was run in two sessions. The first session ran on the 14th of November 2019 and the second session on the 9th of December 2019. Both sessions were conducted in the late afternoon to ensure that both, participants in the US and Europe, could be reached. The experiment was pre-registered with Harvard Dataverse (Weber 2019).

V Results

Table 2.3 reports mean values of descriptive variables by assigned treatment. Most demographics are well-balanced between the treatment groups; however, the proportion of economics students is significantly different across treatment groups. Given that this vari-

Table 2.3: **Balance across treatment groups**

	Impartial Spectator	Non-Veil of Ignorance	Veil of Ignorance
Mean Value or Share			
Female	59.60	61.02	59.70
Age			
18-20	8.51%	9.90%	10.34%
21-29	38.10%	34.14%	34.12%
30-39	26.76%	30.58%	29.27%
40-49	13.32%	14.47%	13.08%
50-59	9.00%	8.12%	8.22%
60+	4.32%	2.79%	4.98%
Students	0.25	0.24	0.27
Economics	0.19**	0.24**	0.22
Income			
Under £20,000	49.80%	51.15%	54.13%
£20,000 to £34,999	26.76%	23.89%	26.53%
£35,000 to £44,999	13.32%	12.15%	9.60%
£50,000 to £74,999	5.59%	8.64%	5.73%
£75,000 to £99,999	2.40%	2.16%	1.60%
Over £100,000	2.13%	2.02%	2.40%
Sample			
United Kingdom	58.32%	56.69%	58.26%
United States	17.02%	19.32%	19.13%
Europe	24.66%	23.99%	22.61%
Left-Right	4.03	4.03	4.02
Risk preference	5.58	5.60	5.42
Quiz performance	2.25	2.55***	2.31
Observations	811	792	805

Notes: Table reports the mean values or category shares for each variable. Category shares indicate the share of respondents within a treatment group who fall into a particular category. Students and Economics are dummy variables. A higher value on the left-right scale from 1 to 5 indicates a more left-wing orientation on economic policy. Risk preferences are self-reported on a scale from 0 to 10 with 10 being the most risk-seeking option. Quiz performance ranges from 0 to 5 depending on how many questions the subject answered correctly. Asterisks indicate significant differences in mean values between treatment groups from a chi-squared test of independence. *** p<0.01, ** p<0.05 , * p<0.1.

able does not appear to influence the main outcome variables[8], it does not appear to be a problem for inference.

The table further reports that quiz performance is significantly higher in the non-veil of ignorance treatment. This is likely to be the case as respondents in this treatment answered two sample quiz questions prior to making their distributive decisions and were therefore better prepared for the actual quiz than respondents in the other two treatments.

Figure 2.4: Frequency distribution of principle, distribution choice, and perceived norms

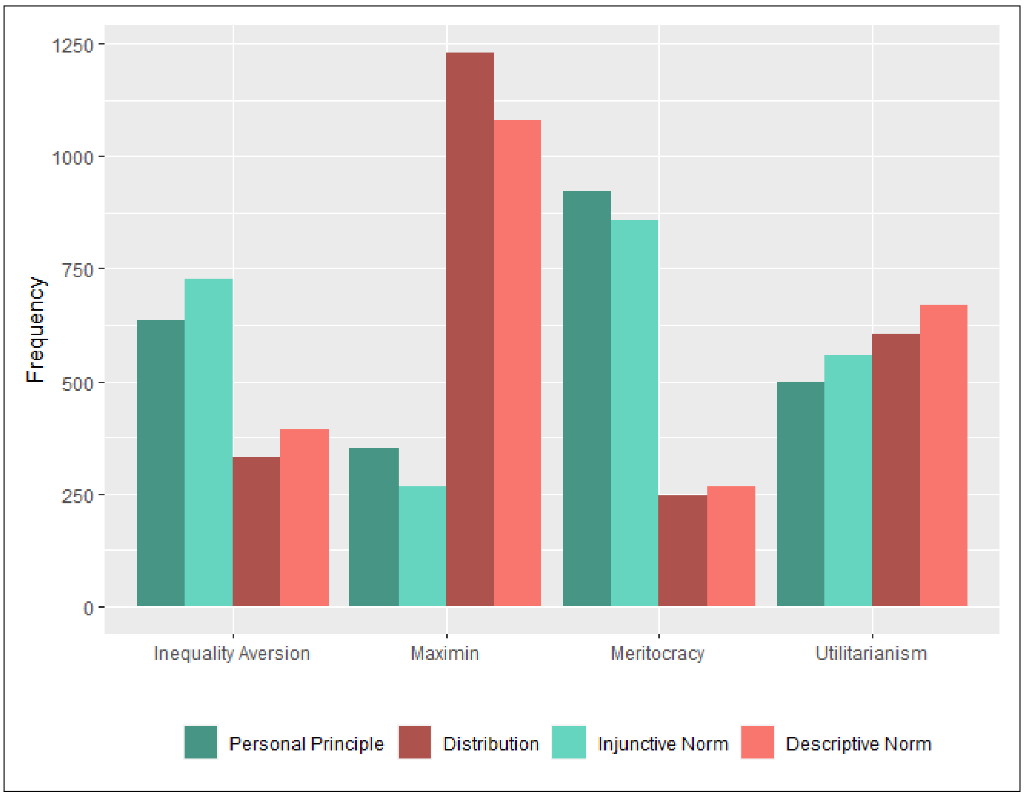


Figure 2.4 shows the aggregate distribution of subjects' chosen principle, distribution and elicited social norms for all treatments. At first glance, these results seem to suggest that distribution choices deviate significantly from chosen principles, in line with perceived descriptive social norms. Most evidently, this appears to be the case for two principles - meritocracy and maximin. 38% of subjects chose meritocracy as their preferred principle, while only 10% chose it in the distributive decision. Conversely, only 15% of subjects choose

[8]See table 2.3b.

the maximin principle while 51% of subjects chose maximin in the distribution choice. In both cases, the frequency of the distributions chosen across all subjects are much more in line with that of perceived descriptive norms than principles. While the contrast is not as stark for utilitarianism or inequality aversion, the same pattern can be observed. This aggregate pattern is reinforced by simple correlation coefficients between distribution choices and principles (-0.87) and between distribution choices and descriptive norms (0.99).

Table 2.4 plots the frequency of principle choice by chosen distribution for all three treatments. The drift to the maximin distribution for each chosen principle is evident. For all subjects, except for those who chose the meritocratic principle, maximin is the most frequently chosen distribution. Indeed, except for those who chose the maximin principle, the principle choice only predicts the distribution choices of 16-19% of subjects. For comparison, suppose a person chose their preferred principle and then randomly selected the distribution: i.e. the principle choice has no influence on the distribution decision. It follows holding principle 'x' would nevertheless 'correctly' predict distribution choices 25% of the time with such random behaviour. This means that for people who choose the inequality aversion, meritocracy and utilitarian principles in our experiment, their actual chosen distribution outcomes are no better predicted than they would be had those distribution outcome decisions actually been random. A chi-squared test of the distribution in table 2.4 is significant at the 99%-confidence level (chi-squared of 405.36, p=0.000) further supporting the finding that the two decisions are significantly different from each other.

Table 2.4: Social Preference by chosen Distribution

	Chosen Distribution			
Social Preference	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Inequality Aversion	**18.90%**	62.52%	5.83%	12.76%
Maximin	11.43%	**68.86%**	6.00%	13.71%
Meritocracy	6.60%	34.42%	**16.56%**	42.42%
Utilitarianism	21.84%	55.11%	6.61%	**16.43%**

Notes: Bold values indicate the share of respondents who chose a distribution consistent with their social preference. Percentages sum to 100% within each row.

Turning now to the individual-level data, table 2.5 reports individual binary logit regressions whereby the outcome variable is equal to 1 of subject i chose the respective distribution. Social preference and social norms are equally dummy variables equal to one if the principle chosen by subject i matched the distribution choice of the model and if the elicited descriptive or injunctive norm of subject i matched the distribution choice of the model, respectively.[9] As we are estimating multiple outcome variables, we also account for multiple hypothesis testing by reporting adjusted p-values based on Anderson (2008). These do not change the main findings. In table 2.3b in the appendix we reproduce this analysis using a combined regression equation for each distribution choice including all three dummies in the same model. This allows us to also introduce control variables. We run separate regressions in table 2.5 because given that personal principles are likely also influenced by injunctive social norms, it is difficult to distinguish their respective influences on the outcome variable in a combined regression model due to potential multicollinearity.[10]

The results reported in table 2.5 show that it always helps when predicting individual distribution decisions to know either a person's personal principle, their perceived injunctive social norm, or their perceived descriptive social norm. However, two considerations point to the primacy of the perceived descriptive social norms in this predictive task.

First, the coefficient of the social preference and injunctive social norm variables are negative for the utilitarian distribution decision. This suggests that it is not helpful to know whether a person's social preference or perceived injunctive norm is utilitarian because this means that person is less likely to also choose the utilitarian distribution. In contrast, the coefficients of the perceived descriptive social norm are always significant and positive, meaning that a person who perceives the descriptive social norm to be utilitarian is also *more likely* to choose the utilitarian distribution.

Second, the coefficients of the perceived descriptive social norm are always significantly larger

[9]In our Pre-Analysis plan, we indicated that we would weigh our results based on demographic variables to create representative samples of the respective regions we include in our experiment. We decided against doing so given the small sample size per country and the skewed age and income distribution outlined in table 2.3. Instead we control for all our demographic variables in our main analysis.

[10]We also decided against using multinomial logit models with categorical independent variables because we are interested in the relative importance of principle or norm x in explaining the choice of distribution x over all other possible options, as opposed to the relative importance of principle or norm x over principle or norm y in explaining distribution x.

Table 2.5: Logistic regressions of distributive choices for all treatments

	All Treatments				Non-Veil of Ignorance Treatment			
	Choice of Distribution				Choice of Distribution			
	Inequality Aversion	Maximin	Meritocracy	Utilitarianism	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Personal Principle	0.580***	0.839***	1.064***	-0.564***	0.368	0.469**	0.746***	-0.079
	(0.134)	(0.128)	(0.147)	(0.137)	(0.226)	(0.221)	(0.265)	(0.217)
	[0.001]	[0.001]	[0.001]	[0.001]	[0.074]	[0.054]	[0.021]	[0.218]
Injunctive Norm	0.338**	0.638***	0.755***	-0.335***	0.436*	0.067	0.763***	-0.423**
	(0.132)	(0.142)	(0.142)	(0.124)	(0.226)	(0.249)	(0.254)	(0.213)
	[0.006]	[0.001]	[0.001]	[0.005]	[0.056]	[0.245]	[0.013]	[0.056]
Descriptive Norm	2.528***	2.093***	2.064***	2.036***	2.416***	2.243***	1.684***	2.020***
	(0.141)	(0.100)	(0.164)	(0.111)	(0.240)	(0.180)	(0.291)	(0.196)
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Selfishness					0.100	-0.321**	0.467*	0.117
					(0.214)	(0.154)	(0.252)	(0.173)
					[0.472]	[0.147]	[0.147]	[0.472]
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
Session Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,219	2,219	2,219	2,219	733	733	733	733

Notes: Estimates come from individual logistic regressions. Personal Principle, Injunctive Norm, and Descriptive Norm are dummy variables equal to 1 if the subject's respective choice of principle or norm matched the distribution choice. Selfishness is a dummy variable equal to 1 if the subject chose the distribution that maximises the payoff of the quintile they were placed in based on their example quiz answers. Robust standard errors are presented in parentheses. Adjusted p-values for multiple hypothesis testing (Anderson 2008) are presented in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

than those of either the social preference or of the perceived injunctive social norm. This contrast becomes even more stark when focusing only on the non-veil of ignorance treatment, where selfish considerations can be entered as an additional motivation for distributive decisions. We can introduce the selfishness variable in this treatment because subjects know their likely quintile position. This is a dummy variable taking the value of 1 when the distribution choice also accords with a person's selfish interest (maximizes own expected material returns); 0 otherwise. Here, perceived descriptive norms always predict the distribution choice positively while, at best, social preferences and perceived injunctive norms only help predict the correct direction in half the distribution decisions. Selfishness itself is generally a poor predictor of distribution choices.

Taking meritocracy as an example, subjects who have a social preference for meritocracy are 2.6 times as likely to choose the meritocratic distribution compared to everyone else, while subjects who perceive the descriptive social norm to be meritocratic are 7.3 times as likely to choose the meritocratic distribution. These findings reinforce the descriptive evidence reported in figure 2.4.

Result 1: There are significant differences between subjects' chosen principle and distribution in the aggregate but principles have some predictive power for individual distribution choices.

While table 2.5 provides evidence that suggests descriptive social norms play a more significant role in explaining distributive choices than social preferences, it does not yet provide evidence on whether social norms can explain subjects' deviation away from their stated principle in the distribution choice. Table 2.6 reports that this is indeed the case, at least for descriptive norm-following. The outcome variable here is a dummy variable equal to 1 if subject i chose a distribution consistent with their previously stated social preference. As the table reports, choosing a distribution consistent with the perceived descriptive norm makes one less likely to choose a distribution consistent with the stated preference. On the other hand, choosing a distribution consistent with the perceived injunctive norm increases the likelihood of being consistent in preference-following substantially. This second result is

Table 2.6: **Principle-consistency by norm-following**

	Principle-Distribution Consistency			
	Descriptive Norm		Injunctive Norm	
	No Controls	Controls	No Controls	Controls
Norm-following	-0.211**	-0.188*	2.110***	2.129***
	(0.096)	(0.101)	(0.109)	(0.116)
Controls		✓		✓
Country Fixed Effects	✓	✓	✓	✓
Session Fixed Effects	✓	✓	✓	✓
Observations	2,408	2,219	2,408	2,219

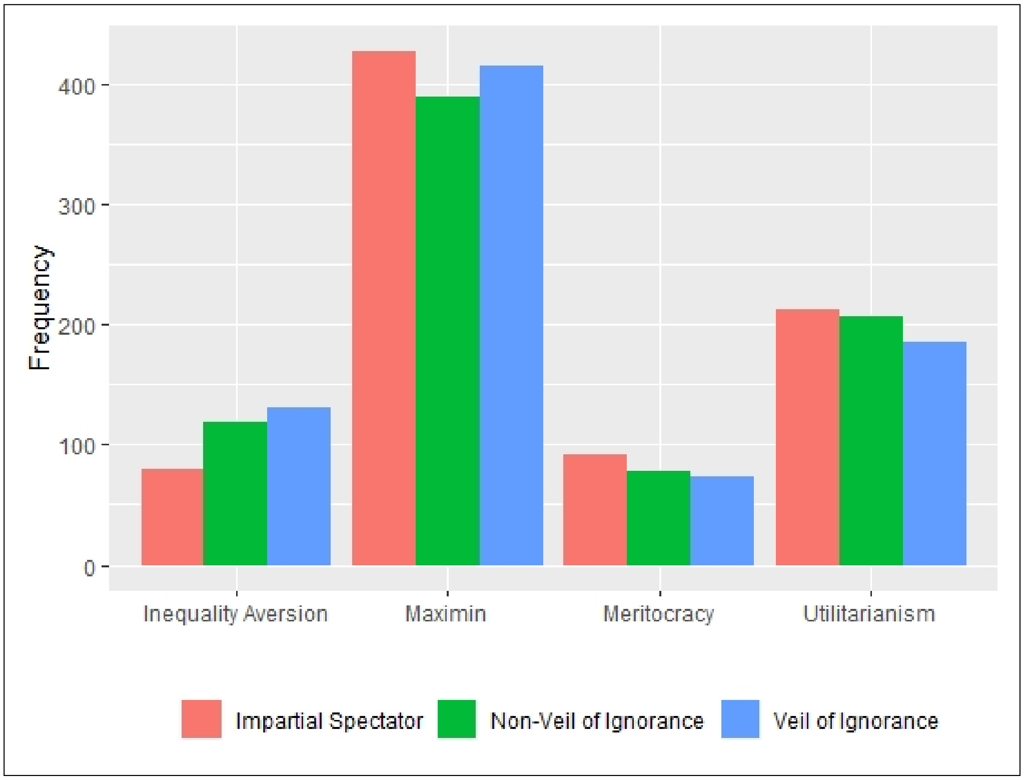
Notes: Estimates come from logistic regressions. Norm-following is a dummy variable equal to 1 if the subject's injunctive or descriptive social norm matched the chosen distribution. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

however most likely a mechanical finding due to the high correlation between personal principles and injunctive norms, which can be seen in figure 2.4 and is reported at the individual level in table 2.3a in the appendix. If a subject is consistent in their preference-following, they are therefore naturally also more likely to be injunctive norm-followers if principles and injunctive norms are correlated. These results provide evidence partially in line with H2:

Result 2: Descriptive but not injunctive norm-following can explain subjects' deviation away from their stated preference in the distribution choice.

While our results indicate that many subjects do not consistently follow their stated social preferences but deviate toward descriptive social norms when making distribution decisions, we also want to ensure that these results are robust to the elicitation method. We therefore compare subjects' decisions across our three treatments: The impartial spectator treatment in which subjects have no stake in the choices they make; the veil of ignorance treatment in which subjects have a stake in their decisions but do not yet know the position they will be placed on in the distribution; and finally, the non-veil of ignorance treatment in which subjects have a known stake in the decisions they make. Figure 2.5 gives the aggregate frequency of distribution choices by treatment. There is a significant difference in chosen distributions by treatments (Chi-squared of 18.63, p=0.05); however, this is driven by inequality aversion as the significance disappears when we exclude this choice from the analysis (Chi-square of

Figure 2.5: Distribution choice by Treatment



2.89, p=0.58). Inequality averse distribution decisions are significantly less frequent under the impartial spectator than veil and non-veil procedures.

Tables 2.7 gives the proportion of social preference followers and the proportion of perceived descriptive and injunctive social norm followers by treatment. There are no significant differences in these frequencies across the three treatments (Chi-squared for preference following is 1.12, p=0.57 and for descriptive norm following is 1.85, p=0.40).

Result 3: There is no significant difference in the level of social preference- or norm-following in the distribution decision across treatments.

Finally, we are interested in how the level of norm-following in subjects' choices affects their subsequent behaviour in a one-shot public goods game. Given the high correlation between injunctive norms and personal principles discussed previously, we exclude injunctive norms from this analysis. As we expect conditional cooperation to be the behaviour which connects

Table 2.7: Principle- and Norm-following by Treatments

	Treatments		
Personal Principle following	Impartial Spectator	Veil of Ignorance	Non-Veil of Ignorance
Inequality Aversion	18.78%	20.51%	17.62%
Maximin	76.24%	70.21%	60.19%
Meritocracy	19.21%	15.33%	14.53%
Utilitarianism	12.59%	13.33%	23.60%
Injunctive Norm following			
Inequality Aversion	14.94%	16.27%	17.52%
Maximin	69.05%	69.31%	50.62%
Meritocracy	16.41%	13.75%	14.93%
Utilitarianism	20.38%	20.75%	20.63%
Descriptive Norm following			
Inequality Aversion	42.45%	50.66%	47.79%
Maximin	75.85%	77.90%	76.08%
Meritocracy	33.65%	43.84%	29.55%
Utilitarianism	52.27%	53.30%	54.75%

norm-following in the distribution task with public goods game contributions, descriptive norm-following is the more relevant version of norm-following in this context anyway.

Table 2.8 reports subjects' contributions for all treatments and the non-veil of ignorance treatment only. We pooled the data from our main experiment and the two subsequent robustness check experiments which included the public goods game for this analysis. Results for only the main experiment and each robustness check can be found in appendix section IX.3. It is evident from the results that norm-following is a strongly significant and positive predictor of higher contribution levels. Being a descriptive norm-follower in the distribution choice increases public good contributions, ceteris paribus, by £0.73-£0.98 in all treatments and by £0.82-£1.09 in the non-veil of ignorance treatment only. Preference-following on the other hand does not help predict contribution levels.

One possible explanation for this result is that given the modal descriptive norm in the distribution task is maximin, the coefficient for norm-following might merely reflect a concern for ensuring some income for all that translates from the distribution experiment to the public goods game and is indirectly captured by norm-following. To account for this possibility, we include the perceived descriptive norm as a control variable in models (2) and (4) of table 2.8. While it is evident that the type of perceived descriptive norm affects contribution levels, the positive effect of norm-following generally on contribution levels remains.

Result 4: There is a significant positive relationship between norm-following in the distributive choice and subjects' contribution levels in the public goods game.

While accounting for selfishness by focusing only on the non-veil of ignorance treatment increases the significance and magnitude of the norm-following variable, selfishness itself does not explain contribution levels. We find, however, that a higher age, being female, a lower income, not being a student, more risk-seeking preferences, and a more left-wing political orientation are associated with higher contribution levels.

VI Motivation for Norm-following

Our main finding, that people deviate significantly from their stated social preference in line with what they perceive to be the descriptive social norm, raises the question of why that is the case. There are several possible explanations for why people may decide to follow a social norm. These can be broadly distinguished by whether they pose a challenge to the preference satisfying model or complement it. The economic approach to norm following falls into the latter category as it assumes norms provide focal points in games with multiple Nash Equilibria and are therefore useful coordination devices in groups (see Binmore 2010). This explanation is precluded by our design as there are no interactive decisions to be made in the distributive part of the experiment. It can therefore not explain why people adhere to their perceived social norm.

An alternative explanation for norm following that is equally not a challenge to the preference satisfying model is that injunctive social norms help constitute preferences (Tajfel et al.

Table 2.8: Public Goods Game

	One-Shot PG Game Contributions (0-20) for all waves			
	All Treatments		Non-Veil of Ignorance Treatment	
	(1)	(2)	(3)	(4)
Norm-following	0.975***	0.726***	1.086***	0.817**
	(0.218)	(0.230)	(0.320)	(0.333)
Preference-following	-0.027		0.198	
	(0.241)		(0.349)	
Descriptive Norm				
Utilitarianism		-0.924***		-1.586***
		(0.262)		(0.385)
Meritocracy		-1.348***		-1.335**
		(0.379)		(0.539)
Inequality Aversion		-0.398		-0.835*
		(0.309)		(0.444)
Selfishness			0.501	0.256
			(0.310)	(0.310)
Treatments				
Veil of Ignorance	0.088	0.081		
	(0.311)	(0.312)		
Non-Veil of Ignorance	0.047	0.057		
	(0.316)	(0.316)		
Sample				
United Kingdom	-0.179	-0.183	-0.432	-0.425
	(0.267)	(0.267)	(0.370)	(0.369)
United States	-0.184	-0.194	0.111	0.097
	(0.367)	(0.366)	(0.559)	(0.554)
Quiz Performance	0.224**	0.205**	0.170	0.146
	(0.100)	(0.100)	(0.143)	(0.143)
Income	-0.262**	-0.243**	-0.280*	-0.250
	(0.103)	(0.103)	(0.155)	(0.155)
Female	0.927***	0.923***	0.971***	0.961***
	(0.231)	(0.231)	(0.340)	(0.339)
Left-Right	0.495***	0.479***	0.379*	0.357*
	(0.147)	(0.146)	(0.209)	(0.210)
Age	0.459***	0.446***	0.570***	0.547***
	(0.099)	(0.098)	(0.147)	(0.147)
Risk seeking	0.137***	0.148***	0.117	0.137*
	(0.051)	(0.051)	(0.075)	(0.074)
Student	-0.480*	-0.482*	-0.774*	-0.768*
	(0.277)	(0.276)	(0.398)	(0.396)
Economics	-0.289	-0.271	-0.248	-0.201
	(0.272)	(0.272)	(0.383)	(0.382)
Constant	9.449***	10.124***	9.750***	10.900***
	(0.852)	(0.866)	(1.245)	(1.265)
Session Fixed Effects	✓	✓	✓	✓
Observations	3,376	3,376	1,619	1,619
R-squared	0.030	0.036	0.038	0.050

Notes: Estimates come from individual linear regressions. The outcome variable is subjects' contribution to the hypothetical one-shot public goods game with values ranging from 0-20. Norm-following is equal to 1 if the subject followed the perceived descriptive norm in their distribution choice and 0 otherwise. Preference-following is equal to 1 if the subject followed their principle in the distribution choice and 0 otherwise. The reference category for the descriptive norm variable is maximin, which was the modal perceived descriptive norm. Selfishness is a dummy variable equal to 1 if the subject chose the distribution that maximises the payoff of the quintile they were placed in based on their example quiz answer. The reference category for the treatment variables is the impartial spectator treatment. The reference category for the Sample variable is Western Europe. Quiz performance ranges from 0 to 5 depending on how many questions the subject answered correctly. A higher value on the left-right variable indicates a more left-wing orientation on economic policy. Risk preferences are self-reported on a scale from 0 to 10 with 10 being the most risk-seeking option. Student is a dummy variable equal to 1 if the subject is currently studying towards a degree and Economics is a dummy variable equal to 1 if the subject has ever studied a course on economics at university. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

1979; Akerlof and Kranton 2000). This explanation is based on social identification theory which argues that people gain a sense of identity by behaving in a way that corresponds to the norms of their group. This gives them a sense of identity because their group's norms differ from those of other groups.

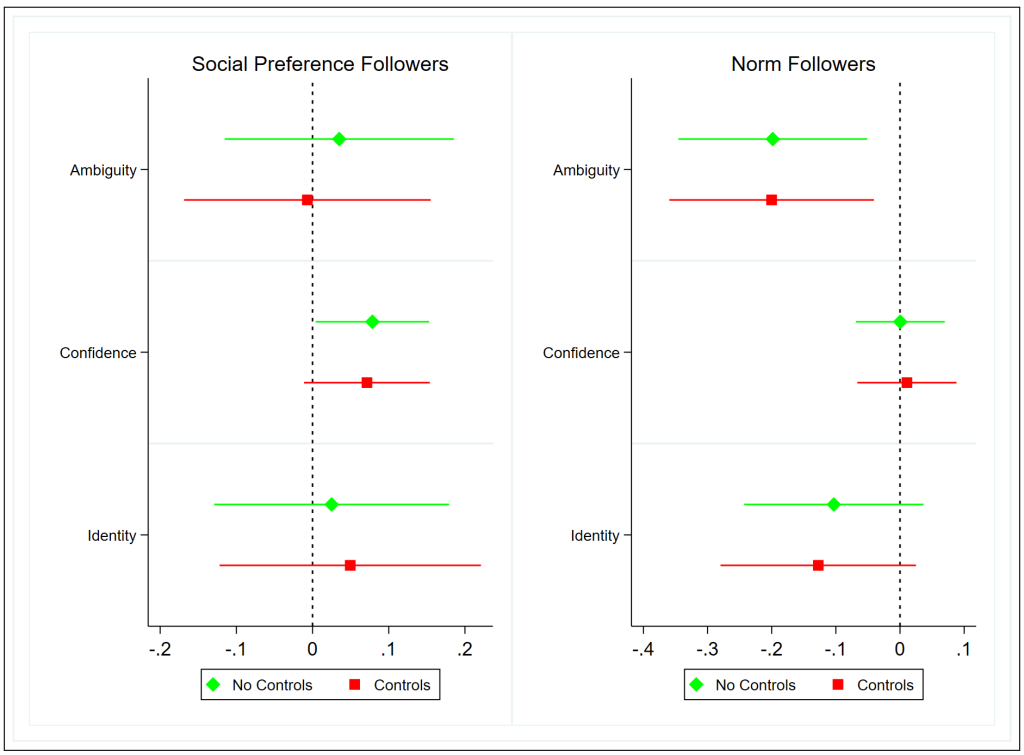
There are two further potential explanations for descriptive norm following that form distinct alternatives to the preference satisfying model. The first is that people may simply not have relevant preferences in the given decision problem. The decision may be so novel or unfamiliar that individuals cannot evaluate the different options and therefore treat other people's behaviour, or expected behaviour, as social information regarding how to value them. There is some experimental evidence in support of this explanation (Fatas et al. 2018).

The second potential explanation of this kind is based on a problem set out by Adam Smith in his *Theory of Moral Sentiments* (1759). He argues that while we have social preferences, it is difficult to be sure whether we authentically act on those social preferences or whether we are actually guided by our own self-interest. To genuinely adhere to social preferences, we therefore need some external standard to act on social preferences authentically. This is what descriptive social norms supply and why there are followed.

In a follow-up experiment we therefore included a number of additional measures to test which of these different explanations best explains descriptive norm-following in the distributive decision. The experiment was run on the 21st of April 2020 with a total of 1,003 subjects. After decision 1, we added a question to the design asking subjects to rate their confidence in their chosen principle on a scale from 0 to 10. In so far as either of the two explanations of norm-following that are distinct from the preference satisfying model are correct, subjects who are less confident in their principle should be more likely to follow a social norm and deviate from their stated principle. We further added questions on ambiguity aversion, social identification (both on strength and type of social identification) and tolerance for deception to the demographic part of the experiment. Both, ambiguity aversion and a higher tolerance for deception would be expected to correlate with norm-following if the authenticity-based explanation of norm-following were to hold. Social identification however would point towards the first explanation of norm-following as complimentary to the preference satisfying model.

As can be seen in figure 2.6, we find that those who followed their social preference in the distribution choice expressed a higher level of confidence in their principle compared to all other subjects. This is consistent with both, the explanation of norm-following based on a lack of defined preferences and a desire for authenticity. As the second part of figure 2.6 shows, higher ambiguity aversion is associated with a higher likelihood of norm-following, which further supports the norm-following as authenticity explanation.

Figure 2.6: Individual Characteristics by Subject Group



Notes: Figures are based on logistic regressions. The outcome variable of the left coefficient plot is equal to 1 if the subject followed their social preference in the distribution choice and 0 otherwise. The outcome variable of the coefficient plot on the right is equal to 1 if the subject followed their perceived social norm and 0 otherwise. Ambiguity ranges from 0 to 7 (with a higher value indicating more ambiguity seeking preferences) and is a standardized scale based on the ambiguity preference survey module developed by Cavatorta and Schröder (2019). Confidence is measured as the subject's response to the question "On a scale from 1 to 10, please rate how confident you are in the choice you just made." which was asked directly after subjects chose a principle. A higher value indicates more confidence. Identity ranges from 1 to 4 with a higher value indicating a higher level of identity. This variable was measured using the module developed by Kuo and Margalit (2012).

Neither, one's degree of social identification nor tolerance for deception, explains people's likelihood to follow either social preferences or norms. We do, however, find a strong and significant relationship between identifying with one's race and being less likely to follow the perceived social norm in the distributive task.[11]

VII Robustness checks

To test the robustness of our results we conducted a series of additional checks. Our immediate concern was to test for the possibility that, by asking subjects to identify their perceived descriptive social norm immediately after they made the actual distribution decision, we might have rendered the distribution choice especially salient to the subjects when eliciting the social norm. Thus, we inverted decision 2 and 3 in our robustness check experiment. This makes decision 1 over the principle of justice immediately precede the social norm perception question which now comes before the actual distribution decision. We find the same patterns. In fact, we find less consistency between preferences and distribution choices and even more norm-following (see appendix section IX.2.1).

To better understand our main results, we also test whether following one's social preference in the distributive task could be predicted by any of our demographic control variables. As table 2.9 reports, there is no demographic variable that consistently predicts preference following. In our main two waves of the experiment, having a more left-wing political orientation is significantly and positively associated with preference following; however, we cannot replicate this finding in our subsequent robustness checks. Equally being a student is positively and significantly associated with preference following in our main robustness check but we can also not replicate this finding in the main experiment or second robustness check. These findings suggest that preference following is not consistently associated with any underlying demographic patterns.

An additional concern is that the complexity of the distributive choice made some subjects choose at random if they did not fully understand the decision problem. We therefore created a dummy variable capturing whether subjects included any of a select list of words and

[11]These results can be found in table 2.3g in appendix IX.3.4.

Table 2.9: Social Preference following

	Social Preference Following		
	Main Experiment	Motivation Test	Average Income Test
Sample			
United Kingdom	-0.013	0.132	-0.185
	(0.135)	(0.196)	(0.389)
United States	-0.031	0.468	-0.093
	(0.166)	(0.358)	(0.438)
Quiz Performance	-0.016	0.021	-0.086
	(0.048)	(0.075)	(0.149)
Income	0.012	0.057	0.177
	(0.048)	(0.083)	(0.119)
Female	0.124	-0.249	-0.105
	(0.111)	(0.180)	(0.317)
Left-Right	0.208***	0.018	0.151
	(0.072)	(0.102)	(0.220)
Age	-0.019	0.062	-0.256*
	(0.045)	(0.089)	(0.154)
Risk seeking	-0.016	-0.013	0.030
	(0.024)	(0.038)	(0.077)
Student	-0.026	0.465**	0.134
	(0.135)	(0.202)	(0.366)
Economics	0.113	-0.093	-0.108
	(0.125)	(0.202)	(0.359)
Constant	-1.861***	-1.515**	-1.360
	(0.385)	(0.586)	(1.108)
Session Fixed Effects	✓		
Observations	2,219	886	271
R-squared	0.322	0.376	0.677

Notes: Estimates come from a logistic regression. The outcome variable Social Preference Following is a dummy variable equal to 1 if the subject's choice of principle matched the distribution. The reference category for the sample variables is Western Europe. Quiz performance ranges from 0 to 5 depending on how many questions the subject answered correctly. A higher value on the left-right variable indicates a more left-wing orientation on economic policy. Risk preferences are self-reported on a scale from 0 to 10 with 10 being the most risk-seeking option. Student is a dummy variable equal to 1 if the subject is currently studying towards a degree and Economics is a dummy variable equal to 1 if the subject has ever studied a course on economics at university. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

terms indicating confusion in their feedback section.[12] We then tested whether comprehension could be predicted by any of our main variables of interest. These results are reported in table 2.10. We find that subjects who chose either the utilitarian or the inequality averse principle as opposed to the meritocratic principle were more likely to indicate confusion

[12]The list of words and terms included: 'confusing', 'confusion', 'confused', 'complicated', 'unsure', 'unclear', 'ambiguous', 'didn't understand', 'did not understand'.

in their feedback section. We further find that subjects who chose the inequality averse distribution as opposed to the meritocratic distribution (which may be viewed as the default option here) were less likely to indicate confusion. We further find some evidence that subjects in the United States were more likely to indicate confusion compared to European subjects. Encouragingly, we find no evidence that either principle- or norm-following predicts comprehension.

A further robustness check relates to the key assumption that we make with respect to individuals using principles of justice when thinking about how to make distribution decisions. In particular, this is crucial in making the connection between individual's chosen principle of justice and their likely social preferences. At the end of the experiment we asked our subjects in an open commentary box to explain how they decided on their distribution option. Table 2.11 lists the most frequently used terms by chosen distribution.

The most used words differ substantially for each distribution choice and, importantly, match the wording of our principle options. This is particularly striking when comparing the terms used to justify the inequality averse and maximin distributions with the meritocratic and utilitarian distributions. In short, the currency that people use to explain their decisions is the same as that of the principles of justice, even though, as we have seen, they are not typically guided by the principle they selected.

Another possible qualification to our conclusion might be that our subjects are guided by more than one justice principle and it is possible that a different secondary principle of justice was triggered when the actual distribution choices were presented in decision 3. We therefore ran a further robustness check with 200 subjects where we asked them after they had identified the principle of justice they thought should be applied (decision 1) if they had a secondary justice principle, and if so, what it was. Just over half (56%) had a secondary principle and of those who did, maximin was again the least chosen (secondary) principle (see appendix IX.3.5. Less than 9% of the 200 subjects identified maximin as their secondary principle and so the possible contribution of a secondary principle in explaining the wholesale shift to maximin in the distribution decision 3 is at best relatively modest even if all these 9% had been guided by their secondary principle alone. Recall in the original experiment

Table 2.10: Comprehension Test

	Comprehension			
	Principle Distribution Choice		Principle- Norm-following	
Principle				
Utilitarianism	0.632***	0.561**		
	(0.242)	(0.252)		
Maximin	0.395	0.184		
	(0.292)	(0.315)		
Inequality Aversion	0.839***	0.647***		
	(0.226)	(0.242)		
Distribution				
Utilitarianism	-0.447	-0.278		
	(0.305)	(0.329)		
Maximin	-0.349	-0.194		
	(0.270)	(0.292)		
Inequality Aversion	-0.944**	-0.854**		
	(0.364)	(0.393)		
Preference-following			0.138	0.090
			(0.192)	(0.206)
Norm-following			0.137	0.256
			(0.177)	(0.193)
Treatments				
Veil of Ignorance	0.081	0.014	0.082	0.003
	(0.212)	(0.229)	(0.212)	(0.230)
Non-Veil of Ignorance	0.132	0.182	0.151	0.187
	(0.210)	(0.226)	(0.209)	(0.227)
Sample				
United Kingdom	0.336	0.423	0.326	0.453
	(0.231)	(0.283)	(0.229)	(0.284)
United States	0.502*	0.535*	0.475*	0.558*
	(0.269)	(0.310)	(0.267)	(0.311)
Constant	-3.205***	-2.808***	-3.266***	-2.943***
	(0.362)	(0.764)	(0.278)	(0.701)
Individual Controls		✓		✓
Session Fixed Effects	✓	✓	✓	✓
Observations	2,408	2,219	2,408	2,219

Notes: Estimates come from a logistic regression. The outcome variable is equal to 1 if subjects mentioned words indicating confusion or misunderstanding in the feedback section. Meritocracy is the reference category for the principle and distribution variables. Preference-following is equal to 1 if the subject followed their social preference in the distribution choice and 0 otherwise. Norm-following is equal to 1 if the subject followed the perceived social norm in their distribution choice and 0 otherwise. The reference category for the treatment variables is the impartial spectator treatment. The reference category for the sample variables is Western Europe. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Table 2.11: Terms most Frequently used to Justify chosen Distribution

	Inequality Aversion				Maximin		
	Total Frequency	Documents	Relative		Total Frequency	Documents	Relative
Equal distribution	5	5	0.011	Hard work	13	13	0.011
Income inequality	4	4	0.009	Equal distribution	12	12	0.010
Basic income	3	3	0.007	Fair distribution	10	10	0.008
Equal amount	3	2	0.004	Greater good	8	8	0.006
Distribute wealth	3	3	0.007	Income inequality	8	8	0.006
Shared equally	2	2	0.004	Many people	7	5	0.004
Best choice	2	2	0.004	make sure	7	7	0.006
Fair distribution	2	2	0.004	Income distribution	7	7	0.006
Observations	447	447	447		1,231	1,231	1,231
	Meritocracy				Utilitarianism		
	Total Frequency	Documents	Relative		Total Frequency	Documents	Relative
Work hard	5	5	0.018	Work hard	12	9	0.015
Felt right	3	3	0.011	Work harder	6	5	0.008
Work harder	3	3	0.011	Hard work	6	6	0.010
worked harder	2	2	0.007	Paid based	4	4	0.007
Hard work	2	2	0.007	Seemed fair	4	4	0.007
Next group	2	1	0.004	Felt right	3	3	0.005
Make sure	2	2	0.007	Worked hard	3	3	0.005
Second game	1	1	0.004	Worth taking	3	3	0.005
Observations	271	271	271		603	603	603

Notes: The table reports the most frequently used terms used by respondents to justify their chosen distribution. Total frequency reports the number of times a term was used overall within the subgroup of respondents who chose a particular distribution. Documents reports the number of responses of individual respondents in which a term was used at least once. Relative reports the proportion of responses within the distribution-dependent subgroup that refer to the given term.

14% identified maximin as their principle and 50% chose the maximin distribution.: even another 9% leaves a big gap.

A final possible qualification that we consider is that, although each principle in decision 1 does identify one of the four distribution outcomes, subjects might have made an execution error when translating their personal principle into an actual distribution decision. Random 'trembling' would, however, introduce 'noise' and weaken the principle-distribution consistency (as it might any norm-distribution consistency); it would not explain why the distribution decisions are actually skewed to the maximin distribution. For this to occur there has to be some reason for supposing that 'errors' are easier to make in the maximin direction because maximin is 'closer' to each of the principles than is any of the others. We test for this possibility by asking another 200 subjects to choose a principle (i.e. decision 1) and then we ask them to identify the distribution (in decision 3) that they associate with their chosen principle. Those who incorrectly identify their chosen principle's distribution do on average err noticeably in the direction of two distribution outcomes: 44% go to utilitarianism and 41% go to maximin. Most (82%) of the trembles to maximin were accounted for by those who identified their chosen principle as inequality aversion, so we re-ran the individual regression in table 2.5 excluding all the subjects who chose the inequality aversion principle in decision 1. The perceived descriptive social norm is still a more important predictor of these remaining subjects' distribution choices than is their chosen principle (see appendix IX.3.6. So, while 'skewed' trembling might explain why those who chose inequality aversion migrated to the maximin distribution, it does not explain why this occurs for subjects that select the other principles (and they are the majority in our sample). Indeed, the errors among the subjects choosing meritocracy (our modal principle choice) were skewed away from maximin (only 8% of their mistakes went to maximin).[13]

Of course, one further explanation of this result may seem possible and so should be touched upon. The questions asked in decisions 3 and 4 both concern the choice of an actual distribution of income, whereas question 1 refers to the choice of a justice principle. Perhaps, therefore, it is not so surprising that decision 4 better predicts decision 3 than does decision 1, given the shared object of decisions in 3 and 4. However, a descriptive norm cannot be defined in a way that is different to that of actual choices and unless social preferences are to be revealed tautologically (and so unfalsifiably) by actual decisions, social preferences cannot be identified through actual choices. Thus, this difference in the object of decision is built into the very competition between the two accounts of why people might behave unselfishly. It is not some artefact of our experimental design; it is integral to a serious test. Indeed, the fact, that decision 3 refers to actual distributions and so does decision 4 on perceived descriptive social norms, does not mean that the one should help predict the other. But they do in our experiment. Nor, incidentally, does the fact, that decision 1 deals in a choice of justice principles as does decision 2 on the perceived injunctive norms, mean that injunctive norms should predict personal principles. But they do. In short, the influence of social

[13]It is perhaps also worth noting that the trembling rate was over 50%: that is only 45% correctly identified the distribution outcome associated with their chosen principle. Again, this suggests that the majority of our subjects were not used to thinking in terms of principles of justice; and if this is the case, it would be difficult for the majority of our subjects to be said to have social preferences that they consult when decision making in this instance.

norms seems not to be limited to that of the descriptive kind, powerful as they appear to be.

VIII Discussion and Conclusion

People frequently behave prosocial in situations where they can reduce their own income to help others. Contrary to the assumptions made by previous studies and by revealed preference theory, we find that such behaviour cannot simply be assumed to reflect underlying social preferences but may instead reflect perceived descriptive norms. In a series of online experiments, we elicited subjects' social preferences and perceived social norms and compared these to decisions made in a simple distributive game. We find that subjects make distributive decisions that deviate significantly from their stated preferences in line with perceived descriptive social norms. Interestingly, the elicitation mechanism used makes surprisingly little difference to either the principles and distributions chosen or the preference-following behaviour of subjects. This is however less surprising given that expected differences in choices made under the different elicitation mechanisms require the preference satisfying model to hold. Finally, we find that descriptive norm-following behaviour in the distributive task is associated with higher contributions in a subsequent one-shot public goods game.

There are several respects in which the behaviour of our subjects is reassuringly consistent with other experimental findings. For example, we find in table 2.3b in the appendix that being trained in economics is a powerful predictor of choosing the utilitarian distribution but not any of the other distributions; and we know, for example, from Fehr et al. (2006) that economics students are more inclined to be influenced by efficiency considerations than non-economics students in such distribution decisions. Likewise, it is known that US subjects hold more meritocratic beliefs than European subjects (see Alesina and Glaeser 2004) and we too find in table 2.3b that the only predictable difference from nationality is that being a US citizen somewhat increases the probability of selecting the meritocratic distribution. Being to the right on a typical left-right political question regarding the role of government in the economy helps predict the utilitarian distribution; whereas being on the left helps predict maximin. This is in line with the common finding that a left-leaning political orientation is associated with a preference for more redistribution (see Alesina and Giuliano 2011). Again, being risk-averse helps predict maximin, as would be expected. Finally, our evidence on the influence of social norms is consistent with what has been found in other studies (e.g. Krupka and Weber 2013; Kimbrough and Vostroknutov 2016).

Our results are important for three main reasons. First, distributive choices form a large part of economic and political decision-making and understanding why people support one economic policy over another is of significance to policymakers and social scientists alike. Our main finding suggests that such distributive preferences do not always reveal underlying social preferences but, more often, reflect perceived social norms. This result is important for any welfare analysis of distributive policies, as it suggests that the use of the Pareto principle has a weak foundation. In particular, it cannot be assumed that prosocial behaviour in this setting reveals social preferences which can then be entered into a social welfare function for the purposes of developing policy recommendations. This, in turn, means that the foundations of welfare economics need reworking to take account of norm-following. This is non-trivial because we have an experiment where the influence of social preferences is carefully distinguished from that of social norms.

Second, our follow-up experiments provide some inside into why people may be guided by social norms as opposed to preferences in distributive decisions. There is evidence that it arises from an epistemic problem with respect to what preferences to act upon. Those who lack confidence in their chosen principle and who are ambiguity averse are inclined to follow their perceived descriptive norm.

Finally, it may be possible to draw some useful substantive insights with respect to the character of prosocial behaviour from this experiment. Some care is required because we only have four actual distributions and had the option set been different, then there might have been different choices. Furthermore, the character of the prosocial behaviour that is revealed may depend on the initial distribution of income that we have assumed. Nevertheless, the average EU actual top 20%/bottom 20% ratio for disposable income is very close to the 5.5 we have assumed (see Eurostat 2018). So, in this respect, the decision problem captures something close to the current post tax relativities and may be relevant to the contemporary discussion regarding how further intervention might be required to alter the income distribution. For example, both the IMF (Ostry et al. 2014) and OECD (OECD 2015) have argued that a move to greater equality would in current circumstances help to boost productivity growth. In this context, our experiment suggests that the majority reveal support for policies that improved the position of the bottom 20%.

Bibliography

George A Akerlof and Rachel E Kranton. Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753, 2000.

Alberto Alesina and George-Marios Angeletos. Fairness and redistribution. *American Economic Review*, 95(4):960–980, 2005.

Alberto Alesina and Paola Giuliano. Preferences for redistribution. In *Handbook of Social Economics*, volume 1, pages 93–131. Elsevier, 2011.

Alberto Alesina and Edward Ludwig Glaeser. *Fighting poverty in the US and Europe: A world of difference*. Oxford University Press, 2004.

Ingvild Almås, Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden. Fairness and the development of inequality acceptance. *Science*, 328(5982):1176–1178, 2010.

Michael L Anderson. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495, 2008.

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

Maria Bigoni and Sigrid Suetens. Feedback and dynamics in public good experiments. *Journal of Economic Behavior & Organization*, 82(1):86–95, 2012.

Ken Binmore. Social norms or social preferences? *Mind & Society*, 9(2):139–157, 2010.

Mariana Blanco, Dirk Engelmann, and Hans Theo Normann. A within-subject analysis of other-regarding preferences. *Games and Economic Behavior*, 72(2):321–338, 2011.

Gary E Bolton and Axel Ockenfels. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: comment. *American Economic Review*, 96(5):1906–1911, 2006.

Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden. Responsibility for what? fairness and individual responsibility. *European Economic Review*, 54(3):429–441, 2010.

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

Alexander W Cappelen, Karl O Moene, Erik Ø Sørensen, and Bertil Tungodden. Needs versus entitlements—an international fairness experiment. *Journal of the European Economic Association*, 11(3):574–598, 2013b.

Fredrik Carlsson, Dinky Daruvala, and Olof Johansson-Stenman. Are people inequality-averse, or just risk-averse? *Economica*, 72(287):375–396, 2005.

Elisa Cavatorta and David Schröder. Measuring ambiguity preferences: A new ambiguity preference survey module. *Journal of Risk and Uncertainty*, 58(1):71–100, 2019.

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

Robert B Cialdini, 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*, 58(6):1015, 1990.

Giovanna d'Adda, Michalis Drouvelis, and Daniele Nosenzo. Norm elicitation in within-subject designs: Testing for order effects. *Journal of Behavioral and Experimental Economics*, 62:1–7, 2016.

Raymond Duch, Denise Laroze, Thomas Robinson, and Pablo Beramendi. Multi-modes for detecting experimental measurement error. *Political Analysis*, 28(2):263–283, 2020.

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

Tore Ellingsen, Magnus Johannesson, Johanna Mollerstrom, and Sara Munkhammar. Social framing effects: Preferences or beliefs? *Games and Economic Behavior*, 76(1):117–130, 2012.

Dirk Engelmann and Martin Strobel. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review*, 94(4):857–869, 2004.

Justin Esarey, Tim Salmon, and Charles Barrilleaux. Social insurance and income redistribution in a laboratory experiment. *Political Research Quarterly*, 65(3):685–698, 2012.

Eurostat. Income inequality in the eu. Accessed: 20/03/2020, `https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20180426-1`, 2018.

Peer Eyal, Rothschild David, Gordon Andrew, Evernden Zak, and Damer Ekaterina. Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, pages 1–20, 2021.

Enrique Fatas, Shaun P Hargreaves Heap, and David Rojo Arjona. Preference conformism: An experiment. *European Economic Review*, 105:71–82, 2018.

Ernst Fehr and Urs Fischbacher. Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2):63–87, 2004.

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

Ernst Fehr, Urs Fischbacher, and Simon Gächter. Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, 13:1–25, 2002.

Ernst Fehr, Michael Naef, and Klaus M Schmidt. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment. *American Economic Review*, 96(5):1912–1917, 2006.

Urs Fischbacher, Simon Gächter, and Ernst Fehr. Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, 71(3):397–404, 2001.

Urs Fischbacher, Simon Gächter, and Simone Quercia. The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology*, 33(4):897–913, 2012.

Raymond Fisman, Ilyana Kuziemko, and Silvia Vannutelli. Distributional preferences in larger groups: Keeping up with the joneses and keeping track of the tails. *Journal of the European Economic Association*, 19(2):1407–1438, 2021.

Axel Franzen and Sonja Pointner. Anonymity in the dictator game revisited. *Journal of Economic Behavior & Organization*, 81(1):74–81, 2012.

Simon Gächter, Daniele Nosenzo, and Martin Sefton. Peer effects in pro-social behavior: Social norms or social preferences? *Journal of the European Economic Association*, 11(3): 548–573, 2013.

Francesco Guala, Luigi Mittone, and Matteo Ploner. Group membership, team preferences, and expectations. *Journal of Economic Behavior & Organization*, 86:183–190, 2013.

John C Harsanyi. Can the maximin principle serve as a basis for morality? a critique of john rawls's theory. In *Essays on Ethics, Social Behavior, and Scientific Explanation*, pages 37–63. Springer, 1980.

Erik O Kimbrough and Alexander Vostroknutov. Norms make preferences social. *Journal of the European Economic Association*, 14(3):608–638, 2016.

Michał Krawczyk. A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. *Journal of Public Economics*, 94(1-2):131–141, 2010.

Erin L Krupka and Roberto A Weber. Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524, 2013.

Alexander Kuo and Yotam Margalit. Measuring individual identity: Experimental evidence. *Comparative Politics*, 44(4):459–479, 2012.

Ilyana Kuziemko, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva. How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, 105(4):1478–1508, 2015.

Lars J Lefgren, David P Sims, and Olga B Stoddard. Effort, luck, and voting for redistribution. *Journal of Public Economics*, 143:89–97, 2016.

Steven D Levitt and John A List. What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2):153–174, 2007.

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

Tim Mulgan. *Understanding utilitarianism.* Routledge, 2014.

Elias Naumann, Christopher Buss, and Johannes Bähr. How unemployment experience affects support for the welfare state: a real panel approach. *European Sociological Review*, 32(1):81–92, 2016.

OECD. Growth and income inequality: trends and policy implications. *OECD Economics Department Policy Notes*, (26), 2015.

Mr Jonathan David Ostry, Mr Andrew Berg, and Mr Charalambos G Tsangarides. *Redistribution, inequality, and growth.* International Monetary Fund, 2014.

Lindsay A Owens and David S Pedulla. Material welfare and changing political preferences: The case of support for redistributive social policies. *Social Forces*, 92(3):1087–1113, 2014.

Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. Beyond the turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70:153–163, 2017.

John Rawls. *A theory of justice.* Harvard university press, 1971.

Philipp Rehm. Citizen support for the welfare state: Determinants of preferences for income redistribution. *WZB Markets and Political Economy Working Paper No. SP II*, 2, 2005.

Paul A Samuelson. A note on the pure theory of consumer's behaviour. *Economica*, 5(17): 61–71, 1938.

Hannah Schildberg-Hörisch. Is the veil of ignorance only a concept about risk? an experiment. *Journal of Public Economics*, 94(11-12):1062–1066, 2010.

Adam Smith. The theory of moral sentiments, ed. *A. Millar, A. Kincaid & J. Bell. AM Kelley.(Originally published in 1759)*, 1759.

Henri Tajfel, John C Turner, William G Austin, and Stephen Worchel. An integrative theory of intergroup conflict. *Organizational Identity: A Reader*, 56:65, 1979.

Björn Vollan and Elinor Ostrom. Cooperation and the commons. *Science*, 330(6006):923–924, 2010.

Nina Weber. Pre-analysis plan: Principle and outcomes - a test of revealed social preferences. `https://doi.org/10.7910/DVN/DA7JKB`, 2019.

IX Appendix

IX.1 Materials and Methods

IX.1.1 Overview

We conducted our online experiment using Qualtrics for the design of the study and Prolific Academic for the recruitment of participants. Prolific Academic is a web-based panel with about 300,000 participants as of October 2021. Participants on Prolific have been found to pay significantly more attention and provide responses of higher quality than those registered on mTurk (Peer et al. 2017; Eyal et al. 2021).

Our main experiment was conducted on the 14th of November and the 9th of December 2019. The average completion time was 8 minutes and 17 seconds and respondents earned on average £1.55 for their participation. The full survey instrument that we used is available in Section IX.5 of this appendix and the original online survey can be accessed here.

IX.1.2 Sampling and Survey Implementation

We conducted a total of two main waves of the experiment, as well as seven additional waves for robustness checks. Table 2.1a provides an overview of all waves. We focused our online

Table 2.1a: Overview of individual waves

	Date	Sample Size	Avg. Time	Returned	Timed Out
First Wave	14/11/2019	1,205	8.11mins	27	16
Second Wage	09/12/2019	1,203	7.45mins	32	25
Average Income Test	30/03/2020	294	14.05mins	59	15
Social Norm Test	30/03/2020	302	11.00mins	36	3
Motivation Test	21/04/2020	1,003	15.08mins	67	37
Second Principle Test	19/11/2020	201	3.48mins	4	2
Distribution Test	25/11/2020	200	4.37mins	5	1
Wording Test	08/10/2021	222	4.14mins	12	1
Order Test	08/10/2021	218	9.52mins	21	1

experiment on participants from the US, UK and the following Western European countries: Belgium, Denmark, Finland, France, Germany, Italy, Luxembourg, Netherlands, Norway, Sweden and Spain. Table 2.1b lists the number of respondents from each geographical area by individual wave. To ensure that we reached respondents from all geographical areas, all waves were ran in the late afternoon GMT time. Our samples are not representative of individual countries.

Table 2.1b: Sample composition of individual waves

	United Kingdom	United States	Western Europe	Total Sample Size
First Wave	768	165	272	1,205
Second Wave	623	280	300	1,203
Average Income Test	153	65	76	294
Social Norm Test	180	48	72	302
Motivation Test	561	48	392	1,003
Second Principle Test	18	120	63	201
Distribution Test	84	9	107	200
Wording Test	89	38	95	222
Order Test	76	42	100	218

IX.1.3 Survey Structure

IX.1.3.1. Basic Set up

Introduction

Subjects are asked for their consent to participate in the study and reminded to read the questions very carefully and answer honestly.

Experimental Part

Using Qualtrics' *Randomizer*, subjects are randomly and evenly allocated to one of three treatments for the following four decisions.

Decision 1. Identify guiding principle of justice.

Decision 2. Incentivised guess of what decision most people made in decision 1.

Decision 3. Select distribution.

Decision 4. Incentivised guess of what decision most people made in decision 3.

Quiz

Demographic Questions

IX.1.3.2. Treatments

Different institutional mechanisms for eliciting justice principles and making distribution decisions (each encoding a different idea over how best to identify what is just).

Treatment 1: *Impartial Spectator.* Decision 1-4 undertaken as an impartial spectator.

Treatment 2: *Veil of Ignorance.* Decision 1-4 undertaken behind a veil of ignorance.

Treatment 3: *Non-veil of Ignorance.* Decision 1-4 undertaken knowing one's own likely position in the distribution.

IX.1.3.3. Robustness Check 1: Average Income Test

In the main two waves of the experiment we referred to "Total" income per distribution choice. We therefore conducted a robustness check where we replaced "Total" with "Average" in all displays of our distribution options.

IX.1.3.4. Robustness Check 2: Social Norm Test

The Krupka and Weber (2013) method uses a separate subject pool to elicit the social norm for a particular decision problem. Our main experiment uses the same subject pool for norm elicitation and so we conducted an additional norm elicitation experiment with a separate subject pool. This experiment only consisted of decision 4 of the experimental part outlined in section IX.1.3.

IX.1.3.5. Robustness Check 3: Motivation Test

Our main robustness check was designed to test the motivations behind norm following and included the following elements in addition to the main experiment:

- **Ambiguity preference elicitation.** We followed the method developed by Cavatorta and Schröder (2019) to measure subjects' ambiguity preferences.
- **Confidence in principle.** After subjects made decision 1, they were asked to rate their confidence in the chosen principle: *On a scale from 1 to 10, please rate how confident you are in the choice you just made.*
- **Identity elicitation.** Following Kuo and Margalit (2012) we asked respondents the following two additional questions in the demographics section:

    1. Some people describe themselves by their nationality, their ethnicity, their race, their religion, or their occupation. How about you? Do you identify first and foremost by:
        - Your nationality
        - Your ethnicity
        - Your race
        - Your religion
        - Your occupation
        - Other (Please specify)
    2. Consider your response to the previous question. How strong would you say your attachment is to the identity you chose? Would you say your attachment is:
        - Not strong at all
        - Slightly strong
        - Somewhat strong
        - Very strong

- **Self-deception elicitation.** To elicit subjects' level of self-deception we asked the following two additional questions in the demographics section:

    1. It has been argued that there will always be occasions when the kindest thing to do is lie. But, on the other hand, if people lie, then who can you believe? Do you agree it is okay to lie sometimes?

  - Scale ranges from 1 (Strongly agree) to 7 (Strongly disagree)

2. There is a big debate in psychology over whether deception in experiments should be permitted. What do you think?
  - Scale ranges from 1 (Never) to 7 (Whenever it helps science)

We further reversed the order of decision 3 and 4 in this robustness check to test whether people simply chose the same distribution option in decision 4 that they chose in decision 3, for example, to appear consistent. The results in section IX.3.4 confirm that this was not the case. This robustness check also only included the impartial spectator treatment as we did not find significant treatment effects in our main waves.

IX.1.3.6. Robustness Check 4: Second Principle Test

To test for the possibility that our subjects have two principles that they take into consideration when making the distribution choice we conducted a further robustness check asking subjects first, whether they had another principle they agreed with and second, which of the other principles it is.

IX.1.3.7. Robustness Check 5: Distribution Test

To ensure that subjects understood which distribution option corresponded to which justice principle we conducted a robustness check asking subjects to identify the distribution corresponding to their chosen principle. This decision was incentivised. If subjects correctly identified the corresponding distribution they received a bonus payment of 50p.

IX.1.3.8. Robustness Check 6: Wording Test

As pointed out by one referee, the wording of our principle statements is not structured in an entirely consistent manner which could have affected subjects' likelihood to choose one principle over another. To test for this possibility, we conducted a robustness check with an alternative wording of the inequality aversion and maximin statements. We also repeated the distribution test introduced in robustness check 5 to check whether subjects are more or less likely to correctly identify the distribution corresponding to their chosen principle given this alternative wording. The wording used in this test is as follows:

Maximin: Income should be distributed to improve the position of the least well-off group in society.

Inequality Aversion: Income should be distributed to reduce inequality by minimizing average differences in income.

IX.1.3.9. Robustness Check 7: Order Test

While we already reversed the order of decisions 3 and 4 in robustness check 2, we added a seventh robustness check to reverse the order of decisions 1 & 2 and 3 & 4. This allows us to test whether making the distribution decision first affects either the chosen distribution and principle, preference consistency, or norm-following.

Table 2.2a: Summary Statistics of Demographics by Wave

	Main Experiment	Average Income Test	Social Norm Test	Motivation Test	Second Principle Test	Distribution Test	Wording Test	Order Test
Demographics (%)								
Female	60.10	49.32	56.61	60.10	47.96	52.53	55.07	48.10
Age								
18-20	9.58	13.65	14.67	14.34	15.58	18.09	10.96	14.49
21-29	35.47	41.98	36.00	43.03	48.24	47.74	40.64	42.99
30-39	28.85	24.91	24.33	25.68	21.11	19.10	28.31	26.64
40-49	13.61	12.63	13.00	10.63	8.54	10.55	14.16	11.21
50-59	8.45	4.78	9.00	5.12	5.53	2.01	3.65	2.34
60+	4.04	2.05	3.00	1.20	1.01	2.51	2.28	2.34
Students	24.92	27.55	29.33	31.70	38.31	34.50	33.78	35.94
Economics	21.47	29.33	21.67	21.38	27.00	26.00	21.62	27.19
Income								
Under £20,000	51.69	50.36	51.60	53.76	58.15	46.84	36.63	38.05
£20,000 to £34,999	25.74	23.36	25.98	27.21	23.37	30.38	33.17	31.22
£35,000 to £44,999	11.69	11.68	10.32	12.17	11.96	17.72	13.86	15.12
£50,000 to £74,999	6.65	7.66	7.12	4.87	3.80	2.53	9.41	9.76
£75,000 to £99,999	2.05	2.19	2.85	1.00	1.09	2.53	4.46	3.90
Over £100,000	2.19	4.74	2.14	1.00	1.63	0.00	2.48	1.95
Sample								
United Kingdom	57.77	52.04	60.00	56.04	8.96	42.00	40.09	34.86
United States	18.48	22.11	16.00	4.80	59.70	4.50	17.12	19.27
Europe	23.75	25.85	24.00	39.16	31.34	53.50	42.79	45.87
Observations	2,408	294	302	1,003	201	200	222	218

IX.2 Additional Descriptive Results

Table 2.2a reports summary statistics of all waves of the study. Our sample is clearly skewed towards younger respondents on low income. Over 50% of our sample has an annual income below £20,000. Except for the Average Income Test, our sample is also predominantly female.

IX.2.1 Distribution of Main Variables

Figures 2.2a and 2.2b report the distribution of respondents' personal principle, injunctive social norm, descriptive social norm, and chosen distribution for the average income and motivation test, respectively. Both distributions show a strikingly similar pattern. Meritocracy is the most chosen personal principle, yet maximin is by far the most chosen distribution and perceived descriptive social norm. In both distributions it is also evident that distribution choices are more closely aligned with perceived descriptive social norms than personal principles.

Figures 2.2c and 2.2d report the distribution of respondents' personal principle, injunctive social norm, descriptive social norm, and chosen distribution for the wording and order tests, respectively. Here, maximin is again the most chosen distribution and meritocracy the modal personal principle in both tests. While maximin is also the most chosen perceived descriptive social norm in the wording test, this is not the case in the order test. Here, utilitarianism is, in fact, the modal perceived descriptive social norm. Importantly however, the difference between the number of respondents who chose maximin and those who chose utilitarianism as their perceived descriptive social norm is only seven out of 218, suggesting that this finding, which is inconsistent compared to all other robustness checks, might be due to sampling.

Figure 2.2a: Distribution of Principle, Distribution Choice, and Norms in Average Income Test

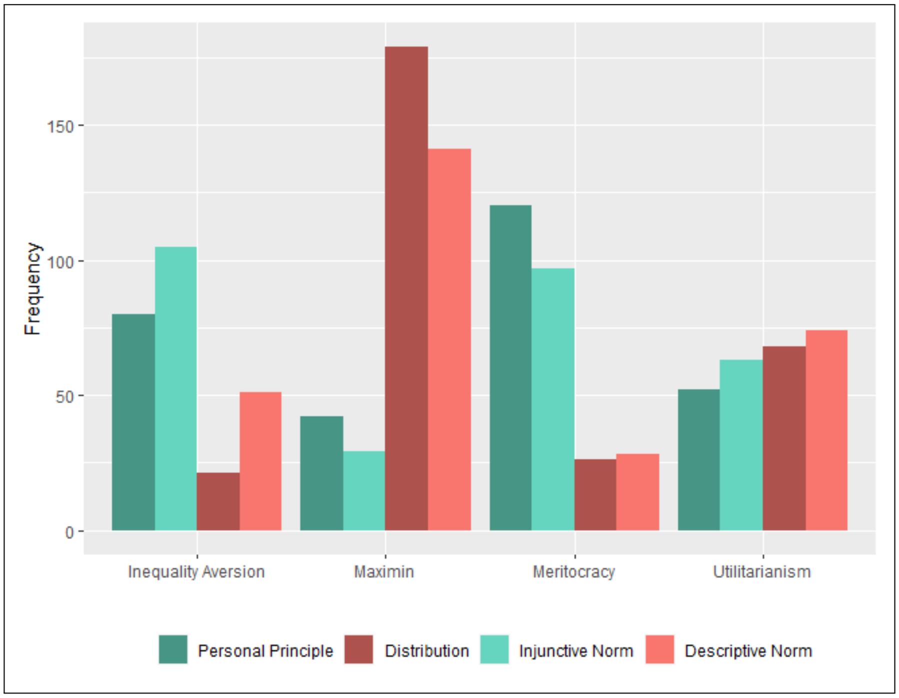


Figure 2.2b: Distribution of Principle, Distribution Choice, and Norms in Motivation Test

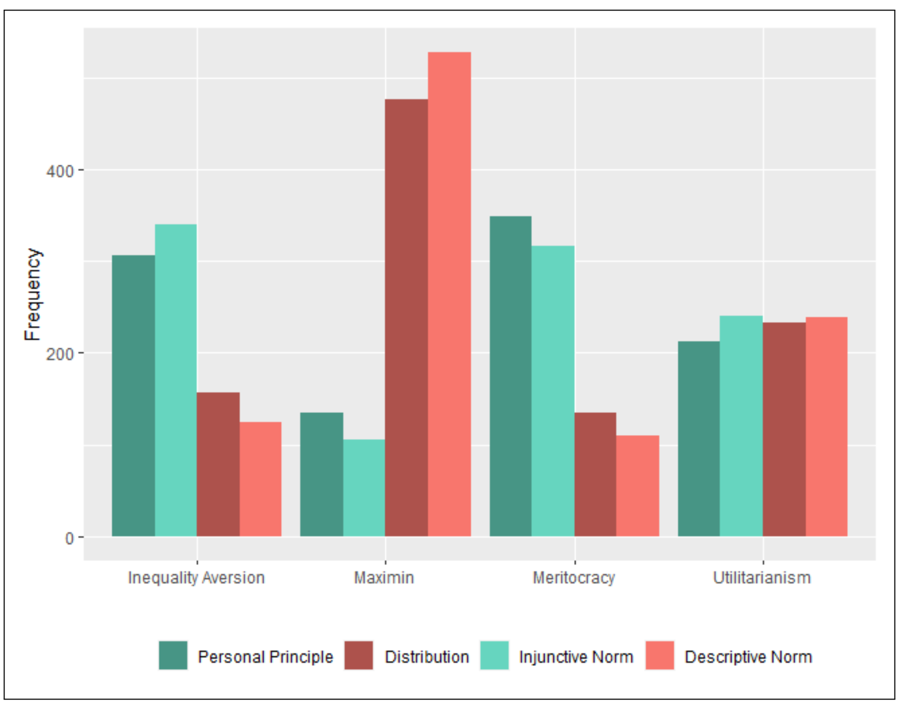


Figure 2.2c: Distribution of Principle, Distribution Choice, and Norms in Wording Test

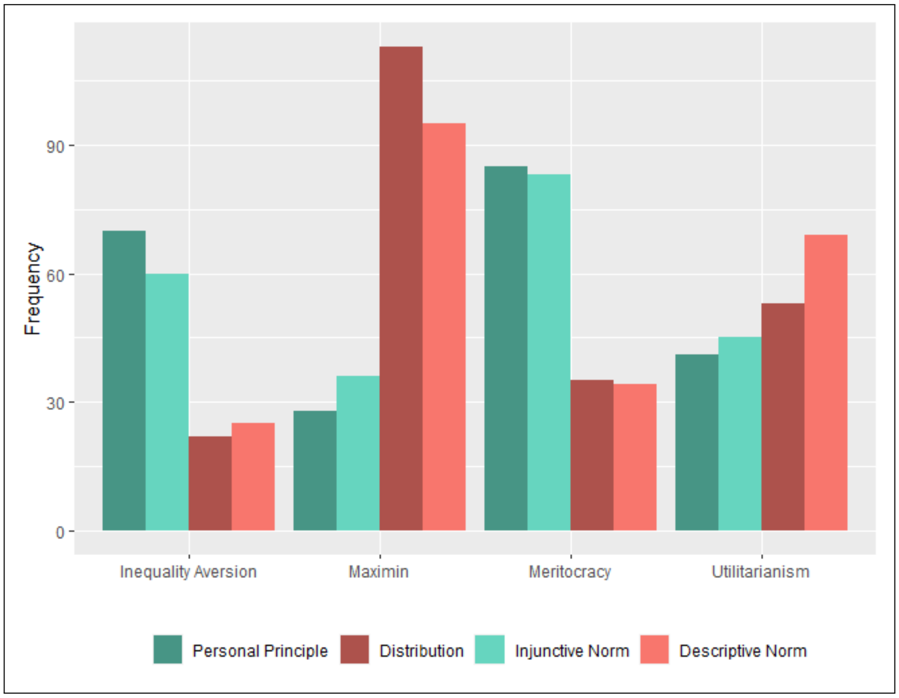
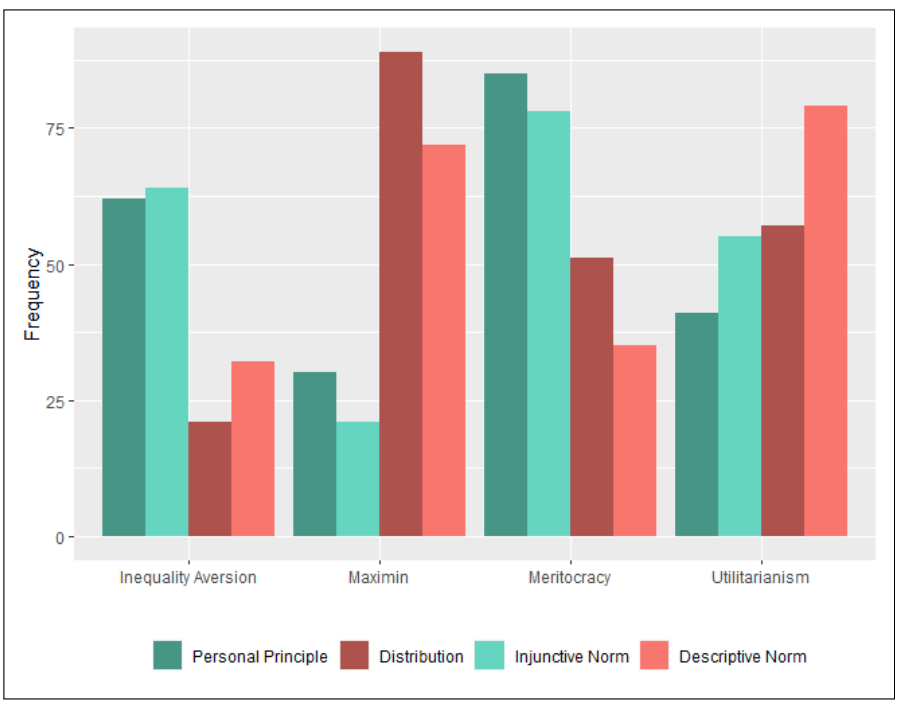


Figure 2.2d: Distribution of Principle, Distribution Choice, and Norms in Order Test



IX.3 Additional Results

IX.3.1 Main Experiment

Table 2.3b reports logistic regressions similar to table 2.5 in the main text; however, each column now corresponds to a regression model including personal principle, perceived descriptive norm, and perceived injunctive norm dummies combined. This allows us to now also report coefficients for our control variables. Our main result, that perceived descriptive norms are the best predictor of distribution choices, holds to this alternative specification. The significance of the injunctive norm coefficients is however reduced compared to the results reported in table 2.5. This is likely due to the fact that injunctive norms also help predict personal principle choices leading to multicollinearity in the combined regression models.

IX.3.1.1. Injunctive Norm and Personal Principle

Table 2.3a reports the results of simple logistic regressions where a person's injunctive norm is used as a predictor of the personal principle. As is evident from the table, there is a strong and highly significant relationship between injunctive norms and personal principles for all four principle options.

Table 2.3a: Logistic regressions of personal principles for all treatments

	Social preference			
	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Injunctive Norm	1.781***	2.160***	1.703***	1.849***
	(0.108)	(0.152)	(0.100)	(0.116)
Constant	-1.652***	-2.346***	-0.460	-3.023***
	(0.415)	(0.553)	(0.368)	(0.437)
Controls	✓	✓	✓	✓
Country Fixed Effects	✓	✓	✓	✓
Session Fixed Effects	✓	✓	✓	✓
Observations	2,219	2,219	2,219	2,219
Pseudo R-squared	0.134	0.124	0.143	0.124

Notes: Estimates come from a logistic regression. Injunctive Norm is a dummy variable equal to 1 if the subject's perceived social norm in the principle choice matched the chosen principle. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Table 2.3b: Logistic regressions of distributive choices for all treatments - complete models

	All Treatments				Non-Veil of Ignorance Treatment			
	Choice of Distribution				Choice of Distribution			
	Inequality Aversion	Maximin	Meritocracy	Utilitarianism	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Personal Principle	0.503***	0.774***	0.837***	-0.684***	0.366	0.641**	0.335	-0.066
	(0.167)	(0.156)	(0.165)	(0.159)	(0.296)	(0.286)	(0.315)	(0.254)
Injunctive Norm	0.043	0.118	0.344**	0.106	0.225	-0.342	0.564*	-0.166
	(0.164)	(0.175)	(0.161)	(0.148)	(0.301)	(0.330)	(0.300)	(0.245)
Descriptive Norm	2.513***	2.081***	1.965***	2.062***	2.431***	2.284***	1.617***	2.004***
	(0.142)	(0.101)	(0.167)	(0.113)	(0.243)	(0.183)	(0.296)	(0.196)
Selfishness					0.227	-0.490***	0.600**	0.063
					(0.236)	(0.178)	(0.254)	(0.187)
Treatments								
Veil of Ignorance	0.611***	-0.062	0.038	-0.202				(0.175)
(0.120)	(0.179)	(0.136)				*Non-Veil of Ignorance*	0.538***	-0.156
-0.073	-0.013					(0.178)	(0.120)	(0.186)
(0.137)				Sample				
			United Kingdom	-0.067	0.061	0.326	-0.122	-0.237
0.199	-0.088	-0.002	(0.185)	(0.132)	(0.220)	(0.144)	(0.328)	(0.247)
(0.370)	(0.238) *United States*	-0.188	-0.082	0.469*	-0.015	-0.325	-0.143	0.770*
-0.132								
	(0.247)	(0.166)	(0.247)	(0.175)	(0.426)	(0.308)	(0.403)	(0.300)
Quiz Performance	-0.082	0.037	-0.121	0.038	-0.008	-0.085	0.081	0.040
	(0.069)	(0.047)	(0.078)	(0.055)	(0.110)	(0.083)	(0.135)	(0.095)
Income	0.020	-0.021	0.056	-0.011	0.061	-0.026	-0.074	0.067
	(0.070)	(0.046)	(0.066)	(0.052)	(0.110)	(0.080)	(0.112)	(0.086)
Female	0.240	0.100	0.243	-0.364***	0.563**	-0.285	0.672**	-0.412**
	(0.155)	(0.109)	(0.173)	(0.121)	(0.268)	(0.204)	(0.340)	(0.208)
Left-Right	0.087	0.248***	-0.094	-0.287***	-0.107	0.379***	-0.085	-0.327***
	(0.095)	(0.067)	(0.096)	(0.074)	(0.159)	(0.118)	(0.173)	(0.122)
Age	0.066	-0.017	0.021	-0.051	0.125	-0.091	0.280**	-0.154
	(0.065)	(0.046)	(0.071)	(0.052)	(0.118)	(0.086)	(0.130)	(0.098)
Risk seeking	0.082**	-0.064***	0.012	0.033	0.120**	-0.117***	0.035	0.069
	(0.033)	(0.023)	(0.033)	(0.026)	(0.058)	(0.041)	(0.056)	(0.046)
Student	-0.005	0.091	0.234	-0.186	0.142	0.113	0.852**	-0.611**
	(0.199)	(0.138)	(0.201)	(0.151)	(0.342)	(0.253)	(0.350)	(0.262)
Economics	-0.181	-0.245*	0.049	0.345**	-0.409	-0.230	0.098	0.384*
	(0.180)	(0.129)	(0.189)	(0.137)	(0.294)	(0.223)	(0.318)	(0.228)
Constant	-4.037***	-1.707***	-3.104***	-0.373	-3.529***	-1.175*	-4.713***	-0.329
	(0.550)	(0.381)	(0.604)	(0.426)	(0.922)	(0.650)	(0.141)	(0.746)
Session Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,219	2,219	2,219	2,219	733	733	733	733

Notes: Estimates come from a logistic regression. Personal Principle, Injunctive Norm, and Descriptive Norm are dummy variables equal to 1 if the subject's respective choice of principle or norm matched the distribution choice. Selfishness is a dummy variable equal to 1 if the subject chose the distribution that maximises the payoff of the quintile they were placed in based on their example quiz answers. The reference category for the treatment variables is the impartial spectator treatment. The reference category for the sample variables is Western Europe. Quiz performance ranges from 0 to 5 depending on how many questions the subject answered correctly. A higher value on the left-right variable indicates a more left-wing orientation on economic policy. Risk preferences are self-reported on a scale from 0 to 10 with 10 being the most risk-seeking option. Student is a dummy variable equal to 1 if the subject is currently studying towards a degree and Economics is a dummy variable equal to 1 if the subject has ever studied a course on Economics at University. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

IX.3.1.2. Public Goods Game

Table 2.3c reports the results of the public goods game analysis for only the main experiment. These results are directly comparable to table 2.8 in the main text. In line with the pooled data of table 2.8, we find that descriptive norm-following significantly increases contributions in the one-shot public goods game. The magnitude of the norm-following variable is now even larger than in the pooled date. Being a descriptive norm-follower in the distribution choice increases public good contributions, ceteris paribus, by £0.96-£1.14 in all treatments and by £1.17-£1.51 in the non-veil of ignorance treatment only. As in the pooled data, preference-following on the other hand does not help predict contribution levels. We again find selfishness to not be a significant predictor of contributions and our demographic variables very show similar patterns to the pooled analysis.

IX.3.2 Average Income Test

IX.3.2.1. Preference- and Norm-following by chosen Distribution

Table 2.3d reports the chosen distribution by personal principle, perceived injunctive norm, and perceived descriptive norm for respondents in the Average Income Test. The pattern visible in table 2.3d is similar to the results of the main experiment: Descriptive social norms are more closely related to distribution choices than personal principles or perceived injunctive norms, except for respondents who chose the maximin distribution. For all subjects, maximin is again the most frequently chosen distribution. Except for those who chose the maximin principle, the principle choice only predicts the distribution choices of 11-19% of subjects which is no better than had the distribution outcome decisions been random. The chi-squared test of the distribution for the principle in table 2.3d is significant at the 99%-confidence level (chi-squared of 37.32, p=0.000) indicating that the two decisions are significantly different from each other.

IX.3.2.2. Main results

Table 2.3e reports the results of logistic regressions with individual distribution choices as the outcome variables for respondents in the Average Income Test. This test was conducted

Table 2.3c: Public Goods Game

	One-Shot PG Game Contributions (0-20)			
	All Treatments		Non-Veil of Ignorance Treatment	
	(1)	(2)	(3)	(4)
Norm-following	1.143***	0.964***	1.508***	1.173**
	(0.268)	(0.284)	(0.467)	(0.504)
Preference-following	0.065		0.358	
	(0.297)		(0.531)	
Descriptive Norm				
Utilitarianism		-0.453		-1.482***
		(0.316)		(0.567)
Meritocracy		-0.930**		-1.184
		(0.461)		(0.790)
Inequality Aversion		-0.397		-1.031
		(0.386)		(0.665)
Selfishness			-0.025	-0.022
			(0.466)	(0.465)
Treatments				
Veil of Ignorance	0.082	0.082		
	(0.311)	(0.312)		
Non-Veil of Ignorance	0.069	0.078		
	(0.317)	(0.317)		
Sample				
United Kingdom	0.157	0.149	0.083	0.071
	(0.348)	(0.348)	(0.612)	(0.614)
United States	0.132	0.128	0.839	0.782
	(0.433)	(0.433)	(0.731)	(0.735)
Quiz Performance	0.176	0.165	0.150	0.142
	(0.123)	(0.123)	(0.217)	(0.216)
Income	-0.321**	-0.311**	-0.569***	-0.549**
	(0.124)	(0.124)	(0.216)	(0.215)
Female	0.842***	0.842***	0.654	0.647
	(0.284)	(0.284)	(0.503)	(0.495)
Left-Right	0.575***	0.566***	0.632**	0.631**
	(0.179)	(0.178)	(0.306)	(0.306)
Age	0.463***	0.452***	0.714***	0.700***
	(0.117)	(0.117)	(0.210)	(0.208)
Risk seeking	0.180***	0.181***	0.165	0.171
	(0.061)	(0.061)	(0.105)	(0.104)
Student	-0.288	-0.295	-0.263	-0.291
	(0.350)	(0.350)	(0.645)	(0.642)
Economics	-0.316	-0.313	-0.526	-0.491
	(0.334)	(0.335)	(0.551)	(0.551)
Constant	8.730***	9.215***	8.269***	9.362***
	(1.015)	(1.036)	(1.768)	(1.810)
Session Fixed Effects	✓	✓	✓	✓
Observations	2,219	2,219	733	733
R-squared	0.036	0.039	0.058	0.068

Notes: Estimates come from individual linear regressions. The outcome variable is subjects' contribution to the hypothetical one-shot public goods game with values ranging from 0-20. Norm-following is equal to 1 if the subject followed the perceived descriptive norm in their distribution choice and 0 otherwise. Preference-following is equal to 1 if the subject followed their principle in the distribution choice and 0 otherwise. The reference category for the descriptive norm variable is maximin, which was the modal perceived descriptive norm. Selfishness is a dummy variable equal to 1 if the subject chose the distribution that maximises the payoff of the quintile they were placed in based on their example quiz answer. The reference category for the treatment variables is the impartial spectator treatment. The reference category for the Sample variable is Western Europe. Quiz performance ranges from 0 to 5 depending on how many questions the subject answered correctly. A higher value on the left-right variable indicates a more left-wing orientation on economic policy. Risk preferences are self-reported on a scale from 0 to 10 with 10 being the most risk-seeking option. Student is a dummy variable equal to 1 if the subject is currently studying towards a degree and Economics is a dummy variable equal to 1 if the subject has ever studied a course on economics at university. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Table 2.3d: Personal Principle and Norms by chosen Distribution

	Chosen Distribution			
Personal Principle	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Inequality Aversion	**11.25%**	75.00%	5.00%	8.75%
Maximin	7.14%	**73.81%**	2.38%	16.67%
Meritocracy	4.17%	45.00%	**14.17%**	36.67%
Utilitarianism	7.69%	65.38%	7.69%	**19.23%**
Injunctive Norm				
Inequality Aversion	**8.57%**	67.62%	5.71%	18.10%
Maximin	6.90%	**58.62%**	20.69%	13.79%
Meritocracy	5.15%	52.58%	**10.31%**	31.96%
Utilitarianism	7.94%	63.49%	6.35%	**22.22%**
Descriptive Norm				
Inequality Aversion	**31.37%**	56.86%	1.96%	9.80%
Maximin	3.55%	**78.72%**	4.96%	12.77%
Meritocracy	0.00%	35.71%	**32.14%**	32.14%
Utilitarianism	0.00%	39.19%	12.16%	**48.65%**

Table 2.3e: Logistic regressions of distributive choices for Average Income Test

	Impartial Spectator Treatment			
	Choice of Distribution			
	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Personal Principle	0.485	0.647*	1.278***	-0.333
	(0.500)	(0.382)	(0.486)	(0.406)
Injunctive Norm	0.169	-0.030	0.312	0.006
	(0.489)	(0.417)	(0.465)	(0.352)
Descriptive Norm	3.648***	1.799***	2.093***	2.039***
	(0.761)	(0.302)	(0.574)	(0.325)
Individual Controls	✓	✓	✓	✓
Session Fixed Effects	✓	✓	✓	✓
Observations	271	271	271	271

Notes: Estimates come from individual logistic regressions. Personal Principle, Injunctive Norm, and Descriptive Norm are dummy variables equal to 1 if the subject's respective choice of principle or norm matched the distribution choice. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

with only the impartial spectator treatment. These regression results are directly comparable to table 2.5 in the main text. Despite the small sample size of this robustness check, descriptive social norms are a highly significant predictor of distribution choices while personal principles only matter for the distribution choices of respondents who chose the meritocratic distribution. Injunctive social norms do not matter at all for distribution choices in those specifications. The descriptive social norm coefficients are similar in magnitude to those of the main regression results.

IX.3.2.3. Public Goods Game

Table 2.3f reports the results of the public goods game analysis for only the average income test. These results are directly comparable to table 2.8 in the main text and table 2.3c of the appendix. In this test, descriptive norm-following is not significant. The lack of significance is likely due to the much smaller sample size of this test compared to the main test and the pooled analysis.

IX.3.3 Social Norm Test

IX.3.3.1. Distribution of perceived Descriptive Social Norm

Figure 2.3a reports the frequency of the perceived descriptive social norms of subjects in the Social Norm Test. Unlike in all our other waves, utilitarianism is the modal choice while maximin is the second most-frequent choice. As this distribution is strikingly different to all other waves of the experiment, it suggests that the respondents make a substantially different choice when asked to decide on the appropriate social norm for a separate group of subjects (as proposed by Krupka and Weber 2013) than when the decision is made on the same subject group.

IX.3.4 Motivation Test

IX.3.4.1. Motivation by Subject Group

Table 2.3g reports individual characteristics for respondents who followed their personal principle and those who followed their perceived descriptive norm in the distribution choice. While confidence in the chosen principle increases preference-following, more ambiguity aversion (a lower ambiguity preference score) is associated with descriptive norm-following. Interestingly, identifying with one's own race significantly decreases the likelihood of following one's perceived descriptive norm social norm.

Table 2.3f: Public Goods Game

	One-Shot PG Game Contributions (0-20)	
	Average Income Test	
	(1)	(2)
Norm-following	0.240	-0.254
	(0.791)	(0.862)
Preference-following	-0.736	
	(0.918)	
Descriptive Norm		
Utilitarianism		-2.061*
		(1.060)
Meritocracy		-4.164***
		(1.530)
Inequality Aversion		0.859
		(1.060)
Constant	9.305***	10.272***
	(2.797)	(2.765)
Controls	✓	✓
Observations	271	271
R-squared	0.046	0.097

Notes: Estimates come from individual linear regressions. The outcome variable is subjects' contribution to the hypothetical one-shot public goods game with values ranging from 0-20. Norm-following is equal to 1 if the subject followed the perceived descriptive norm in their distribution choice and 0 otherwise. Preference-following is equal to 1 if the subject followed their principle in the distribution choice and 0 otherwise. The reference category for the descriptive norm variable is maximin, which was the modal perceived descriptive norm. Selfishness is a dummy variable equal to 1 if the subject chose the distribution that maximises the payoff of the quintile they were placed in based on their example quiz answer. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Figure 2.3a: Distribution of perceived Descriptive Social Norm

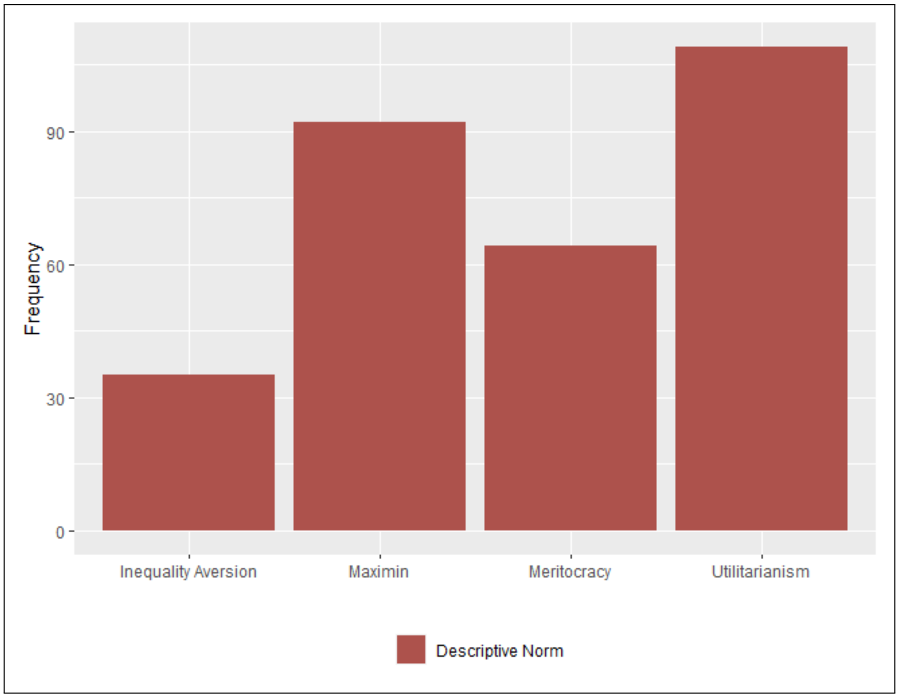


Table 2.3h reports individual predictors of respondents' confidence in their chosen principle. A stronger social identity is thereby associated with a higher level of confidence in one's chosen principle.

IX.3.4.2. Preference- and Norm-following by chosen Distribution

Table 2.3i reports the chosen distribution by personal principle, perceived injunctive norm, and perceived descriptive norm for respondents in the Motivation Test. The pattern visible in table 2.3i is again similar to the results of the main experiment: Descriptive social norms are more closely related to distribution choices than personal principles, except for respondents who chose the maximin distribution. The proportion of respondents who chose the distribution that matches their perceived descriptive social norm is somewhat larger than the proportion of respondents in the Average Income Test (see table 2.3d).

For all subjects, except for those who chose the meritocratic principle, maximin is the most frequently chosen distribution. Except for those who chose the maximin principle, the principle choice predicts the distribution choices of 15-19% of subjects which is no better than had

Table 2.3g: Logistic regressions of individual characteristics by subject group

	Non-Veil of Ignorance Treatment			
	Subject Group			
	Personal Principle Followers		Descriptive Norm Followers	
Ambiguity preference	0.035	-0.007	-0.198***	-0.200**
	(0.077)	(0.083)	(0.075)	(0.081)
Confidence	0.079**	0.071*	0.000	0.011
	(0.038)	(0.042)	(0.035)	(0.039)
Identity	0.025	0.049	-0.103	-0.127
	(0.079)	(0.088)	(0.071)	(0.078)
Identity group				
Ethnicity	0.478	0.747*	-0.677*	-0.619
	(0.387)	(0.428)	(0.347)	(0.390)
Nationality	0.441	0.630	-0.487*	-0.530*
	(0.301)	(0.339)	(0.267)	(0.288)
Occupation	0.367	0.487	-0.479*	-0.503
	(0.318)	(0.356)	(0.285)	(0.308)
Race	0.339	0.796	-1.352***	-1.269**
	(0.524)	(0.564)	(0.457)	(0.542)
Religion	0.580	0.653	-0.035	0.077
	(0.505)	(0.548)	(0.475)	(0.502)
Self-deception 1	0.004	0.005	-0.050	-0.028
	(0.049)	(0.053)	(0.043)	(0.048)
Self-deception 2	0.019	-0.020	-0.023	0.013
	(0.043)	(0.046)	(0.038)	(0.042)
Constant	-2.344***	-2.474***	2.300***	2.651***
	(0.619)	(0.892)	(0.569)	(0.798)
Individual Controls		✓		✓
Observations	971	859	971	859
Pseudo R-squared	0.006	0.020	0.017	0.041

Notes: Estimates come from a logistic regression. The outcome variable 'Personal Principle Followers' is equal to 1 if the subject followed their personal principle in the distribution choice and 0 otherwise. The outcome variable "Descriptive Norm Followers" is equal to 1 if the subject followed the perceived descriptive social norm in their distribution choice and 0 otherwise. Ambiguity preference ranges from 0 to 7 (with a higher value indicating more ambiguity seeking preferences) and is a standardized scale based on the ambiguity preference survey module developed by Cavatorta and Schröder (2019). Confidence is measured from 1 to 10 and a higher value indicates more confidence in the chosen principle. Identity ranges from 1 to 4 with a higher value indicating a higher level of identity. This variable was measured using the module developed by Kuo and Margalit (2012). 'Other' is the reference group for identity group. Self-deception 1 ranges from 1 to 7 with a lower value indicating more self-deception. Self-deception 2 ranges from 1 to 7 with a higher value indicating more tolerance for deception. *** p<0.01, ** p<0.05 , * p<0.1.

the distribution outcome decisions been random. The chi-squared test of the distribution for the personal principle in table 2.3d is significant at the 99%-confidence level (chi-squared of 99.40, p=0.000) indicating that the two decisions are significantly different from each other.

IX.3.4.3. Main result

Table 2.3j reports the results of logistic regressions with individual distribution choices as the outcome variables for respondents in the Motivation Test. This test was conducted with only the non-veil of ignorance treatment. These regression results are also directly comparable to

Table 2.3h: **Linear Regression of Confidence in Principle**

	Confidence in Principle	
Ambiguity preference	-0.001	-0.001
	(0.003)	(0.003)
Identity	0.175**	0.162**
	(0.070)	(0.074)
Identity group		
Ethnicity	-0.506	-0.114
	(0.318)	(0.327)
Nationality	-0.451*	-0.256
	(0.236)	(0.257)
Occupation	-0.497*	-0.254
	(0.260)	(0.278)
Race	-0.219	-0.046
	(0.422)	(0.471)
Religion	-0.684*	-0.178
	(0.395)	(0.407)
Self-deception 1	0.039	0.022
	(0.042)	(0.043)
Self-deception 2	0.051	0.020
	(0.035)	(0.038)
Constant	6.683***	6.683***
	(0.815)	(0.951)
Individual Controls		✓
Observations	971	859
Pseudo R-squared	0.018	0.083

Notes: Estimates come from a linear regression. The outcome variable 'Confidence in Principle' is measured from 1 to 10 and a higher value indicates more confidence in the chosen principle. Ambiguity preference ranges from 0 to 7 (with a higher value indicating more ambiguity seeking preferences) and is a standardized scale based on the ambiguity preference survey module developed by Cavatorta and Schröder (2019). Identity ranges from 1 to 4 with a higher value indicating a higher level of identity. This variable was measured using the module developed by Kuo and Margalit (2012). 'Other' is the reference group for identity group. Self-deception 1 ranges from 1 to 7 with a lower value indicating more self-deception. Self-deception 2 ranges from 1 to 7 with a higher value indicating more tolerance for deception. *** p<0.01, ** p<0.05 , * p<0.1.

table 2.5 in the main text. Descriptive social norms are again a highly significant predictor of distribution choices while personal principles only matter for the distribution choices of respondents who chose the meritocratic or maximin distribution with much smaller coeffi-

Table 2.3i: Personal Principle and Norms by chosen Distribution

	Chosen Distribution			
Personal Principle	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Inequality Aversion	**18.63%**	55.23%	11.11%	15.03%
Maximin	11.94%	**63.43%**	5.97%	18.66%
Meritocracy	12.61%	30.95%	**19.48%**	36.96%
Utilitarianism	18.40%	54.25%	11.79%	**15.57%**
Injunctive Norm				
Inequality Aversion	**17.65%**	54.71%	11.18%	16.47%
Maximin	13.33%	**55.24%**	10.48%	20.95%
Meritocracy	9.49%	34.49%	**19.62%**	36.39%
Utilitarianism	21.67%	51.67%	10.00%	**16.67%**
Descriptive Norm				
Inequality Aversion	**58.06%**	24.19%	5.65%	12.10%
Maximin	10.61%	**73.86%**	6.44%	9.09%
Meritocracy	7.27%	20.00%	**44.55%**	28.18%
Utilitarianism	8.37%	14.64%	18.83%	**58.16%**

Table 2.3j: Logistic regressions of distributive choices for Motivation Test

	Non-Veil of Ignorance Treatment			
	Choice of Distribution			
	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Personal Principle	0.240	0.643***	0.661***	-0.588***
	(0.194)	(0.208)	(0.202)	(0.220)
Injunctive Norm	0.224	0.306	0.712***	-0.436**
	(0.192)	(0.223)	(0.204)	(0.207)
Descriptive Norm	2.408***	2.649***	1.991***	2.309***
	(0.227)	(0.171)	(0.256)	(0.190)
Selfishness	0.396**	2.303***	-3.972***	-3.236***
	(0.193)	(0.171)	(0.720)	(0.372)
Individual Controls	✓	✓	✓	✓
Session Fixed Effects	✓	✓	✓	✓
Observations	886	886	886	886

Notes: Estimates come from individual logistic regressions. Personal Principle, Injunctive Norm, and Descriptive Norm are dummy variables equal to 1 if the subject's respective choice of principle or norm matched the distribution choice. Selfishness is a dummy variable equal to 1 if the subject chose the distribution that maximises the payoff of the quintile they were placed in based on their example quiz answers. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

cients. As in the main results reported in table 2.5 in the main text, holding a utilitarian principle or having a perceived utilitarian injunctive norm is again negatively associated with choosing the utilitarian distribution. The descriptive social norm coefficients are similar to those of the main regression results.

As this test included only the non-veil of ignorance treatment we could also include a selfishness variable. Contrary to our main results, selfishness is negatively associated with choosing the meritocratic and utilitarian distribution, yet positively associated with choosing the maximin distribution. This finding further supports the conclusion that selfishness is not a consistent predictor of behaviour in our experiment.

IX.3.4.4. Public Goods Game

Table 2.3k reports the results of the public goods game analysis for only the Motivation Test. These results are directly comparable to table 2.8 in the main text and table 2.3c and 2.3f of the appendix. For this test, we find that norm-following is no longer significant in models (1) and (2). The lack of significance might be due to the smaller sample size of this test compared to the main experiment and the pooled analysis.

Interestingly, contrary to the pooled analysis and the main experimental results in table 2.3c of the appendix, we find that selfishness is now significantly and positively associated with contributions in model (1).

As we were able to identify subjects who are conditional co-operators in this robustness check, we can test whether descriptive norm following in the distribution task is indeed related to conditional cooperation in the public goods game. In table 2.3l, we report the results of simple logit models with conditional cooperation as the outcome variable and an interaction between descriptive norm-following and each specific norm as the explanatory variable. Importantly, as norm-following itself was not actually significant in table 2.3k, this test will likely underestimate the relationship between descriptive norm following and conditional cooperation we would be able to observe in the pooled data or the main experiment. Unfortunately, in those experiments we did ,not elicit conditional cooperation. Nonetheless, as can be seen in table 2.3l, at least for those who followed a descriptive norm of maximin, norm-following is positively related to being a conditional co-operator.

Table 2.3k: Public Goods Game - Motivation Test

	One-Shot PG Game Contributions (0-20)	
	Motivation Test	
	(1)	(2)
Norm-following	0.675	0.530
	(0.450)	(0.456)
Preference-following	0.115	
	(0.475)	
Descriptive Norm		
Utilitarianism		-1.580***
		(0.538)
Meritocracy		-1.234
		(0.749)
Inequality Aversion		-0.594
		(0.601)
Selfishness	1.004**	0.550
	(0.424)	(0.436)
Constant	10.994***	11.837***
	(1.662)	(1.685)
Controls	✓	✓
Observations	886	886
R-squared	0.036	0.046

Notes: Estimates come from individual linear regressions. The outcome variable is subjects' contribution to the hypothetical one-shot public goods game with values ranging from 0-20. Norm-following is equal to 1 if the subject followed the perceived descriptive norm in their distribution choice and 0 otherwise. Preference-following is equal to 1 if the subject followed their principle in the distribution choice and 0 otherwise. The reference category for the descriptive norm variable is maximin, which was the modal perceived descriptive norm. Selfishness is a dummy variable equal to 1 if the subject chose the distribution that maximises the payoff of the quintile they were placed in based on their example quiz answer. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

IX.3.4.5. Preference-following in perceived Social Norm

Table 2.3m reports individual characteristics for respondents who chose a perceived descriptive social norm which is equivalent to their personal principle. A stronger social identification is a weakly significant predictor of having a personal principle that is equivalent to the perceived descriptive norm.

Table 2.3l: Conditional Cooperation

	Conditional Cooperation			
	Utilitarianism	Meritocracy	Maximin	Inequality Aversion
Norm-following	-0.148	-0.370	0.351**	0.018
	(0.240)	(0.417)	(0.162)	(0.309)
Constant	-1.419***	-1.422***	-1.584***	-1.440***
	(0.086)	(0.082)	(0.108)	(0.083)
Observations	1,001	1,001	1,001	1,001

Notes: Estimates come from individual logistic regressions. The outcome variable is conditional cooperation with values equal to 0 or 1. Norm-following is an interaction between a dummy variable for each descriptive norm and a dummy variable equal to 1 if that norm was followed in the distribution task, and 0 otherwise. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

IX.3.5 Second Principle Test

IX.3.5.1. Second Principle Distribution

Out of the 201 subjects included in the second principle test 113 indicated that they would take a second principle into consideration when deciding on how to distribute income in the group.

Table 2.3n reports the chosen second principle by first principle. The first thing to note is that maximin is not the most chosen second choice of any of the first principles. In fact, it is the least chosen second option. We additionally find that subjects are on average significantly (p=0.002) more confident in their first choice of principle (average of 7.325 on a 10-point scale) as opposed to their second choice (average of 6.673).

IX.3.6 Distribution Test

IX.3.6.1. Assumed distribution by chosen principle

Out of the 200 subjects included in the distribution test, 89 correctly identified the distribution associated with their chosen principle.

Table 2.3o reports the distribution subjects assumed to represent the chosen principle by chosen principle. Subjects who chose maximin as their principle were by far the best at identifying the distribution corresponding to their principle (77.14% correctly identified the distribution). Out of those who chose meritocracy as their principle (which is the majority

Table 2.3m: Logistic Regression of Principle-following in perceived Descriptive Norm

	Principle-followers in perceived Descriptive Norm	
Ambiguity preference	0.108	0.088
	(0.084)	(0.091)
Identity	0.156*	0.165*
	(0.084)	(0.095)
Identity group		
Ethnicity	0.402	0.697*
	(0.365)	(0.406)
Nationality	-0.056	0.121
	(0.282)	(0.318)
Occupation	-0.056	0.072
	0.305)	(0.342)
Race	0.588	0.762
	(0.479)	(0.541)
Religion	0.457	0.403
	(0.480)	(0.529)
Self-deception 1	-0.029	-0.036
	(0.050)	(0.055)
Self-deception 2	0.021	0.003
	(0.045)	(0.049)
Constant	-2.558***	-2.257**
	(0.619)	(0.881)
Individual Controls		✓
Observations	971	859
Pseudo R-squared	0.017	0.027

Notes: Estimates come from a linear regression. The outcome variable 'Principle-followers in perceive Descriptive Norm' is equal to 1 if the subject's perceived descriptive social norm is equivalent to their personal principle and 0 otherwise. Ambiguity preference ranges from 0 to 7 (with a higher value indicating more ambiguity seeking preferences) and is a standardized scale based on the ambiguity preference survey module developed by Cavatorta and Schröder (2019). Confidence is measured from 1 to 10 and a higher value indicates more confidence in the chosen principle. Identity ranges from 1 to 4 with a higher value indicating a higher level of identity. This variable was measured using the module developed by Kuo and Margalit (2012). 'Other' is the reference group for identity group. Self-deception 1 ranges from 1 to 7 with a lower value indicating more self-deception. Self-deception 2 ranges from 1 to 7 with a higher value indicating more tolerance for deception. *** p<0.01, ** p<0.05 , * p<0.1.

of subjects in our main waves), only 8% confused the maximin distribution with the meritocratic distribution. Most of those subjects thought the utilitarian distribution to be the meritocratic one. This emphasises the robustness of our main result, as meritocrats did not move towards maximin out of confusion.

Table 2.3n: Second Principle Choice by First Principle

	First Principle			
Second Principle	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Inequality Aversion	**0.00%**	58.62%	51.35%	70.59%
Maximin	26.67%	**0.00%**	21.62%	5.88%
Meritocracy	33.33%	27.59%	**0.00%**	23.53%
Utilitarianism	40.00%	13.79%	27.03%	**0.00%**
Total	100%	100%	100%	100%

IX.3.6.2. Main analysis excluding subjects with inequality aversion as a first principle

Out of those subjects who chose inequality aversion as their principle, 52.54% confused the maximin distribution with the inequality averse distribution. Given that this probably explains some of the movement towards maximin in the distribution choice, we repeated our main analysis excluding those who chose inequality aversion as their principle in table 2.3p. It is evident from the results reported in the table that excluding those with inequality aversion as their principle does not affect our main result - descriptive social norms are still significantly better predictors of distribution choices than personal principles or injunctive social norms. This result holds even when we only look at subjects in the non-veil of ignorance treatment and control for selfishness.

Table 2.3o: Assumed distribution by chosen principle

	Chosen Principle			
Distribution	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Inequality Aversion	**37.29%**	5.71%	4.00%	22.58%
Maximin	52.54%	**77.14%**	8.00%	25.81%
Meritocracy	3.39%	5.71%	**33.33%**	3.23%
Utilitarianism	6.78%	11.43%	54.67%	**48.39%**
Total	100%	100%	100%	100%

Table 2.3p: Logistic regressions of distributive choices (excluding inequality aversion principle-holders)

	All Treatments				Non-Veil of Ignorance Treatment			
	Choice of Distribution				Choice of Distribution			
	Inequality Aversion	Maximin	Meritocracy	Utilitarianism	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Personal Principle		1.126***	1.018***	-0.897***		0.789***	0.687**	-0.389*
		(0.134)	(0.174)	(0.141)		(0.231)	(0.302)	(0.227)
Injunctive Norm	0.169	0.687***	0.718***	-0.533***	0.137	0.148	0.659**	-0.763***
	(0.194)	(0.153)	(0.156)	(0.133)	(0.351)	(0.264)	(0.285)	(0.237)
Descriptive Norm	2.601***	2.158***	1.932***	1.993***	2.614***	2.258***	1.682***	1.924***
	(0.176)	(0.117)	(0.179)	(0.124)	(0.303)	(0.215)	(0.319)	(0.220)
Selfishness					0.301	-0.420**	0.371	0.118
					(0.257)	(0.183)	(0.267)	(0.187)
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
Session Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1,645	1,645	1,645	1,645	530	530	530	530

Notes: Estimates come from individual logistic regressions. Personal Principle, Injunctive Norm, and Descriptive Norm are dummy variables equal to 1 if the subject's respective choice of principle or norm matched the distribution choice. Selfishness is a dummy variable equal to 1 if the subject chose the distribution that maximises the payoff of the quintile they were placed in based on their example quiz answers. Robust standard errors are presented in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

IX.3.7 Wording Test

IX.3.7.1. Assumed distribution by chosen principle

Given the alternative wording of the maximin and inequality aversion statements used in our wording test, we first check whether the proportion of subjects correctly identifying the corresponding distribution has changed. Out of the 222 subjects included in the wording test, 88 correctly identified the distribution associated with their chosen principle. This is a significantly smaller proportion than subjects who correctly identified the distribution associated with their chosen principle when we used the original wording (39.64% compared to 44.50%). This finding therefore supports the use of our original statements in our main analysis. Table 2.3q reports the distribution subjects assumed to represent the chosen principle by chosen principle. The percentages are strikingly similar to those reported in table 2.3o of this appendix. Importantly, however, the proportion of respondents who correctly identified maximin and inequality aversion, the two principles for which the wording changed, decreased. In fact, the percentage of subjects correctly identifying inequality aversion as the distribution corresponding to their chosen principle decreased from just over 37% to about 29%.

Table 2.3q: Assumed distribution by chosen principle - alternative wording

	Chosen Principle			
Distribution	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Inequality Aversion	**28.99%**	3.57%	9.41%	19.51%
Maximin	53.62%	**75.00%**	11.76%	26.83%
Meritocracy	10.14%	7.14%	**31.76%**	2.44%
Utilitarianism	7.25%	14.29%	47.06%	**51.22%**
Total	100%	100%	100%	100%

IX.3.7.2. Main results

Table 2.3r reports the results of logistic regressions with individual distribution choices as the outcome variables for respondents in the Wording Test. This test was conducted with only the impartial spectator treatment. These regression results are directly comparable to table 2.5 in the main text. Despite the small sample size of the robustness check and the lower proportion of subjects who correctly identified the distribution corresponding to their principle, the main results are strikingly robust. Descriptive social norms are a consistent and highly significant predictor of distribution choices while personal principles are mostly not. Only those choosing the meritocratic distribution are significantly affected by their personal principle. The descriptive social norm coefficients are again similar to those of the main regression results.

Table 2.3r: Logistic regressions of distributive choices for Wording Test

	Impartial Spectator Treatment			
	Choice of Distribution			
	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Personal Principle	-0.248	0.404	1.534***	-0.442
	(0.657)	(0.455)	(0.463)	(0.555)
Injunctive Norm	1.265**	0.335	0.026	0.223
	(0.572)	(0.429)	(0.477)	(0.456)
Descriptive Norm	3.252***	1.649***	2.851***	1.976***
	(0.677)	(0.343)	(0.598)	(0.399)
Individual Controls	✓	✓	✓	✓
Session Fixed Effects	✓	✓	✓	✓
Observations	187	187	187	187

Notes: Estimates come from individual logistic regressions. Personal Principle, Injunctive Norm, and Descriptive Norm are dummy variables equal to 1 if the subject's respective choice of principle or norm matched the distribution choice. Robust standard errors are presented in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

IX.3.8 Order Test

IX.3.8.1. Preference- and Norm-following by chosen Distribution

Table 2.3s reports the chosen distribution by personal principle, perceived injunctive norm, and perceived descriptive norm for respondents in the Order Test. The pattern visible in table 2.3s is similar to the results of the main experiment: Descriptive social norms are more closely related to distribution choices than personal principles or injunctive norms. The percentage of descriptive norm followers is especially high for maximin with over 72% of respondents who chose the maximin distribution following their perceived descriptive norm. Interestingly, given this reversed order of decisions, the percentage of those who chose the inequality averse and utilitarian distributions and follow their perceived descriptive social norm decreased while the opposite is the case for those who chose the meritocratic distribution, compared to the results of our main waves.

Table 2.3s: Personal Principle and Norms by chosen Distribution

	Chosen Distribution			
Personal Principle	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Inequality Aversion	**16.13%**	62.90%	11.29%	9.68%
Maximin	3.33%	**53.33%**	13.33%	30.00%
Meritocracy	5.88%	16.47%	**42.35%**	35.29%
Utilitarianism	12.20%	48.78%	9.76%	**29.27%**
Injunctive Norm				
Inequality Aversion	**14.06%**	51.56%	25.00%	9.38%
Maximin	9.52%	**38.10%**	19.05%	33.33%
Meritocracy	6.41%	33.33%	**29.49%**	30.77%
Utilitarianism	9.09%	40.00%	14.55%	**36.36%**
Descriptive Norm				
Inequality Aversion	**40.63%**	28.13%	12.50%	18.75%
Maximin	5.56%	**72.22%**	13.89%	8.33%
Meritocracy	8.57%	17.14%	**51.43%**	22.86%
Utilitarianism	1.27%	27.85%	24.05%	**46.84%**

IX.3.8.2. Main results

Table 2.3t reports the results of logistic regressions with individual distribution choices as the outcome variables for respondents in the order test. This test was conducted with only the impartial spectator treatment. These regression results are directly comparable to table 2.5 in the main text. Similar to all previous robustness checks, the main results hold again. Despite the small sample size of this robustness check, descriptive social norms are a highly significant predictor of choices across all possible distributions. The descriptive social norm coefficients are again similar to those of the main regression results, although, given the smaller sample size, there is more variation. Personal principles are also significant predictors of the inequality averse and meritocratic distribution choices. Perceived injunctive norms however do not reach conventional levels of statistical significance for any of the distribution options.

Table 2.3t: Logistic regressions of distributive choices for Order Test

	Impartial Spectator Treatment			
	Choice of Distribution			
	Inequality Aversion	Maximin	Meritocracy	Utilitarianism
Personal Principle	1.415**	0.672	2.159***	0.137
	(0.576)	(0.461)	(0.441)	(0.436)
Injunctive Norm	0.959	-0.187	0.714*	0.724
	(0.592)	(0.559)	(0.364)	(0.377)
Descriptive Norm	3.738***	2.236***	1.900***	1.873***
	(0.773)	(0.362)	(0.460)	(0.382)
Individual Controls	✓	✓	✓	✓
Session Fixed Effects	✓	✓	✓	✓
Observations	196	196	196	196

Notes: Estimates come from individual logistic regressions. Personal Principle, Injunctive Norm, and Descriptive Norm are dummy variables equal to 1 if the subject's respective choice of principle or norm matched the distribution choice. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

IX.4 Description of Variables

Principle. Categorical variable capturing the principle selected by subject i.

1: Utilitarianism

2: Meritocracy

3: Maximin

4: Inequalty Aversion.

Distribution. Categorical variable capturing the distribution selected by subject i.

1: Utilitarianism

2: Meritocracy

3: Maximin

4: Inequalty Aversion.

Injunctive Norm. Categorical variable capturing the perceived social norm for the principle choice selected by subject i.

1: Utilitarianism

2: Meritocracy

3: Maximin

4: Inequalty Aversion.

Descriptive Norm. Categorical variable capturing the perceived social norm for the distribution choice selected by subject i.

1: Utilitarianism

2: Meritocracy

3: Maximin

4: Inequalty Aversion.

Treatment. Categorical variable capturing the treatment subject i is assigned to.

1: Impartial Spectator

2: Non-Veil of Ignorance

3: Veil of Ignorance

Gender. Dummy variable coded as 1 if subject i indicated to be female, 0 if subject i indicated to be male. Subjects who indicated "other" or "prefer not to say" were coded as missing values (n=22).

Age. Categorical variable capturing the age bracket of subject i.

1: 18-20 years old

2: 21-29 years old

3: 30-39 years old

4: 40-49 years old

5: 50-59 years old

6: 60 years or older

Student. Dummy variable coded as 1 if subject i is currently in full-time education, 0 otherwise.

Economics. Dummy variable coded as 1 if subject i indicated that they have taken a module in economics or a related subject at University. A value of 0 indicates that subject i either has not taken a module in economics or has never attended higher education.

Left-Right. Categorical variable capturing how much subject i agrees with the statement: "On economic policy matters, there is a role for the government".

1: Strongly Disagree

2: Disagree

3: Neither Agree nor Disagree

4: Agree

5: Strongly agree

Risk preferences. Variable capturing subject i's willingness to take risks on a scale from 0 to 10, where 0 means "completely unwilling to take risks" and a 10 means "very willing to take risks".

Income. Categorical variable capturing the income bracket of subject i. Values are stated in Pound Sterling (£) for subjects from the UK, in US Dollars ($) for subjects from the US and in Euros (€) for subjects from Western Europe.

1: Less than 20,000

2: 20,000 to 34,999

3: 35,000 to 49,999

4: 50,000 to 74,999

5: 75,000 to 99,999

6: Over 100,000

Sample. Categorical variable indicating whether subject i is a resident in the US, UK or Western Europe.

1: Europe

2: United Kingdom

3: United States

Quiz Performance. Variable ranging from 0 to 5, capturing the number of questions subject i correctly answered in the main Quiz.

Example Quiz Performance. Variable ranging from 0 to 2, capturing the number of questions subject i correctly answered in the example quiz of the non-veil of ignorance treatment.

Study. Variable indicating whether subject i was part of the first wave of the main study in November 2019 or the second wave in December 2019.

Principle Following. Dummy variable coded as 1 if subject i's chosen distribution is equal to their chosen principle.

Descriptive Norm Following. Dummy variable coded as 1 if subject i's chosen distribution is equal to their perceived descriptive social norm in the distribution choice.

Injunctive Norm Following. Dummy variable coded as 1 if subject i's chosen principle is equal to their perceived injunctive social norm in the principle choice.

Principle Following in perceived Descriptive Norm. Dummy variable coded as 1 if subject i's perceived descriptive social norm is equal to their chosen principle.

Injunctive Norm Following in Principle. Dummy variable coded as 1 if subject i's chosen principle is equal to their perceived injunctive norm.

Selfish. Dummy variable coded as 1 if subject i's chosen distribution is the distribution which maximises the income of their predicted quintile position from the example quiz in the non-veil of ignorance treatment.

Decision Group. Categorical variable indicating whether subject i is a norm-follower, principle-follower or selfish in the non-veil of ignorance treatment. Subjects that are both, norm- and principle-followers, are coded as principle-followers. Subjects that are both, norm-followers and selfish, are coded as selfish. Subjects that are both, principle-followers and selfish, are coded as selfish. This coding is used to ensure the most robust test of our hypotheses.
1: Norm-Following
2: Principle-Following
3: Selfish

Confidence in Principle. Variable capturing subject i's confidence in their chosen principle on a scale from 0 to 10, where 0 means "not confident at all" and a 10 means "Very confident".

Public Good Contribution. Variable ranging from 0 to 20 capturing the hypothetical contribution of subject i in the one-shot public goods game.

Conditional Co-operator. Dummy variable capturing whether subject i is a conditional co-operator. This variable is equal to 1 if subject i entered increasing values for each increasing hypothetical average contribution-level of the other group members.

Identity. Variable capturing subject i's social identification with a self-defined reference group, ranging from 1 to 4 with 1 indicating "Not strong at all" and 4 indicating "Very strong" social identity.

Identity Group. Categorical variable capturing the group subject i most identifies with. This variable is also used as the reference group for the Identity variable.

1: Your ethnicity

2: Your nationality

3: Your occupation

4: Your race

5: Your religion

6: Other

Self-Deception 1. Variable capturing subject i's self-deception measured as the level of agreement with the statement "It is okay to lie sometimes", ranging from 1 to 7 whereby 1 means "Strongly agree" and 7 means "Strongly disagree".

Self-Deception 2. Variable capturing subject i's self-deception measured as their response to the statement "There is a big debate in psychology over whether deception in experiments should be permitted. What do you think?", ranging from 1 to 7 whereby 1 means "Never" and 7 means "Whenever it helps science".

Ambiguity preference. Variable capturing subject i's preference for ambiguity ranging from 0 to 7 with 0 indicating ambiguity aversion and 7 ambiguity seeking preferences.

IX.5 Survey Instrument

All values below are given in Pound Sterling (£). This was changed to US Dollars ($) and Euros (€) depending on respondents' country of residence. All options in decisions 1-4 were presented in randomized order during the survey experiment. Distribution options in decisions 3 and 4 were presented as separate tables.

IX.5.1 Impartial Spectator Treatment

Background

A group of people are asked to do a quiz and their answers generate income. We rank their performance from the bottom 20% of performers to the top 20% in the table below and give the average income generated for a person in each 20% performance band. For example, this shows someone who performs in the middle band (the 3rd 20%) generates an income of £40 on average. Please click on the arrow below to proceed.

Performance Level	*Average Income*
Bottom 20% of performers	£20
2nd 20%	£30
3rd 20%	£40
4th 20%	£70
5th 20%	£110

Decision 1

Which of the following statements best describes how you think income should be distributed in this group? Please note, you are not part of this group.

- Inequalities are only justifiable if they improve the position of the least well-off group in society.
- Inequalities should be minimized.
- Individual income should be based exclusively on his/her ability and talents.
- Income should be distributed to maximize the average income in society.

Decision 2

All the participants of the study are now asked to select a statement. Each of you will be rewarded with a bonus payment of 50p if you select the statement chosen by most of the participants.

- Inequalities are only justifiable if they improve the position of the least well-off group in society.
- Inequalities should be minimized.
- Individual income should be based exclusively on his/her ability and talents.
- Income should be distributed to maximize the average income in society.

Decision 3

Below you can see four options for distributing the income generated by the quiz. It shows for each option how much a performer in each 20% band will receive. For example, a performer in the bottom 20% can either receive £20, £30 or £40 depending on the distribution. As mentioned earlier, performance on the quiz generates income for this group on average as in the table below:

Performance Level	*Average Income*
Bottom 20% of performers	£20
2nd 20%	£30
3rd 20%	£40
4th 20%	£70
5th 20%	£110

Which distribution option would you choose for this group? Please note, you are not part of this group.

	Average Income			
Performance Level	*Inequality Aversion*	*Maximin*	*Meritocracy*	*Utilitarianism*
Bottom 20%	£30	£40	£20	£20
2nd 20%	£60	£40	£30	£30
3rd 20%	£60	£50	£40	£50
4th 20%	£60	£60	£70	£70
5th 20%	£60	£80	£110	£110
Total	£270	£270	£270	£280

Decision 4

All the participants of the study are now asked to select a distribution. Each of you will be rewarded with a bonus payment of 50p if you select the distribution chosen by most of the participants.

	Average Income			
Performance Level	*Inequality Aversion*	*Maximin*	*Meritocracy*	*Utilitarianism*
Bottom 20%	£30	£40	£20	£20
2nd 20%	£60	£40	£30	£30
3rd 20%	£60	£50	£40	£50
4th 20%	£60	£60	£70	£70
5th 20%	£60	£80	£110	£110
Total	£270	£270	£270	£280

Quiz Introduction

You will now take part in the previously mentioned quiz which is the final part of this study. You will have 30 seconds to answer as many questions as possible. For your participation in the quiz you will receive an additional bonus payment of 50ct after completing the study. However, how well you perform on the quiz does not influence the size of this bonus payment.

IX.5.2 Veil of Ignorance Treatment

Background

People in a group that you belong to are asked to do a quiz and their answers generate income. We rank performance from the bottom 20% of performers to the top 20% in the table below and give the average income generated for a person in each 20% performance band. For example, the table below shows someone who performs in the middle band (the 3rd 20%) generates an income of £40 on average. In the following, you will participate in the above mentioned quiz and your performance will affect the bonus payment you will receive after completing the study. Please click on the arrow below to continue.

Performance Level	*Average Income*
Bottom 20% of performers	£20
2nd 20%	£30
3rd 20%	£40
4th 20%	£70
5th 20%	£110

Decision 1

Which of the following statements best describes how you think income should be distributed in your group?

- Inequalities are only justifiable if they improve the position of the least well-off group in society.
- Inequalities should be minimized.
- Individual income should be based exclusively on his/her ability and talents.
- Income should be distributed to maximize the average income in society.

Decision 2

All the participants in your group are now asked to select a statement. Each of you will be rewarded with a bonus payment of 50p if you select the statement chosen by most of the members of your group.

- Inequalities are only justifiable if they improve the position of the least well-off group in society.
- Inequalities should be minimized.
- Individual income should be based exclusively on his/her ability and talents.
- Income should be distributed to maximize the average income in society.

Decision 3

Below you can see four options for distributing the income generated by the quiz. It shows for each option how much a performer in each 20% band will receive. For example, a performer in the bottom 20% can either receive £20, £30 or £40 depending on the distribution. As mentioned earlier, performance on the quiz generates income for your group on average as in the table below:

Performance Level	*Average Income*
Bottom 20% of performers	£20
2nd 20%	£30
3rd 20%	£40
4th 20%	£70
5th 20%	£110

Which distribution option would you like to choose for your group? The distribution you choose will be implemented and affect the bonus payment you can earn through your performance on the quiz. The conversion rate for the bonus payment is £1=1p so if your performance puts you into the top 20% you can receive a bonus payment of 60p-110p depending on the distribution you have chosen.

	Average Income			
Performance Level	*Inequality Aversion*	*Maximin*	*Meritocracy*	*Utilitarianism*
Bottom 20%	£30	£40	£20	£20
2nd 20%	£60	£40	£30	£30
3rd 20%	£60	£50	£40	£50
4th 20%	£60	£60	£70	£70
5th 20%	£60	£80	£110	£110
Total	£270	£270	£270	£280

Decision 4

All the participants in your group are now asked to select a distribution. Each of you will be rewarded with a bonus payment of 50p if you select the distribution chosen by most of the members of your group.

	Average Income			
Performance Level	*Inequality Aversion*	*Maximin*	*Meritocracy*	*Utilitarianism*
Bottom 20%	£30	£40	£20	£20
2nd 20%	£60	£40	£30	£30
3rd 20%	£60	£50	£40	£50
4th 20%	£60	£60	£70	£70
5th 20%	£60	£80	£110	£110
Total	£270	£270	£270	£280

Quiz Introduction

You will now take part in the previously mentioned quiz which is the final part of this study. You will have 30 seconds to answer as many questions as possible. How well you perform on this quiz compared to the other participants determines in which of the five performance quintiles you will be placed. Your previously chosen distribution and your performance on this quiz therefore influence the bonus payment you will receive after completing the study.

IX.5.3 Non-veil of Ignorance Treatment

Background

People in a group that you belong to are asked to do a quiz and their answers generate income. We rank performance from the bottom 20% of performers to the top 20% in the table below and give the average income generated for a person in each 20% performance band. For example, the table below shows someone who performs in the middle band (the 3rd 20%) generates an income of £40 on average. In the following, you will participate in the above mentioned quiz and your performance will affect the bonus payment you will receive after completing the study. Please click on the arrow below to continue.

Performance Level	*Average Income*
Bottom 20% of performers	£20
2nd 20%	£30
3rd 20%	£40
4th 20%	£70
5th 20%	£110

Example Quiz

Please answer the following two questions. Based on your answers to these two questions we will predict how well you will perform on the quiz. You have 15 seconds to answer the questions.

- $9 \times 13 =$
- $80/2.5 =$

On the basis of your answer to these questions we predict that you would belong to the top/middle/bottom 20% of performers in the full quiz.

Decision 1

Which of the following statements best describes how you think income should be distributed in your group?

- Inequalities are only justifiable if they improve the position of the least well-off group in society.

- Inequalities should be minimized.
- Individual income should be based exclusively on his/her ability and talents.
- Income should be distributed to maximize the average income in society.

Decision 2

All the participants in your group are now asked to select a statement. Each of you will be rewarded with a bonus payment of 50p if you select the statement chosen by most of the members of your group.

- Inequalities are only justifiable if they improve the position of the least well-off group in society.
- Inequalities should be minimized.
- Individual income should be based exclusively on his/her ability and talents.
- Income should be distributed to maximize the average income in society.

Decision 3

Below you can see four options for distributing the income generated by the quiz. It shows for each option how much a performer in each 20% band will receive. For example, a performer in the bottom 20% can either receive £20, £30 or £40 depending on the distribution. As mentioned earlier, performance on the quiz generates income for your group on average as in the table below:

Performance Level	*Average Income*
Bottom 20% of performers	£20
2nd 20%	£30
3rd 20%	£40
4th 20%	£70
5th 20%	£110

Which distribution option would you like to choose for your group? The distribution you choose will be implemented and affect the bonus payment you can earn through your performance on the quiz. The conversion rate for the bonus payment is £1=1p so if your performance puts you into the top 20% you can receive a bonus payment of 60p-110p depending on the distribution you have chosen.

	Average Income			
Performance Level	*Inequality Aversion*	*Maximin*	*Meritocracy*	*Utilitarianism*
Bottom 20%	£30	£40	£20	£20
2nd 20%	£60	£40	£30	£30
3rd 20%	£60	£50	£40	£50
4th 20%	£60	£60	£70	£70
5th 20%	£60	£80	£110	£110
Total	£270	£270	£270	£280

Decision 4

All the participants in your group are now asked to select a distribution. Each of you will be rewarded with a bonus payment of 50p if you select the distribution chosen by most of the members of your group.

	Average Income			
Performance Level	*Inequality Aversion*	*Maximin*	*Meritocracy*	*Utilitarianism*
Bottom 20%	£30	£40	£20	£20
2nd 20%	£60	£40	£30	£30
3rd 20%	£60	£50	£40	£50
4th 20%	£60	£60	£70	£70
5th 20%	£60	£80	£110	£110
Total	£270	£270	£270	£280

Quiz Introduction

You will now take part in the previously mentioned quiz which is the final part of this study. You will have 30 seconds to answer as many questions as possible. How well you perform on this quiz compared to the other participants determines in which of the five performance quintiles you will be placed. Your previously chosen distribution and your performance on this quiz therefore influence the bonus payment you will receive after completing the study.

IX.5.4 Quiz

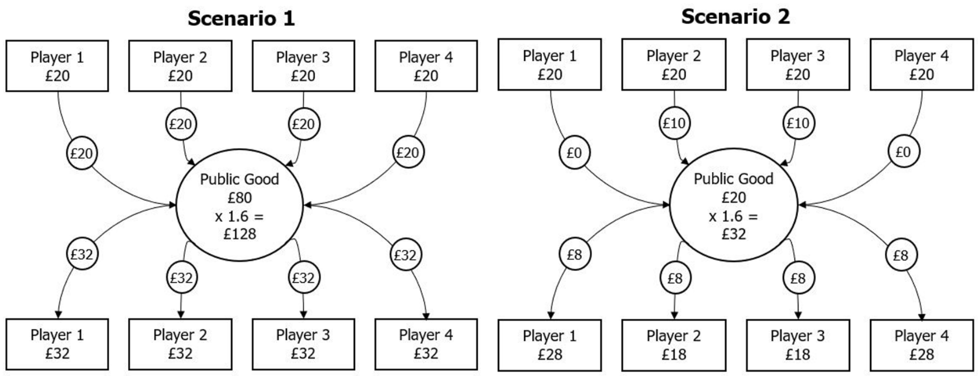
Please answer as many of the below questions as possible.

- $3 + 5 =$
- $8 \times 16 =$
- $(5 \times 8) - 12.2 =$
- $100 \times 10/5 =$
- $40/2.5 =$

IX.5.5 Public Goods Game

Before taking part in the previously mentioned quiz, you will now participate in a short hypothetical game within a group of four participants. Each of you has an endowment of £20. You can either contribute all or part of your endowment to a public good. The sum of the contributions of all four participants will then be multiplied by 1.6 and divided evenly amongst all group members.

The below table illustrates two potential scenarios of the game. In the first scenario all participants contribute their entire endowment (£20) to the public good. In the second scenario only the second and third player contribute half of their endowment to the public good. Please click on the arrow below to proceed to the game.



How much would you like to contribute? Please remember, the sum of the contributions of all four participants will be multiplied by 1.6 and divided evenly amongst all group members.

Your endowment: £20

Your contribution in £:

IX.5.6 Demographics

Nationality. What is your country of birth?

Gender. What is your gender?

- Female
- Male
- Other
- Prefer not to say

Age. How old are you?

- 18-20
- 21-29
- 30-39
- 40-49

- 50-59
- 60 or older
- Prefer not to say

Student. Are you currently studying towards a degree at University?

- Yes
- No

Economics. Have you ever taken a module on economics or a related subject area at University?

- Yes
- No
- I have never attended higher education

Income. What is your total personal income per year?

- Less than £20,000
- £20,000 to £34,999
- £35,000 to £49,999
- £50,000 to £74,999
- £75,000 to £99,999
- Over £100,000
- Prefer not to say

Risk preferences. Please tell us, in general, how willing or unwilling you are to take risks. Please use a scale from 0 to 10, where 0 means "completely unwilling to take risks" and a 10 means you are "very willing to take risks". You can also use any numbers between 0 and

10 to indicate where you fall on the scale.

Left-Right. How much do you agree or disagree with the following statement: "On economic policy matters, there is a role for the government"?

- Strongly agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

Rational. Were there any particular reasons for the principles and distributions you chose? Please use the field below to explain your choices.

Feedback. Please let us know in the field below whether you have any feedback regarding the study. Were any of the questions or tasks unclear?

IX.6 Pre-Analysis Plan

IX.6.1 Summary

Experimental economics has provided ample evidence that people frequently behave non-selfishly in situations where they can reduce their own payoff to help others (e.g. see Fehr et al. 2006). It is typically assumed that such pro-social behaviour arises because experimental subjects are motivated by a social preference. However, this assumption has so far not been tested. We test this assumption by conducting a survey experiment focusing on the consistency between people's preferences towards distributive justice and their choices over outcomes. In particular, we will look at the role of social norms in shaping choices as opposed to preferences and how this might affect the consistency between the two.

IX.6.2 Research Design

The survey experiment will be conducted on Prolific Academic with 1,200 subjects from the UK, US and continental Europe. The experiment consists of three treatments, each using a different institutional mechanism to elicit people's preferences over distributive justice: An impartial spectator treatment, a veil of ignorance treatment and a treatment in which one's position in the group is known. Subjects then have to firstly decide on the principle of distributive justice they most agree with (inequality aversion, maximin, meritocracy or utilitarianism) and play a coordination game to identify whether there is a social norm guiding their choice. Secondly, subjects are told that they will participate in a performance quiz and have to choose a payoff distribution they would like to implement in the group. Each payoff distribution thereby corresponds to one of the four previously mentioned principles of justice. Their performance on the quiz and the distribution chosen will determine their payoff after completing the experiment. Subjects are then asked again to play a coordination game to test whether their choice of distribution was guided by a social norm. Lastly, subjects are asked to play a one-shot public goods game.

The experiment addresses four main questions:

1. Is people's pro-social behaviour in experimental settings consistent with their social preferences? In other words, is there evidence consistent with the assumption that pro-social behaviour is motivated by social preferences?

2. Does the mechanism for eliciting social preferences affect the consistency of these preferences with pro-social behaviour?

3. Do the elicited social preferences and their consistency with pro-social behaviour in a distributive task affect pro-social behaviour in a public goods game?

4. Do people's pro-social behaviour reflect their beliefs about social norms within a group?

We will answer question 1. by comparing subjects' choice of principle with their chosen distribution. Question 2. will be tested by comparing people's choices across the three different treatments. Question 3. will be answered by testing whether subjects' public goods game contribution can be explained by their chosen principle and chosen distribution and to answer question 4. we will examine whether social norms exist in the coordination game and whether beliefs about these norms determine subjects' choices.

IX.6.3 Empirical Strategy

To identify whether pro-social behaviour of subjects is consistent with their social preferences, reflects social norms and to test whether the mechanism for eliciting these preferences affects their consistency we will estimate the following model:

$$D_{t,i} = \beta_0 + \beta_1 P_{t,i} + \beta_2 N_{t,i} + \beta_3 \lambda_{t,i} + \epsilon_{t,i} \tag{2.4}$$

Whereby D is the outcome measured as the chosen distribution by subject i in treatment t, P is the chosen social preference, N the elicited social norm in the distributive task, λ a vector of controls including age, gender, nationality, risk-aversion, income and student status and ϵ is the error term.

To identify whether there is a causal relationship between elicited social preferences and public goods game contributions we will estimate the following additional model:

$$C_i = \beta_0 + \beta_1 P_i + \beta_2 D_i + \beta_3 \lambda_i + \epsilon_i \tag{2.5}$$

Whereby C is the outcome variable measured as subject i's contribution to the one-shot public goods game, P is the chosen principle, D the chosen distribution, λ the vector of controls and ϵ the error term.

Lastly, to estimate whether beliefs about social norms shape social preferences, rather than just social behaviour in the distributive task, we will estimate a third additional model:

$$P_{t,i} = \beta_0 + \beta_1 N_{t,i} + \beta_2 \lambda_{t,i} + \epsilon_{t,i} \tag{2.6}$$

Whereby P is the principle chosen by subject i in treatment t, N is the elicited social norm in the principle decision, λ the vector of controls and ϵ the error term.

IX.6.4 Hypotheses

H1: There is a significant difference between subjects' chosen principle and chosen distribution.

H2: There is a significant difference in the level of consistency between chosen principles and chosen distributions across treatments.

H3: There is a significant positive relationship between the level of pro-sociality of subjects' chosen principle and distribution and their contribution in the public goods game.

H4: Social norms primarily influence subjects' choice of distribution in the group but not their choice of principle.

IX.6.5 Model specification

Previous research has found risk-aversion to be an important factor in determining peoples' distributive preferences, particularly behind a veil of ignorance (Carlsson et al. 2005; Schildberg-Hörisch 2010). To account for this potentially biasing factor we will include a control variable for risk-aversion into our model. We will further control for participants' income as various studies have previously found a causal link between income-levels and preferences for redistribution (e.g. see Esarey et al. 2012; Owens and Pedulla 2014; Naumann et al. 2016).

Previous research has also found participants' gender (Alesina and Angeletos 2005; Rehm 2005) and nationality to influence distributive preferences; particularly, whether participants live in the United States or Europe affects their preferred level of redistribution (Alesina and Glaeser 2004). In addition to these two factors we will also control for age and student status, as economics students have been found to be less inequality averse in experimental settings than the average population (Fehr et al. 2006), which may further bias our estimation.

IX.6.6 Weighting

The subjects participating in our experiment will be registered Prolific Academic participants from the US, UK and continental Europe. Unlike other online subject pools like Amazon's Mechanical Turk, Prolific verifies participants' identity; however, our sample is nonetheless not representative of the general populations of each country. We will therefore weight respondents, where appropriate, based on the age, gender and income distributions of the countries included in our sample.

IX.6.7 Robustness Checks

Self-interest: A further potential principle which could guide subjects' choice of principle and outcome, other than the four principles mentioned previously, is self-interest. To account for this potential explanation, we will estimate whether people's performance in the sample quiz in the non-veil of ignorance treatment determines their subsequent choice of distribution and principle by estimating the following additional two models:

$$P_i = \beta_0 + \beta_1 Q_i + \beta_2 \lambda_{t,i} + \epsilon_{t,i} \tag{2.7}$$

$$D_i = \beta_0 + \beta_1 Q_i + \beta_2 \lambda_{t,i} + \epsilon_{t,i} \tag{2.8}$$

Whereby P and D are the principle and distribution chosen by subject i, respectively, Q is the estimated quintile of subject i in the final quiz based on the answers to the sample questions, λ the vector of controls and ϵ the error term. If subjects are motivated by self-interest in their choice of principle and distribution and are aiming to maximise their own pay-off, those estimated to perform at the top of the distribution would choose either meritocracy or utilitarianism, those estimated to perform in the middle would choose inequality aversion and those at the bottom would choose maximin.

Consistency: A further concern regarding the validity of our estimation may be related to the determinants of why subjects are consistent or not consistent across the two choices. It may be the case that subjects are aware of the distributions corresponding to particular principles and want to appear consistent in front of the experimenters. Whilst we cannot directly test for this possibility, we can estimate whether the probability of being consistent across choices is determined by any of the demographic control variables. On the other hand, it may also be the case that subjects are more consistent in a particular treatment, as the institutional mechanism employed allows preferences to influence distributive choices more closely or social norms influence distributive choices less (*H2*). We will therefore estimate the following binary logit model:

$$Y_i = \beta_0 + \beta_1 T_{t,i} + \beta_2 \lambda_{t,i} + \epsilon_{t,i} \tag{2.9}$$

Whereby Y is a binary outcome variable equal to 1 if subject i's choice of principle and outcome is consistent, T is the treatment assigned to subject i, λ the vector of controls and ϵ the error term.

Comprehensiveness: Given the complexity of the distributive choice participants are expected to make, it may be the case that some participants choose a distribution at random if they do not understand the decision problem, instead of expressing a genuine preference or reflecting expectations about social norms. To account for this potential biasing factor, we are including a question asking for feedback at the end of the demographic questions. As an additional robustness check we will estimate our models including all participants as well as excluding those who mentioned being confused by the distributive choice in the feedback section. Based on our pilot results we estimate this group to be very small (1 out of 22 pilot participants mentioned the distributive choice to be somewhat confusing).

IX.6.8 Additional Analysis

Heterogenous treatment effects: We will test the effects of our three treatments on subjects' choices of principle, outcome and public goods game contribution conditional on subjects' characteristics such as level of risk-aversion, income-level, gender, age, student status and nationality. We will further test whether subjects' choice of principle conditional on these demographic characteristics determines their decisions over outcomes.

Sub-group analysis: We will use demographic variables such as risk aversion, income-level, nationality and gender to split our sample in different subgroups and test whether the consistency of subjects' choice of principle and choice of outcome differs between these groups. We will further test whether our treatments affect subjects' choices differently among these subgroups and whether the effects on public goods game contributions differ across these groups.

Multi-Mode Analysis: We will replicate the experiment in a laboratory setting at King's College London to account for potential measurement error caused by the experimental mode (Duch et al. 2020). We will do so by testing whether subjects' choices differ systematically between the online survey experiment and the laboratory setting.

Chapter 3

Experience of social mobility and support for redistribution: Accepting or blaming the system?

I Introduction

The level of social mobility in a society, or how much of a person's income and education can be predicted by that of their parents, is an important measure of economic opportunities within that society. Social mobility has therefore received much attention as a potential factor in explaining distributive preferences: If social mobility is high, economic outcomes appear to be the result of effort rather than a person's background and so demand for redistribution is expected to be low (Alesina and Glaeser 2004; Alesina and Angeletos 2005; Cappelen et al. 2013). While there is now substantial evidence that individuals' perceptions of societal mobility indeed affect their support for redistributive policies in this way (Corneo and Grüner 2002; Bjørnskov et al. 2013; Davidai and Gilovich 2015; Shariff et al. 2016; Alesina et al. 2018), less is known about how one's own experience of mobility affects these preferences. In fact, the limited existing evidence on the effects of personal mobility experience on distributive preferences suggests that there is no clear relationship between the two (Corneo and Grüner 2002; Alesina and Angeletos 2005; Clark et al. 2010; Guillaud

2013).

Using cross-country survey data from 26 countries collected across four waves and a survey experiment, I test a potential behavioural explanation for the previously missing link between own mobility experience and support for redistribution – the self-serving bias. This attribution bias states that people tend to blame external circumstances for their failures and take excessive personal credit for successes (Campbell and Sedikides 1999; Gilovich et al. 2002; Hestermann and Le Yaouanq 2021). Applying this bias to the case of social mobility experience suggests that people who have experienced upward mobility may be more likely to accept the system and to not extrapolate from their own experience onto society at large. On the other hand, those who experienced downward mobility may be more likely to blame the system and, therefore, to extrapolate from their experience onto society. This would suggest that the experience of social mobility has an asymmetric relationship with perceptions of societal social mobility and, in turn, distributive preferences.

Using the ISSP Social Inequality Cumulative (ISSP 2014), I find that such an asymmetric relationship between the experience of social mobility and distributive preferences indeed exists in observational cross-country data. Importantly, this relationship is not driven by personal income levels which are well known to affect distributive preferences (Alesina and Giuliano 2011). I further find that this asymmetric relationship also holds between the experience of social mobility and perceptions of societal mobility, suggesting that the mechanism through which mobility experience affects distributive preferences might be how it shapes beliefs about opportunities in society.

In an information provision experiment, I test the causality of this finding by providing subjects with an experimental shock to their mobility experience. The basic design of the experiment is similar to Hoy and Mager (2021), Karadja et al. (2017), and Cruces et al. (2013) who each provide subjects with a shock to their relative income position. Subjects in my experiment are asked to identify their own occupation and that of their parents when growing up. They are also asked to subjectively estimate their own mobility experience relative to their parents. I then calculate an experimental mobility measure for each subject equal to the difference between the subjective estimate and an objective mobility estimate based on income and education data for each occupation type. Holding this measure constant, I test the effect of being informed of one's objectively calculated mobility in the treatment condition relative to a control condition where subjects receive unrelated but similarly framed information.

I find that subjects who experience a negative mobility shock, by being informed of an objective mobility experience that is lower than their subjective estimate during the treatment condition, increase their support for redistribution significantly compared to subjects with the same experimental mobility score in the control group. Especially support for higher governmental spending on the poor and higher taxes on the rich increases for these subjects. Those who experience a positive mobility shock in the treatment condition do not change their distributive preferences relative to comparable subjects in the control group.

In line with the self-serving bias, only those who experience a negative mobility shock also change their perceptions of social mobility in society. However, neither group change their perceived personal benefits from redistribution, suggesting that this change in distributive preferences is not due to a rational change in perceived benefits.

I probe the robustness of the experimental results in a number of ways. First, I run a placebo test to account for the possibility that merely over- or underestimating something and being informed of it has an effect on perceptions and preferences. I do this by asking all subjects to estimate the difference in length between two rivers in North America and inform those in the control group about whether they over- or underestimated the length. I find no effects of this information on either preferences or perceptions.

Second, I check whether the experimental mobility measure I calculate, which is essentially a measure of misperception, is associated with any particular preferences, beliefs or demographics. I also test whether subjects who correctly identify their own mobility experience, and therefore do not experience a shock during the treatment, differ from the other subjects on any relevant measure. Neither is the case.

Third, I exclude subjects who do not believe the information they are provided with in both, the treatment and control group, and check whether those subjects (49 out of 1,100) differ on any relevant measure. They do not, except that subjects in the treatment group are somewhat more likely to not believe the information than those in the control, which is not surprising. I also restrict the main analysis to subjects who spent enough time on the treatment and control screens to read the information.[1]

Fourth, I test two plausible alternative models for the relationship between mobility experience and distributive preferences. I first run models with a continuous experimental mobility measure allowing for a linear relationship between the experienced mobility shock during the experiment and the outcome variables of interest. All preference coefficients hereby remain insignificant and near zero. Only overall mobility perceptions are positively related to the continuous experimental mobility measure but with a smaller coefficient than in the main models. I also test for the possibility that subjects' reference point is a weakly positive mobility experience rather than no mobility experience. The experimental results do not support this.

Finally, I report models with various alternative measures of mobility experience, such as only looking at subjects who experienced extreme mobility shocks during the experiment, using different measures of parents' income and education levels, and using simple information treatment effects without calculating the experimental mobility measure. The main results remain robust to all of these tests: Subjects who experience negative mobility increase their support for redistribution and decrease their perception of social mobility in society. Subjects who experience no mobility or positive mobility do not change their preferences or perceptions.

These findings are important for three main reasons. First, the new mechanism considered here can explain a puzzling observation which directly contradicts standard political economy models of redistribution (Meltzer and Richard 1981): Despite the increase in inequality over the past decades (Dabla-Norris et al. 2015) and the fall in social mobility, especially in the United States (Chetty et al. 2014a), there has been no significant increase in support for redistribution (Kenworthy and McCall 2008). Given that the results of this study suggest

[1] I also report the main results for all subjects, irrespective of time spent on the treatment and control screens and find no notable differences.

that only those with downward mobility experiences adjust their demand for redistribution, a decrease in absolute mobility, which means that there are both, less people with upward and less people with downward mobility experiences, ceteris paribus, then leads to less demand for redistribution overall. This somewhat counter-intuitive relationship is entirely consistent with the self-serving bias and supported by the descriptive and experimental evidence provided in this paper.

Second, the suggested relationship between people's experience of social mobility and distributive preferences allows to make predictions about changes in distributive preferences across time. This is more difficult when only looking at people's perceptions of societal mobility, given that little is known about how these perceptions are formed or affected by real-world events.

Third, this paper suggests a demand-side explanation for the Great Gatsby Curve with a different causal direction than the mechanisms usually discussed (e.g. see Jerrim and Macmillan 2015 and Sakamoto et al. 2014) - countries with lower levels of social mobility may see higher levels of inequality because, as the findings of this paper suggest, the lack of absolute mobility experience decreases demand for redistribution.

The outline of this paper is as follows. Section II provides a short overview of the relevant literature and conceptual framework, section III uses descriptive data to look at aggregate correlations between mobility experience and distributive preferences, and section IV describes the survey experiment and the experimental results. Section V concludes.

II Conceptual Framework

This paper builds on the substantial existing literature on determinants of preferences for redistribution at the individual level (e.g. Corneo and Grüner 2002; Klor and Shayo 2010; Alesina and Giuliano 2011; Luttmer and Singhal 2011; Durante et al. 2014; Kuziemko et al. 2015); specifically, on the growing literature on the relationship between social mobility and demand for redistribution (e.g. Piketty 1995; Benabou and Ok 2001; Alesina and Angeletos 2005; Clark et al. 2010; Esarey et al. 2012; Bjørnskov et al. 2013; Alesina et al. 2018;

Fehr et al. 2020) and on the literature on the effects of procedural fairness on distributive preferences (e.g. Alesina and Glaeser 2004; Alesina and Angeletos 2005; Cappelen et al. 2013). I also follow other studies looking at the effects of personal experiences on economic preferences more broadly (e.g. Malmendier and Nagel 2011; Fuchs-Schündeln and Schündeln 2015; Malmendier and Nagel 2016; Roth and Wohlfart 2018).

Prior to reviewing some of the findings of these studies in more detail, I will briefly discuss how social mobility has been conceptualized in the literature. Both, social mobility experience and perceptions of societal social mobility, are generally defined across two dimensions: absolute versus relative mobility and inter- versus intragenerational mobility. Absolute mobility, as I define it in this paper, is commonly measured as the correlation between children's and parents' income or, more broadly, the elasticity of income from one generation to the next. Relative mobility tends to be measured as the opportunity of a child born into the bottom quintile to rise to the top quintile (Chetty et al. 2014a,b, 2017).

A second dimension to take into consideration when discussing social mobility is inter- versus intragenerational mobility. While intergenerational mobility captures the effect of upbringing and family background on a person's socio-economic status, intragenerational mobility captures fluctuations in socio-economic status across a person's lifetime. Both, empirical estimates of real social mobility in society and studies of perceptions of mobility, tend to focus on intergenerational mobility. This is at least partly due to data limitations, as intragenerational mobility measures require long-term panels of individuals including their income fluctuations across time.[2] In this paper, I will include measures of both, absolute and relative mobility, but follow previous research by focusing on intergenerational as opposed to intragenerational mobility.

The idea that mobility experience affects distributive preferences is not new. As argued by Piketty (1995), mobility experience may affect distributive preferences at the individual level by shaping beliefs about societal mobility. That is because learning about the actual level of mobility in society by experimenting with effort levels is too costly. Few papers have so far however empirically examined how mobility experience affects distributive preferences.

[2] An example of an empirical study that uses intragenerational mobility measures is Kopczuk et al. (2010).

Alesina and Angeletos (2005) and Corneo and Grüner (2002) find that upward experienced mobility is associated with reduced support for redistribution. In contrast, Clark et al. (2010) and Guillaud (2013) find the exact opposite. These studies measure mobility experience as a dummy variable capturing whether a person believes to be better off than their parents or not. In other words, they do not differentiate between people who experienced negative or no mobility. They also only report self-assessed mobility experience and none use experimental methods to test the causality of the relationship.

The relationship between perceptions of social mobility in society and preferences for redistribution is, on the other hand, much better established. Most studies find that if social mobility is perceived to be low, demand for redistribution rises and vice versa (Corneo and Grüner 2002; Bjørnskov et al. 2013; Davidai and Gilovich 2015; Shariff et al. 2016; Alesina et al. 2018). One notable exception is a recent study by Fehr et al. (2020) who do not find evidence for this relationship on a sample of German participants. Nonetheless, in most studies this relationship is significant.

The mixed and somewhat contradictory existing descriptive evidence on the relationship between social mobility experience and distributive preferences can easily be explained by applying the self-serving bias and introducing perceptions of social mobility as a moderating variable. The self-serving bias states that people blame external circumstances for their failures and take excessive personal credit for successes (Campbell and Sedikides 1999; Gilovich et al. 2002). Applying this bias to the case of social mobility experience suggests that people who have experienced upward mobility may feel acceptance towards the economic system and do not extrapolate from their own experience onto society at large. They therefore do not update their perceptions of social mobility in society. On the other hand, those who experienced downward mobility may blame the system and, therefore, extrapolate from their experience onto society. Their perceptions of social mobility in society decrease, which subsequently increases their demand for redistribution.

This then suggests that, ceteris paribus, the experience of upward mobility does not actually have a particular effect on support for redistribution, which would explain the contradictory evidence in the existing literature. There may, of course, be other factors influencing distributive preferences for this group but the experience of mobility itself would not affect distributive preferences. On the other hand, the experience of downward mobility would lead to an increase in support for redistribution, holding other factors constant. This leads to my first two hypotheses:

Hypothesis 1: Individuals who have experienced upward social mobility, ceteris paribus, do not change their support for redistribution.

Hypothesis 2: Individuals who have experienced downward social mobility, ceteris paribus, increase their support for redistribution.

If both, H1 and H2, can be supported, it would provide evidence in line with the proposed relationship between the experience of social mobility and support for redistribution. The self-serving bias suggests that the mechanism underlying this relationship is how the experience of mobility affects perceptions of overall mobility in society. There is however a plausible alternative mechanism which would also be consistent with H2 and, potentially, H1: Differences in distributive preferences could be explained by differences in beliefs about marginal benefits from taxation. As one experiences downward mobility, perceived marginal benefits from redistribution rise and vice versa, leading to more (less) demand for redistribution. This would however not be consistent with H1, as it predicts a linear relationship between mobility experience and distributive preferences. If the perceived marginal benefits from redistribution for those with downward mobility experiences however outweigh the perceived marginal costs of those who moved up, the relationship could still be asymmetric and H1 and H2 could both be consistent with this explanation. Such an asymmetric relationship is entirely plausible if one takes loss aversion (Gilovich et al. 2002) into account. Given this possibility, merely testing H1 and H2 does not provide conclusive evidence for the self-serving bias explanation. Testing the two suggested mechanisms is therefore the secondary aim of this paper. Hypotheses 3 and 4 follow:

Hypothesis 3: Personal social mobility experience asymmetrically affects perceptions of societal mobility.

Hypothesis 4: Personal social mobility experience asymmetrically affects perceived marginal gains from redistribution.

In the following, I first look at descriptive cross-country data to test whether hypotheses 1 and 2 hold in the aggregate. To then get at the causality of the relationship and to test hypotheses 3 and 4, I report the results of an information provision experiment.

III Descriptive Data

The descriptive dataset used in this study is the ISSP Social Inequality Cumulative (ISSP 2014) which includes individual-level, representative data for all countries that participated in at least two waves of the ISSP Social Inequality Module, a total of 26. The individual waves of the module were conducted in 1987, 1992, 1999 and 2009, respectively and variables included in the cumulative dataset were included in at least two waves of the Social Inequality module. Out of these four waves, three can be used for the analysis as they include data on all the variables of interest for each individual respondent.[3] Overall, there are 103,538 respondents included in the dataset of which 26,866 respondents have provided responses to all the relevant questions for this estimation.[4]

Support for Redistribution: The main dependent variable, support for redistribution, is based on indicator V33 in the cumulative dataset of the ISSP Social Inequality Module. The indicator reports respondents' agreement with the statement "It is the responsibility of the government to reduce the differences in income between people with high incomes and those with low incomes". Respondents can indicate that they either "Strongly agree", "Agree", "Neither agree nor disagree", "Disagree" or "Strongly disagree".

Following Alesina and Giuliano (2011), I also look at support for redistribution as a dummy variable given that differences between individual points on the scale (e.g. "Strongly agree" versus "Agree") may not be as meaningful for some respondents as the difference between

[3]Data for West Germany and East Germany were collected separately in all waves but will not be treated separately in the main regression estimations. The data available for Slovakia in 1992 was in fact collected for the whole of Czechoslovakia, which had not yet split into Slovakia and the Czech Republic at that point.

[4]Some relevant questions were not asked in all countries and waves which significantly reduces the available sample size.

overall agreeing or disagreeing with the statement.[5]

Additionally, I look at respondents' agreement with the statements "Government should spend less on benefits for the poor", "Income differences in (R's country) are too large" and "Government should provide basic income for all", also coded as dummy variables. Lastly, I include item V40 which asks respondents about taxes in their country for those with high incomes. Possible answers range from "much too high" to "much too low" on a scale from 1 to 5.

Perceptions of Social Mobility: I measure perceptions of social mobility by generating an indicator based on individuals' answers to three questions, focused on the relative importance of family wealth, education and social connections in determining people's success in life using principle component analysis (PCA).[6] The resulting index ranges from 0 to 100 with a higher value indicating a higher level of perceived upward social mobility.[7]

Experienced social mobility: To measure people's own mobility experience I generate three indicators. First, I match the occupations of respondents and their parents, which are included as ISCO88 codes in the ISSP survey (ILO 1990), to the ISEI index of socio-economic status (Ganzeboom et al. 1992; Ganzeboom and Treiman 1996; Ganzeboom 2010), following in particular Yaish and Andersen (2012). This index captures the mean education and mean income of each occupation while controlling for age. The resulting scale ranges from 16 to 90 with a higher score indicating a higher level of socio-economic status. The individual-level experienced social mobility values (eSM) are then derived by subtracting the parental ISEI score ($ISEI_p$) from the respondents ISEI score ($ISEI_r$):

[5]To transform item V33 into a dummy variable I have followed the methodology of Corneo and Grüner (2002) and have coded respondents who answered with "Strongly agree" or "Agree" as 1 and respondents who answered with "Neither agree nor disagree", "Disagree" or "Strongly disagree" as 0.

[6]There is an ongoing debate in the literature about how to best measure people's perceptions of social mobility. A common measure of perceived social mobility is asking respondents about the likelihood of a person born into one quintile moving to another quintile within an income distribution, most commonly from the bottom to the top quintile. The ISSP does not include such a question but I have included it in the survey experiment.

[7]A detailed description of the individual components and the distribution of the generated index can be found in appendix section VI.1. Country-year-level estimates of the generated index can be found in table 3.1c also in section VI.1 of the appendix.

$$eSM_r = ISEI_r - ISEI_p \tag{3.1}$$

Whereby the parental ISEI score is derived based on the below equation:

$$ISEI_p = \max\{ISEI_f, ISEI_m\} \tag{3.2}$$

Hereby, $ISEI_f$ is the father's ISEI score and $ISEI_m$ the mother's score. The parental ISEI score ($ISEI_p$) is always equal to the score of the parent with the higher socio-economic status and the formula (3.1) used to derive the index ensures that the sign of the generated social mobility scale is equivalent to the direction of the experienced social mobility.[8] The generated index then ranges from -72, very negative mobility, to 72, very positive mobility, with $eSM_r = 0$ indicating no social mobility.[9]

Second, I follow previous research (e.g. Corneo and Grüner 2002) and use item V67 in the ISSP Social Inequality Cumulative which asks respondents about their relative occupational status compared to their father: "Please think of your present job (or your last one if you don't have one now). If you compare this job with the job your father had when you were [14/15/16], would you say that the level or status of your job is (or was)...". Respondents can then answer with "Much higher than your father's", "Higher", "About equal", "Lower" or "Much lower than your father's". I have coded respondents who did not know how or could not answer the question as missing variables. The resulting index then ranges from -2 to 2 with negative values indicating a subjective negative experience of social mobility and vice versa.

Third, I match country-level average hourly earnings from the Luxembourg Income Study (LIS 2019) with the ISSP Social Inequality Cumulative based on the ten major groups of the ISCO88 job classifications. I aggregated the ISCO88 classifications for the respondents, the mother, and the father in the ISSP survey to the ten major groups and match the

[8] Taking the average of the sum of the scores of both parents would decrease the score of a respondent with one working parent relative to a respondent with two working parents, where the scores of the respective parents with the higher status are equal. Given that the comparison is made to an individual respondent, the sum of the scores of both parents can also not be used.

[9] Further details of the matching procedure can be found in appendix section VI.1.2.

average hourly earnings with the respective group of respondents' in their country and year of surveying.[10] Unfortunately, the LIS does not go back far enough to provide accurate income data for the parents of respondents in the ISSP survey. I have therefore estimated the income of parents in the same way as that of respondents by matching the average hourly earnings at the time of surveying with the ISCO88 classification of each individual parent.

III.1 ISEI elasticity

To check the validity of the individual-level social mobility scores of respondents and to provide country-level social mobility estimates, I calculate the intergenerational elasticity of ISEI scores for each country and wave available in the dataset. I estimate the intergenerational mobility of the ISEI score by using the Poisson Pseudo Maximum Likelihood (PPML) estimator, which has been identified as one of the most robust estimators for mobility research (Mitnik 2017). The following model is generally used to estimate the IGE (intergenerational elasticity) of income, which I adapt for the ISEI scores following Andrews and Leigh (2009):

$$y_r = \beta_0 + \beta_1 X_p + AGE_r + AGE_r^2 + \epsilon \tag{3.3}$$

Whereby y_r denotes the ISEI score of the respondent and X_p the ISEI score of the parents. A polynomial for age is included as a control in the equation. β_1 is then the estimate of the intergenerational elasticity of the ISEI score. To ensure that only respondents of working age are included in the estimation, I restrict the model to respondents between the ages of 25 and 55.

Table 3.1 reports this estimate of the intergenerational elasticity of the ISEI score for each country and wave included in the sample. A low elasticity score indicates more social mobility and vice versa. For example, the elasticity score of 0.27 for the US in 2009 indicates that 27% of the difference between the average ISEI score in the US and that of a respondents' parents will be transferred to the respondent.

[10] The waves available in the LIS database do not match directly onto the waves of the ISSP dataset. Appendix section VI.1.3, therefore, provides an overview of the waves used for matching by country and year.

Table 3.1: ISEI elasticity by year (including age 25 to 55)

	1987	1992	1999	2009	Average
Country					
Australia		0.20	0.20	0.20	0.20
Austria	0.45	0.42	0.39	0.37	0.41
Bulgaria				0.38	0.38
Canada			0.19		0.19
Chile			0.47	0.45	0.46
Cyprus			0.36	0.32	0.34
Czech Republic		0.25	0.33	0.35	0.31
France			0.19	0.37	0.28
Germany (East)		0.30	0.27	0.47	0.35
Germany (West)	0.35	0.37	0.40	0.36	0.37
Hungary	0.32	0.33	0.25	0.42	0.33
Israel				0.31	0.31
Italy				0.35	0.35
Japan				0.16	0.16
Latvia			0.20	0.24	0.22
New Zealand		0.25	0.19		0.22
Norway		0.26	0.29	0.28	0.28
Philippines				0.23	0.23
Poland		0.39	0.37	0.35	0.37
Portugal			0.47	0.38	0.43
Russia		0.18	0.28	0.28	0.24
Slovak Republic		0.37	0.25	0.27	0.29
Slovenia			0.42	0.39	0.41
Spain			0.43	0.34	0.39
Sweden			0.28	0.34	0.31
Switzerland	0.24			0.38	0.31
United States		0.20	0.23	0.27	0.23
Average	0.34	0.29	0.31	0.34	0.32

Notes: ISEI elasticity values range from 0 to 1 with larger values indicating less socio-economic elasticity across generations and, therefore, less mobility.

Table 3.2: Income elasticity by year (including age 25 to 55)

	1987	1992	1999	2009	Average
Country					
Austria			0.32	0.39	0.36
Canada			0.16		0.16
Czech Republic			0.28	0.24	0.26
Germany (West)	0.24	0.23	0.39	0.36	0.31
Israel				0.26	0.26
Slovak Republic				0.25	0.25
Spain			0.38	0.33	0.36
Switzerland				0.33	0.33
United States		0.19	0.19	0.24	0.21
Average	0.24	0.21	0.29	0.30	0.28

Notes: Income elasticity values range from 0 to 1 with larger values indicating less income elasticity across generations and, therefore, less mobility.

Several patterns can be observed in table 3.1. First, across all countries, social mobility increased from 1987 to 1992 but has since then steadily decreased. This trend can already be observed in the 1999 wave but is further increased in the 2009 wave which, of course, also coincided with the financial crisis. There are further large differences across countries with Canada having the highest average level of social mobility with an elasticity value of 0.19 and Chile and Portugal having the lowest levels of mobility averaged across all waves.

As a preliminary robustness check of the ISEI indicator, I compare the derived estimates to the values obtained by Yaish and Andersen (2012), who equally match the ISEI indicator to the ISCO88 codes of respondents; however, for the 1992 and 1999 waves only. They further compare the scores of respondents and their fathers only and use a Full Maximum-Likelihood estimation model. The correlation of 0.96 suggests that the generated dataset, including two more waves and parental occupational status rather than the father's status only, is a suitable expansion of this existing dataset.

III.2 Income elasticity

I estimate the income elasticity similarly to the ISEI elasticity by using the PPML estimator and the model outlined in Section III.1. Income elasticity estimators are given in table 3.2.

Unfortunately, the LIS income data is only available for nine out of the 26 countries. These countries show a similar trend for the income data as for the ISEI data discussed before: From 1987 to 1992, income mobility appears to have improved on average but since then has significantly decreased again with the 2009 wave having, on average, the lowest level of income mobility.

III.3 Descriptive results

I estimate the correlation between social mobility experience eSM_i and respondent i's support for redistribution SfR_i:

$$SfR_i = eSM_i + \gamma_i + yearFE + countryFE + \epsilon_i \tag{3.4}$$

I include a vector of controls, γ_i, including own and parental ISEI scores, political orientation, education, gender, and age, as well as year- and country-fixed effects to account for any macroeconomic events that may have occured at the national level or between waves and could influence support for redistribution. Standard errors are clustered at the country-year level. As I am estimating multiple outcome variables, I account for multiple hypothesis testing by reporting adjusted p-values based on Anderson (2008).

The main descriptive results, which are all relative to respondents who experienced no mobility, are reported in table 3.3. In the first part of the table, I report the results for self-assessed mobility experience with the father. Based on this measure, there are 7,447 respondents who experienced negative mobility and 16,625 respondents who experienced positive mobility. No mobility is hereby defined as a self-assessed mobility score of 0 on a scale from -2 to 2. The second part uses objective personal mobility experience based on the ISEI score. Here, 8,553 respondents experienced negative mobility and 14,982 respondents experienced positive mobility. No mobility, on this measure, is defined as an ISEI mobility score between -7.2 and 7.2 on a scale from -72 to 72. The third part uses the same objective mobility experience but excludes subjects who misperceive the direction of their objective mobility experience. This leaves 3,131 respondents who experienced negative mobility and 9,087 respondents who experienced positive mobility. The final part of the table reports the results for income mobility using the LIS data. Based on this measure, there are 6,790 respondents who experienced negative income mobility and 29,359 respondents who experienced positive income mobility. This income mobility measure uses standardized earnings by occupation and no mobility is defined as a standardized earnings difference within +/- 5% of the mean.

There is a clear pattern observable in table 3.3. Using the self-reported mobility measure, a negative mobility experience is consistently associated with stronger support for redistribution on all measures, as well as a significantly more negative perception of societal social mobility. Specifically, having experienced negative mobility as opposed to no mobility on the self-reported measure increases support for redistribution on the binary outcome variable by 13.5 percentage points and on the ordered one by 3.2 percentage points. Additionally, agreement with the statement that income differences are too large increases by 20.3 percentage

Table 3.3: Support for Redistributive Policies

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich	Perception of Social mobility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Self-reported mobility experience							
Negative	0.135***	0.159***	0.203***	0.140**	0.203**	0.122***	-2.397***
	(0.047)	(0.040)	(0.047)	(0.064)	(0.105)	(0.043)	(0.382)
	[0.004]	[0.001]	[0.001]	[0.010]	[0.016]	[0.004]	[0.001]
Positive	0.020	0.007	0.023	0.027	0.025	0.041	-0.458
	(0.042)	(0.034)	(0.059)	(0.078)	(0.043)	(0.037)	(0.474)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Controls	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Country Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Observations	24,986	24,986	25,265	12,189	6,738	24,971	16,633
ISEI mobility experience							
Negative	0.159**	0.116*	0.010	0.100	0.007	0.018	-0.477
	(0.054)	(0.052)	(0.057)	(0.092)	(0.114)	(0.037)	(0.592)
	[0.022]	[0.082]	[1.000]	[0.852]	[1.000]	[1.000]	[1.000]
Positive	-0.098	-0.057	0.095	-0.116	-0.040	-0.023	0.497
	(0.040)	(0.031)	(0.054)	(0.090)	(0.086)	(0.044)	(0.499)
	[0.118]	[0.191]	[0.191]	[0.243]	[0.519]	[0.519]	[0.361]
Controls	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Country Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Observations	26,056	26,056	26,360	12,823	6,902	26,027	17,400
ISEI mobility experience (if aware of direction)							
Negative	0.260***	0.230***	0.169**	0.180*	0.217**	0.121**	-2.867**
	(0.070)	(0.062)	(0.087)	(0.157)	(0.084)	(0.070)	(1.005)
	[0.001]	[0.001]	[0.033]	[0.079]	[0.013]	[0.045]	[0.013]
Positive	-0.080	-0.049	0.007	-0.049	-0.139	0.024	0.097
	(0.066)	(0.059)	(0.087)	(0.129)	(0.138)	(0.056)	(0.951)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Controls	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Country Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Observations	12,525	12,525	12,637	5,995	3,471	12,514	8,260
Income mobility experience							
Negative	-0.070	-0.109	-0.153	0.174	0.052	-0.080	0.279
	(0.075)	(0.062)	(0.094)	(0.094)	(0.151)	(0.062)	(0.843)
	[0.394]	[0.329]	[0.329]	[0.329]	[0.527]	[0.329]	[0.527]
Positive	0.021	-0.006	-0.136	0.052	0.295	-0.070	-0.226
	(0.099)	(0.092)	(0.124)	(0.099)	(0.250)	(0.060)	(0.757)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Controls	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Country Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Observations	26,056	26,056	26,360	12,823	6,902	26,027	17,400

Notes: Estimates come from logistic (models (1), (3), (4) and (5)), ordered logit (models (2) and 6)) and linear (model (7)) regressions. Robust standard errors are clustered on a country-year level are presented in parentheses. Adjusted p-values for multiple hypothesis testing (Anderson 2008) are presented in brackets. No mobility is defined as either a self-reported score of 0 on a scale from -2 to 2, as an ISEI mobility score of -7.2 to 7.2 on a scale from -72 to 72 or as a standardised average earnings difference within +/- 5% of the mean. Positive and Negative mobility are then defined as above or below the no mobility threshold of the respective measure. Controls include the personal ISEI score, the parental score, political orientation, education, gender and age. *** p<0.01, ** p<0.05 , * p<0.1.

points, support for more spending on the poor by 14 percentage points, support for UBI by 20.3 percentage points and higher taxes on the rich by 2.4 percentage points. Finally, perceived societal mobility decreases by 2.4 percentage points. The same is not the case for those respondents who experienced upward mobility. These estimates are all relative to respondents who have experienced no mobility. The results for the objective ISEI mobility measure are less striking but, at least in the first two models, reveal the same pattern.

Using the objective measure only for respondents who are aware of the general direction of their mobility experience also results in a highly significant relationship between negative mobility experience and distributive preferences as well as mobility perceptions.[11] There is again no significant relationship between positive mobility experiences and preferences or perceptions. Finally, the LIS income measure shows no significant results at all.[12] This is potentially the case due to the high level of aggregation required to match the ISSP and LIS data or because of the small number of observations which could be matched successfully. Alternatively, it may suggest that income mobility alone does not affect perceptions and preferences in the same way as a socio-economic mobility measure.

Overall, these descriptive results suggest that an asymmetric relationship between mobility experience and distributive preferences as well as societal mobility perceptions indeed exists. However, and maybe unsurprisingly, this relationship is particularly strong for the self-assessed measure of mobility and for the objective mobility measure when respondents are aware of the direction of their mobility experience. If mobility experience was affecting preferences through, for example, some intergenerational transmission of beliefs as proposed, amongst others, by Piketty (1995), then one would not necessarily have to be aware of the direction of the own mobility experience. The results in table 3.3 do not support such an explanation.

Arguably, respondents who experienced positive or negative social mobility differ in other

[11] Tables 3.2b and 3.2c in appendix section VI.2.3. report the same regressions for respondents who experienced very high or very low mobility on both, the subjective and the objective ISEI mobility measures. The results show the same pattern. The effect sizes are even larger however than those in table 3.3 for respondents who experienced very negative objective mobility and are aware of the direction.

[12] I also report the results of the main regressions with an alternative threshold for upward and downward income mobility in table 3.2d in appendix section VI.2.3. The results are identical - income mobility, based on the LIS data, is not correlated with preferences or perceptions.

aspects besides their mobility experience, which may influence their preferences for redistribution and perceptions of social mobility at the societal level.[13]

While these preliminary results indicate significant and robust correlations in line with hypotheses 1,2 and 3, they do not allow for any causal statements about the effects of mobility experience on distributive preferences. The ISSP also does not include a question on perceived personal benefits from redistribution and so hypothesis 4 cannot be tested with this dataset.

Specifically, there are two important issues with using observational data to make inferences about the relationship between mobility experience and preferences. First, using mobility experience as an explanatory variable means that one has to disentangle the effect of mobility experience from that of a change in personal income and education-levels, which are both known to affect distributive preferences. Controlling for personal income therefore means that the social mobility measure captures parental income and education-levels. Vice versa, if one were to control for parental income levels the social mobility measure would capture own income. Arguably, the latter of these two options makes little sense. The benefit of an experimental test is that the experience of mobility can be affected by changing perceptions of own experiences, without actually changing personal or parental income or education levels. It therefore isolates the experience of mobility from these factors as much as possible, which is not feasible when using observational data.

The second fundamental issue with using observational data in this case is that there are good reasons to believe that the relationship between social mobility experience and distributive preferences suffers from reverse causality. In particular, if the mechanism underlying the relationship is the perception of societal mobility, then it is plausible that mobility experience does not just affect perceptions of societal mobility, but that the reverse is also true: Beliefs about opportunities in society could impact people's effort-levels which, in turn, might influence their mobility experience. There is, in fact, evidence in the existing literature that the perceived fairness of reward structures in workplace environments impacts

[13] Details of which factors are associated with upward and downward mobility in the ISSP dataset can be found in appendix section VI.2.1.

people's willingness to exert effort (e.g. Janssen 2000). Whether such a relationship exists at the societal level between perceived social mobility and exerted effort-levels has, as far as I am aware, not been tested yet. Nonetheless, this poses a fundamental issue to any inference using observational data only.

IV Survey Experiment

To account for these conceptual issues and to test the causality of the proposed relationship, I conducted a survey experiment in April 2021 with a sample of 1,100 subjects from the United States.[14] The United States hereby provides a particularly strong test of the self-serving bias as the US is a prime example of an individualistic country (Alesina and Glaeser 2004). Subjects might therefore be less likely to 'blame the system' and to extrapolate from their own experience onto society, which provides an additional hurdle to finding a significant relationship.

The aim of the survey experiment is to isolate the causal effect of social mobility experience on support for redistribution and perceptions of society. Therefore, the social mobility experience of subjects has to be changed exogenously. Given that the intergenerational aspect of mobility experience is difficult to model experimentally, an information provision experiment appears to be the best option. While I cannot change the real mobility experience of subjects, I can make use of the fact that about 50% of respondents in the ISSP survey are somewhat misinformed about the degree of their own mobility experience.[15] In particular, I provide subjects with information on their personal intergenerational mobility experience and test how this information, if contradictory to their previously held beliefs, changes their support for redistribution. In other words, I provide subjects with a shock to their personal mobility experience. Given that the descriptive results indicate that it is the mobility ex-

[14]The experiment was pre-registered via the American Economic Association registry for Randomized Controlled Trials with RCT ID AEARCTR-0007580 and was granted ethical clearance from the Research Ethics Committee at King's College London (reference number MRSP-19/20-21021).

[15]Table 3.2a in section VI.2.2 and table 3.3d in section VI.3.2 of the appendix test for differences between respondents who misperceive and correctly perceive their own mobility experience in the ISSP dataset and the experimental data, respectively. While there are notable differences in the ISSP dataset, this is not the case in the experimental data. In both datasets, perceived societal mobility also does not differ between the two groups.

perience one is aware of that affects preferences and perceptions, this experimental design allows me to change the part of mobility experience that appears to be most important for the purpose of this study - the mobility experience subjects are aware of. The basic design of the experiment is similar to Hoy and Mager (2021), Karadja et al. (2017), and Cruces et al. (2013), who each provide subjects with a shock to their relative income position.

To avoid deception and to ensure that the information provided is believable, I use subjects' real experienced social mobility to tailor the information provision conditional on the actual experience. I also ask subjects whether they find the information provided believable and exclude those from the analysis who do not find it believable (49 out of 1,100).[16] The basic structure of the experiment is outlined below:[17]

Part I: Subjects state their own occupation and that of their parents when they were growing up. Based on the given answers, each subject is assigned an ISCO88 code for their occupation and one for the occupation of each parent. They also state how they personally assess their social mobility experience to date, relative to their parents.

Part II: Subjects are divided into control and treatment group. The treatment group is given a short paragraph describing the person's mobility experience and the data used to calculate the mobility experience. To provide subjects with information about their mobility experience, the ISEI value of the parent with the highest ISEI score is subtracted from the subject's ISEI score. The ISEI scores used during the experiment for each ISCO88 code are available in Ganzeboom (2010). The control group is given similarly framed information about the difference in length between two rivers in the US, the Missouri and the Arkansas river.

Part III: Post-treatment questions about distributive preferences and beliefs about social mobility.

[16]Table 3.3f reports a balance test of subjects who did and did not believe the provided information. There are no significant differences between the two groups; however, subjects assigned to the control condition were somewhat more likely to believe the provided information than those in the treatment condition.

[17]The full survey instrument can be found in appendix VI.4.

Figure 3.1: Example Treatment Information Screen

> We will now tell you, based on the information you gave us earlier about your own job and the jobs your parents had when you grew up, whether you have objectively experienced upward, downward or no social mobility.
>
> The data we use is based on the International Standard Classification of Occupations (ISCO88) and the International Standard of Occupational Status (ISEI). This measure takes both, the required education-level and potential income of different jobs, into account.
>
> Based on the information you gave us and this measure, you experienced **upward mobility.**

To provide subjects with the information during the treatment condition, the experiment was coded to automatically calculate an objective mobility measure using subjects' responses and to then display a text based on the calculated value. Figure 3.1 displays the text given to a subject who experienced upward mobility and is randomly assigned to the treatment group.

To make the analysis as comparable as possible to the descriptive analysis using the ISSP data, I only sampled subjects between 25 and 55 years old who were not studying full-time towards a university degree at the time the experiment was conducted.

IV.1 Empirical strategy

Similar to the descriptive analysis, I estimate the following model, whereby SfR_i is subject i's support for governmental redistribution measured by a series of survey questions, eSM_i is the experimental mobility score of subject i, D is the treatment assignment, γ is a vector of controls, and ϵ is the error term:

$$SfR_i = eSM_i \times D_i + \gamma_i + \epsilon_i \tag{3.5}$$

The experimental mobility score (eSM) is thereby defined as the difference between self-assessed mobility experience compared with the father (identical to item V67 in the ISSP) and the objective mobility experience, calculated using the ISSP and ISEI scores. For the calculation of the experimental mobility score, the objective mobility experience is aggregated into five groups ranging from -2 to 2, equal to the self-assessed mobility scores. The experimental mobility measure can therefore range from -4 to 4 with a higher value indicating a more positive experience. It is important to emphasize that this measure is calculated for all subjects, irrespective of treatment assignment. Treatment effects are therefore measuring the effects of being informed of the objective mobility experience, in other words *experiencing a mobility shock*, while holding the underlying experimental mobility measure constant. This ensures that both, the subjective as well as the objective mobility experience of subjects, is controlled for and not confounding the estimation.[18]

Given that the assumption that subjects will compare the treatment information to their experienced mobility with the father is quite strong, I also look at an additional measure of mobility experience, based on household income, in the experimental results. Here, I subtract the family income of the subject when growing up from the current household income. Figure 3.2 reports density plots of both measures by control (blue) and treatment group (green). There are no notable differences between the two groups and both measures show approximately normal distributions.

Hypothesis 1 would predict that the interaction term $eSM_i \times D_i$ is positive for those who have a negative eSM value and were assigned to the treatment as opposed to the control group, while hypothesis 2 would predict this to not be the case for those with a positive eSM value.

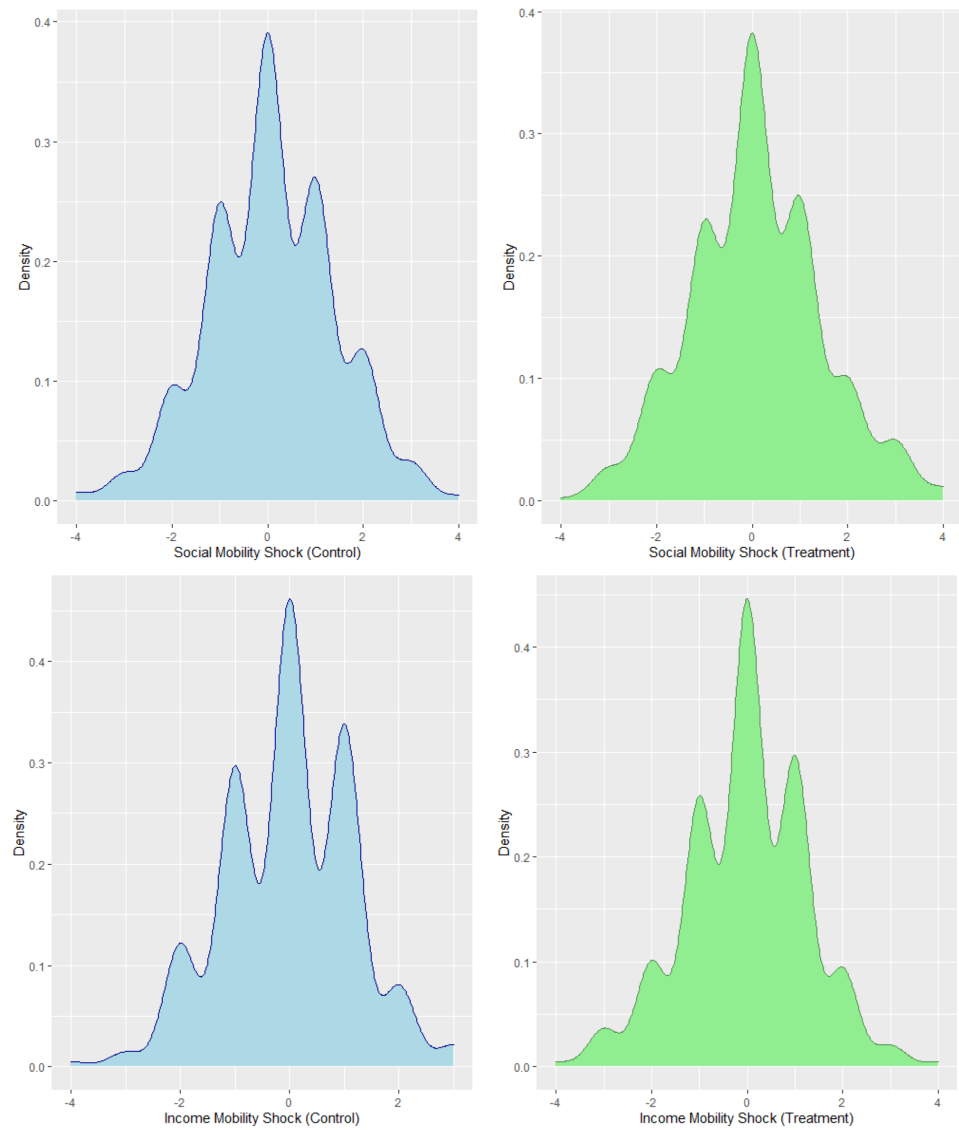
To answer the secondary research question, I then regress the same set of explanatory variables on both, perceived societal mobility in the United States, and perceived personal gain from governmental redistribution:

$$Y_i = eSM_i \times D_i + \gamma_i + \epsilon_i \tag{3.6}$$

Hypothesis 3 would predict the interaction term to be negative for those who have a negative

[18] A balance test by random treatment assignment is reported in table 3.3e in section VI.3.3 of the appendix. There is no variable that differs significantly between the treatment and control group, suggesting that randomization was successful.

Figure 3.2: Distribution of Mobility Shock by Treatment and Control



eSM value and were assigned to the treatment as opposed to the control group when perceived societal mobility is the outcome variable. Hypothesis 4 would predict the interaction term to be positive for the same group when perceived personal gain from governmental redistribution is the outcome variable.

This experimental test is, of course, by no means a perfect test of the effect of mobility experience on distributive preferences. However, if a short piece of information about own mobility experience can significantly change preferences and perceptions, it would suggest that real changes in mobility experience likely have quite a substantial impact.

IV.2 Experimental Results

Table 3.4: Experiment: Support for Redistribution

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
Upward mobility	0.130	0.249*	0.133	0.138	0.198	0.054
	(0.275)	(0.131)	(0.106)	(0.101)	(0.141)	(0.040)
	[0.340]	[0.340]	[0.340]	[0.340]	[0.340]	[0.340]
No mobility	0.185	0.118	0.116	0.047	0.134	-0.022
	(0.272)	(0.126)	(0.106)	(0.122)	(0.132)	(0.048)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Downward mobility	0.518*	0.247*	0.140	0.275***	0.015	0.133***
	(0.293)	(0.149)	(0.113)	(0.102)	(0.159)	(0.039)
	[0.109]	[0.109]	[0.173]	[0.018]	[0.348]	[0.007]
Controls	✓	✓	✓	✓	✓	✓
Observations	596	592	593	582	590	593
Treatment						
Upward income mobility	-0.206	0.083	0.216	0.219*	0.271	0.070
	(0.363)	(0.166)	(0.142)	(0.127)	(0.173)	(0.055)
	[0.447]	[0.447]	[0.348]	[0.348]	[0.348]	[0.348]
No income mobility	0.502	0.347	0.125	-0.057	0.248	0.053
	(0.772)	(0.266)	(0.151)	(0.265)	(0.318)	(0.105)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Downward income mobility	0.323	0.268	0.210	0.400***	-0.003	0.186***
	(0.361)	(0.184)	(0.143)	(0.137)	(0.204)	(0.053)
	[0.278]	[0.170]	[0.170]	[0.011]	[0.492]	[0.007]
Controls	✓	✓	✓	✓	✓	✓
Observations	283	283	283	279	282	283

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Robust standard errors are presented in parentheses. Adjusted p-values for multiple hypothesis testing (Anderson 2008) are presented in brackets. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** p<0.01, ** p<0.05 , * p<0.1.

Table 3.4 reports the main experimental results using the two different measures of experimental mobility experience. The first part of the table reports the effect of being in the treatment group as opposed to being in the control group for subjects depending on the experimental mobility score based on the comparison with the father, as previously defined. Using this measure, there are 322 subjects with a negative score, 455 subjects with a positive score, and 342 subjects with a score of zero.

The second part uses the difference between current household income and family income when subjects grew up to calculate the experimental mobility measure. Here, 175 subjects have a negative score, 253 subjects have a positive score, and 61 subjects have a score of zero.[19] The total number of subjects reported in table 3.4 is smaller for both measures, as the analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds.

While less striking than the descriptive results, which may be due to the significantly smaller sample size, for four of the outcome variables the pattern of the experimental data is the same. The effects of two measures even survive when adjusting the p-values for multiple hypothesis testing, despite the small sample size: Those who experienced a downward mobility shock in the treatment condition are significantly more likely to support more spending on the poor and higher taxes on the rich. The coefficients are also not negligible - being in the treatment as opposed to the control group as someone with a negative mobility score, increases support for more spending on the poor by 5 percentage points and support for higher taxes on the rich by 13 percentage points. Using the income measure, these effect sizes are even larger - support for more spending on the poor is increased by 8 percentage points and support for higher taxes on the rich by 19 percentage points for those with negative mobility scores in the treatment group.

General support for governmental redistribution is only weakly affected by the treatment with a weakly significant positive effect for those who experienced a downward mobility shock on the first measure. A possible reason for this might be that the question explicitly asks about whether it is the *responsibility of the government* to reduce differences in income. This may be viewed more negatively in the US than in some of the other, especially European, countries (Alesina and Glaeser 2004) included in the descriptive ISSP dataset.

Result 1: Experiencing a positive experimental mobility shock does not significantly increase support for redistribution on any of the measures (in support of H1).

Result 2: Experiencing a negative experimental mobility shock significantly increases support for governmental spending on the poor and higher taxes on the rich on both measures (in support of H2).

[19] As these values are self-reported and I am interested in the experimental mobility shock, I do not adjust family income when growing up for inflation. Rather, I assume that subjects compared their responses to D8 and D9 directly. The full survey experiment including the exact wording of these items can be found in appendix section VI.4.

IV.3 Mechanisms

To test the secondary research question and hypotheses 3 and 4, I look at the treatment effects on perceptions of societal mobility and perceived personal benefits from redistribution. The results are reported in table 3.5. As the models included in the table test different hypotheses, I have not adjust the p-values for multiple hypothesis testing.

Table 3.5: Experiment: Mobility and societal perceptions

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
Upward mobility	-0.061	0.048	-0.054	0.197
	(0.134)	(0.142)	(0.135)	(0.334)
No mobility	-0.122	-0.014	0.126	0.464
	(0.121)	(0.137)	(0.140)	(0.352)
Downward mobility	-0.252*	-0.359**	-0.077	0.311
	(0.148)	(0.152)	(0.155)	(0.358)
Controls	✓	✓	✓	✓
Observations	590	588	586	558

Treatment				
Upward income mobility	-0.066	-0.015	-0.064	0.675
	(0.160)	(0.170)	(0.166)	(0.413)
No income mobility	0.327	0.087	0.162	1.040
	(0.350)	(0.297)	(0.300)	(0.743)
Downward income mobility	-0.178	-0.493**	-0.041	-0.185
	(0.171)	(0.199)	(0.183)	(0.462)
Controls	✓	✓	✓	✓
Observations	281	282	282	264

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** p<0.01, ** p<0.05 , * p<0.1.

While neither those who experienced an upward mobility shock, no shock, or a downward shock in the treatment condition adjust their perception of the personal benefits from redistribution in either of the specifications, those who experienced a negative shock clearly changed their perception of overall social mobility in society compared to those in the control group. This is the case for both mobility measures. Being in the treatment as opposed to the control group as someone with a negative experimental mobility score decreases perceived overall societal mobility by 7 percentage points on the first measure and 10 percentage points on the second measure.

Interestingly, the perceived mobility of the lowest quintile is only weakly significantly affected for those who experienced a downward shock on the first measure and beliefs about whether income differences are due to effort or luck in society are not at all affected by the treatment. This suggests that the experimental mobility shock affects overall beliefs about social mobility in society, but that these beliefs do not translate into fairness perceptions as they are usually defined in the economics literature (Alesina and Glaeser 2004; Alesina and Angeletos 2005; Cappelen et al. 2013). These findings nonetheless provide support for hypothesis 3 and against hypothesis 4.

Result 3: Experiencing a negative experimental mobility shock on both measures significantly decreases perceptions of overall mobility in society, while experiencing none or a positive experimental shock has no effect on perceptions (in support of H3).

Result 4: Positive and negative experimental mobility shocks have no significant effect on beliefs about personal benefits from redistribution (against H4).

IV.4 Robustness

Given that the experiment does not change real social mobility experience, but only creates an experimental mobility shock for those in the treatment condition, I test the robustness of the results in multiple ways.

A first possible concern of the experimental design is that rather than the change in subjective personal mobility experience causing a difference in preferences, it may be the case that simply under- or overestimating something, such as own mobility experience, causes some negative reaction that could affect preferences. To account for this possibility, subjects in the control group were given a placebo treatment which asked them to estimate the difference between two of the longest rivers in North America and then gave them information framed similar to the treatment information, telling them whether they over- or underestimated the difference in length between the rivers. Figure 3.3 displays the text given to a subject who underestimated the difference in length between the two rivers and was randomly assigned to the control group. Tables 3.3a and 3.3b in appendix section VI.3.1 test the effect of the

Figure 3.3: Control Information Screen

> We will now tell you, based on the answer you gave us earlier about the two rivers in North America, the Missouri and the Arkansas river, whether you have objectively overestimated, underestimated or correctly estimated the difference in length between the two rivers.
>
> The data we use is based on the book "Rivers of North America" by Arthur C. Benke and Colbert E. Cushing.
>
> Based on the response you gave us, you **underestimated** the difference in length between the two rivers.

placebo information on preferences for redistribution, perceptions of societal mobility and perceived personal benefit from redistribution. Neither over-, nor underestimating the difference between the two rivers and being informed of the false estimation affect any of the preferences or perceptions.

A further concern is that having a negative experimental mobility score as opposed to a positive score may be correlated with individual characteristics that could also predict the outcome variables of interest. Section VI.3.2 in the appendix tests for potential differences in individual-level characteristics of subjects depending on their experimental mobility scores for both, the income and ISEI mobility measures. There are no concerning differences between those with positive and negative scores that would affect the main results. Neither is the case for those with experimental mobility scores of zero.

A significant factor influencing the results of information provision experiments is whether subjects believe and pay attention to the provided information (Haaland et al. 2020). While I was able to ask subjects directly about the believably of the information, measuring whether subjects actually read the provided text is more difficult. A potential proxy of paid attention is, however, the time spent on the treatment and control screens. The average time spent on these across all subjects was 14.5 seconds.[20] In the main estimations, I therefore exclude subjects who spent less than 8 seconds on the screens to ensure that I only include subjects who actually paid enough attention to fully read the provided information. This number is, of course, somewhat arbitrary and so I report main treatment effects for all subjects, irrespective of time spent on the treatment and control screens in section VI.3.4 of the appendix. The main results remain the same although the treatment effect on perceived social mobility in society for those with negative experimental mobility scores is even stronger.

A strong assumption made by the main experimental models in tables 3.4 and 3.5 is that subjects' reference point for evaluating their own mobility experience is no mobility relative to their parents. It is, however, entirely plausible that subjects actually expect to be better or worse off than their parents and so dividing the sample into groups of positive, negative, and no mobility may not be be most accurate. I test this possibility in multiple ways by looking at the effects of only extreme experimental mobility shocks and by aggregating subjects assuming a positive reference point.[21] I find no evidence that would suggest the main models reported in tables 3.4 and 3.5 are not suitable.

Finally, I test three alternative measures of the experimental mobility measure. First, I calculate mobility only with the parent who the subject believes themselves to have experienced the highest mobility with and then with the parent the subjects believes themselves to have experienced the least mobility with. Both are plausible alternatives for a comparison with the father.[22] I do not find evidence that would suggest that these models are better estimates than the main models reported in tables 3.4 and 3.5. Second, I test treatment effects without calculating the experimental mobility measure, but instead look at simple information effects while controlling for pre-treatment beliefs about the own mobility experience.

[20] Section VI.3.4 in the appendix provides further analysis of respondents who did and did not believe the information provided.

[21] The results of these tests can be found in tables 3.3i and 3.3j in section VI.3.5 and tables 3.3u and 3.3v in section VI.3.6 of the appendix.

[22] Section VI.3.5 in the appendix.

The main results remain the same, however, the effects on preferences for redistribution are only weakly significant in these models. The effects on societal mobility beliefs are somewhat stronger than the main models which differentiate between the effects of positive and negative mobility.[23] Third, I test whether a continuous mobility measure better predicts preferences and beliefs. This is not the case.[24] In fact, when using the continuous measure, none of the effects on preferences for redistribution remain.

IV.5 Party Affiliation

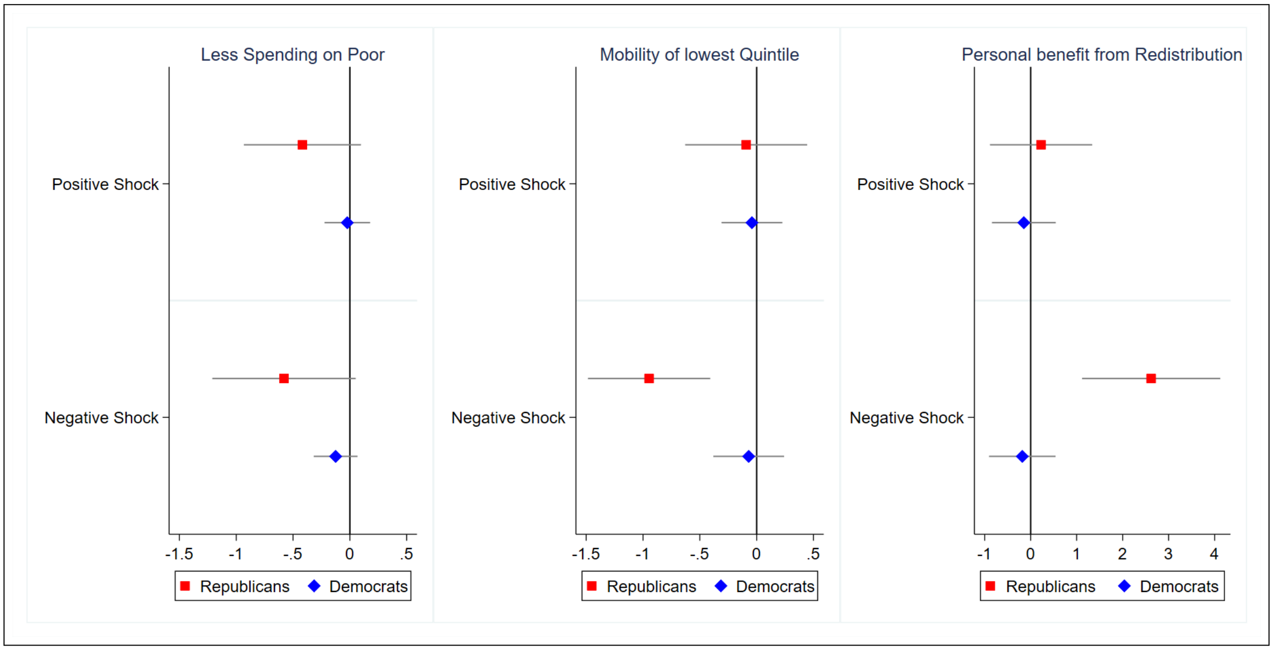
Following Alesina et al. (2018), who find that information about societal mobility only changes distributive preferences of Democrats, I also test treatment effects for Democrats and Republicans separately. Contrary to expectations, the effect of the self-serving bias is stronger for Republicans. In particular, as illustrated in figure 3.4, experiencing an effective negative mobility shock decreases agreement with the statement "The government should spend less on the poor" somewhat for Republicans in the treatment ($p<0.1$) as opposed to the control group. The same is not the case for those Republicans who experienced a positive shock or for Democrats. Given this interesting treatment effect on Republicans, I also test differences in the mechanism variables for Democrats and Republicans. An even more striking pattern emerges here. The effect of the treatment on perceptions of mobility is primarily driven by Republicans who have a more negative view of societal mobility if they were assigned to be in the treatment as opposed to the control group and experienced a negative mobility shock. Even more interesting however is that for Republicans, not for Democrats, the perceived personal benefits from redistribution are significantly increased if they are assigned to the treatment group and experienced an effective negative mobility shock. The coefficient here is also very large. While hypothesis 4 can therefore be rejected in the aggregate, it cannot be rejected for Republicans.

Result 5: While beliefs about personal benefits from redistribution are not affected by either a downward or upward mobility shock in the aggregate, Republicans increase their perceived benefit from redistribution if they experience a downward mobility shock significantly.

[23] Subsection 3 of section VI.3.5 in the appendix.

[24] Subsection 4 of section VI.3.5 in the appendix.

Figure 3.4: Treatment effects by political affiliaton



Notes: Figures are based on linear regressions. 95% confidence intervals are shown. The mobility shock is measured as the difference between self-assessed mobility and objective mobility. The effects are treatment effects relative to respondents with the same hypothetical shock in the control group. Controls include self-assessed mobility with the father, and personal income. The analysis is restricted to subjects who indicated that they believed the provided information.

V Conclusion

How does the experience of social mobility affect distributive preferences? The results of this paper suggest that the experience of social mobility is asymmetrically related to distributive preferences and mediated by perceptions of overall mobility within society. While negative mobility experience increases support for redistribution by changing perceptions of social mobility in society, positive mobility experience neither changes perceptions nor preferences. Therefore, as absolute social mobility decreases, which has been the case over the past couple of decades (see tables 3.1 and 3.2), ceteris paribus, demand for redistribution also decreases. This somewhat counter-intuitive finding is entirely consistent with a common attribution bias, the self-serving bias. Those with negative mobility experiences blame the system and extrapolate from their negative experience onto society at large, which increases their demand for redistribution. On the other hand, those with positive mobility experiences accept the system and, do not extrapolate from their experience onto perceptions of societal mobility, leading to no less support for redistribution.

In this paper, I have first estimated the correlation between personal mobility experience and distributive preferences with observational data from 26 countries collected across four waves spanning three decades. I have calculated three different measures of mobility experience for each individual respondent and found that experiencing negative mobility, if one is aware of the direction of the own mobility experience, significantly increases support for redistribution and decreases perceptions of societal mobility. This is not the case for those who experienced upward mobility. In the survey experiment, I find a similar patter. Experiencing a negative mobility shock during the experiment significantly increases support for governmental spending on the poor and higher taxes on the rich. The mechanism driving this effect appears to be a change in perceptions of societal mobility. Experiencing the negative experimental mobility shock significantly decreases perceptions of overall social mobility in society while experiencing a positive mobility shock does not. Interestingly, this effect is strongest for Republicans who adjust their beliefs and preferences substantially after experiencing the negative shock. For this subgroup, beliefs about personal benefits from redistribution also change significantly.

Given the nature of the research question, this study has, of course, some limitations. An information provision experiment can only provide a weak, short term shock to personal mobility experience and not all subjects believed the information provided. Given this, it is however even more surprising that the experiment resulted in any significant changes in preferences and beliefs. Future research in this area may be able to do more to simulate the experience of intergenerational social mobility, for example, in the laboratory.

An aspect that may also be addressed in future research is that of potential differential effects for men and for women. While I have followed previous research by using a question on personal mobility experience with the father as one of my main explanatory variables (Corneo and Grüner 2002), this may not be the best measure for everyone. Specifically, women might compare their own income or status to that of their mothers, rather than their fathers. This possibility is supported by the fact that women are more likely to misperceive their own mobility experience in the ISSP dataset.[25]. Given that this measure of misperception uses mobility with the father, it may be the case that women do not actually misperceive their mobility, but simply do not compare themselves to their fathers, but instead to their mothers. Unfortunately, many participants in the survey experiment stated that their mothers did not work when they were growing up and so such a comparison is difficult due to a lack of available data, at least with this sample.

Finally, there is an important qualification to the results of this paper. While objective mobility experience matters, it primarily does so when respondents are aware of the direction of their own experience. This may be unsurprising (e.g. see Gugushvili (2016) who finds that perceptions matter more than experience) but potentially also makes it more difficult to use the results of this study to make predictions about distributive preferences across time when only objective mobility measures are available. About half of the respondents in the ISSP and in the experimental data are, however, in fact, aware of their own mobility experience. Additionally, given that the relationship between objective experience and preferences is weak for those that are unaware of their experience, the predictions should still hold in the aggregate.

Despite these qualifications, the implications of the findings reported in this paper are significant. As the descriptive data reveals, social mobility has decreased in a majority of countries included in the ISSP dataset over the last decades. Ample research has also shown that income inequalities have increased over the same time period (Dabla-Norris et al. 2015). This paper provides a potential explanation for why these trends have not resulted in an increased demand for governmental redistribution. That is, because mobility experience is not linearly related to distributive preferences, but asymmetrically. A decrease in absolute mobility, therefore, leads, ceteris paribus, to less demand for redistribution. This finding also provides a potential demand-side explanation for the Great Gatsby Curve – countries with lower levels of social mobility may see higher levels of inequality because the lack of experience of absolute mobility decreases demand for redistribution.

[25] See table 3.2a in section VI.2.2 in the appendix.

Bibliography

Alberto Alesina and George-Marios Angeletos. Fairness and redistribution. *American Economic Review*, 95(4):960–980, 2005.

Alberto Alesina and Paola Giuliano. Preferences for redistribution. In *Handbook of Social Economics*, volume 1, pages 93–131. Elsevier, 2011.

Alberto Alesina and Edward Ludwig Glaeser. *Fighting poverty in the US and Europe: A world of difference*. Oxford University Press, 2004.

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

Michael L Anderson. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495, 2008.

Dan Andrews and Andrew Leigh. More inequality, less social mobility. *Applied Economics Letters*, 16(15):1489–1492, 2009.

Martha Augoustinos and Stephanie De Garis. 'too black or not black enough': Social identity complexity in the political rhetoric of barack obama. *European Journal of Social Psychology*, 42(5):564–577, 2012.

Emily Beller. Bringing intergenerational social mobility research into the twenty-first century: Why mothers matter. *American Sociological Review*, 74(4):507–528, 2009.

Roland Benabou and Efe A Ok. Social mobility and the demand for redistribution: the poum hypothesis. *The Quarterly Journal of Economics*, 116(2):447–487, 2001.

Christian Bjørnskov, Axel Dreher, Justina AV Fischer, Jan Schnellenbach, and Kai Gehring. Inequality and happiness: When perceived social mobility and economic reality do not match. *Journal of Economic Behavior & Organization*, 91:75–92, 2013.

Pierre Bourdieu. The forms of capital. In *The Sociology of Economic Life*. Routledge, 1986.

W Keith Campbell and Constantine Sedikides. Self-threat magnifies the self-serving bias: A meta-analytic integration. *Review of General Psychology*, 3(1):23–43, 1999.

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

Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623, 2014a.

Raj Chetty, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. Is the united states still a land of opportunity? recent trends in intergenerational mobility. *American Economic Review*, 104(5):141–47, 2014b.

Raj Chetty, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang. The fading american dream: Trends in absolute income mobility since 1940. *Science*, 356(6336):398–406, 2017.

Andrew Clark, Emanuela D'Angelo, et al. *Upward social mobility, well-being and political preferences: Evidence from the BHPS*. Università politecnica delle Marche, Dipartimento di economia, 2010.

Giacomo Corneo and Hans Peter Grüner. Individual preferences for political redistribution. *Journal of Public Economics*, 83(1):83–107, 2002.

Guillermo Cruces, Ricardo Perez-Truglia, and Martin Tetaz. Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, 98:100–112, 2013.

Ms Era Dabla-Norris, Ms Kalpana Kochhar, Mrs Nujin Suphaphiphat, Mr Frantisek Ricka, and Ms Evridiki Tsounta. *Causes and consequences of income inequality: A global perspective*. International Monetary Fund, 2015.

Shai Davidai and Thomas Gilovich. Building a more mobile america—one income quintile at a time. *Perspectives on Psychological Science*, 10(1):60–71, 2015.

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

Justin Esarey, Tim Salmon, and Charles Barrilleaux. Social insurance and income redistribution in a laboratory experiment. *Political Research Quarterly*, 65(3):685–698, 2012.

Dietmar Fehr, Daniel Muller, Marcel Preuss, et al. Social mobility perceptions and inequality acceptance. Technical report, 2020.

Sam Friedman. Habitus clivé and the emotional imprint of social mobility. *The Sociological Review*, 64(1):129–147, 2016.

Nicola Fuchs-Schündeln and Matthias Schündeln. On the endogeneity of political preferences: Evidence from individual experience with democracy. *Science*, 347(6226):1145–1148, 2015.

Harry BG Ganzeboom. A new international socio-economic index (isei) of occupational status for the international standard classification of occupation 2008 (isco-08) constructed with data from the issp 2002-2007. In *annual conference of international social survey programme, Lisbon*, volume 1, 2010.

Harry BG Ganzeboom and Donald J Treiman. Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research*, 25(3):201–239, 1996.

Harry BG Ganzeboom, Paul M De Graaf, and Donald J Treiman. A standard international socio-economic index of occupational status. *Social Science Research*, 21(1):1–56, 1992.

Thomas Gilovich, Dale Griffin, and Daniel Kahneman. *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge university press, 2002.

Alexi Gugushvili. Intergenerational objective and subjective mobility and attitudes towards income differences: evidence from transition societies. *Journal of International and Comparative Social Policy*, 32(3):199–219, 2016.

Elvire Guillaud. Preferences for redistribution: an empirical analysis over 33 countries. *The Journal of Economic Inequality*, 11(1):57–78, 2013.

Ingar Haaland, Christopher Roth, and Johannes Wohlfart. Designing information provision experiments. 2020.

Nina Hestermann and Yves Le Yaouanq. Experimentation with self-serving attribution biases. *American Economic Journal: Microeconomics*, 13(3):198–237, 2021.

Christopher Hoy and Franziska Mager. Why are relatively poor people not more supportive of redistribution? evidence from a randomized survey experiment across ten countries. *American Economic Journal: Economic Policy*, 13(4):299–328, 2021.

ILO. *International standard classification of occupations: ISCO-88*. International Labour Organization, 1990.

ISSP. International social survey programme: Social inequality i-iv-issp 1987-1992-1999-2009. *GESIS Data Archive, Cologne. ZA5890 Data file Version*, 1(0), 2014.

Onne Janssen. Job demands, perceptions of effort-reward fairness and innovative work behaviour. *Journal of Occupational and Organizational Psychology*, 73(3):287–302, 2000.

John Jerrim and Lindsey Macmillan. Income inequality, intergenerational mobility, and the great gatsby curve: Is education the key? *Social Forces*, 94(2):505–533, 2015.

Mounir Karadja, Johanna Mollerstrom, and David Seim. Richer (and holier) than thou? the effect of relative income improvements on demand for redistribution. *Review of Economics and Statistics*, 99(2):201–212, 2017.

Lane Kenworthy and Leslie McCall. Inequality, public opinion and redistribution. *Socio-Economic Review*, 6(1):35–68, 2008.

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

Wojciech Kopczuk, Emmanuel Saez, and Jae Song. Earnings inequality and mobility in the united states: Evidence from social security data since 1937. *The Quarterly Journal of Economics*, 125(1):91–128, 2010.

Ilyana Kuziemko, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva. How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, 105(4):1478–1508, 2015.

LIS. Luxembourg income study (lis) database, http://www.lisdatacenter.org (multiple countries; 1987-2009). *Luxembourg: LIS*, 1(0), 2019.

Erzo FP Luttmer and Monica Singhal. Culture, context, and the taste for redistribution. *American Economic Journal: Economic Policy*, 3(1):157–79, 2011.

Ulrike Malmendier and Stefan Nagel. Depression babies: do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics*, 126(1):373–416, 2011.

Ulrike Malmendier and Stefan Nagel. Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87, 2016.

Allan H Meltzer and Scott F Richard. A rational theory of the size of government. *Journal of Political Economy*, 89(5):914–927, 1981.

Pablo Mitnik. Estimators of the intergenerational elasticity of expected income: A tutorial, 2017.

Thomas Piketty. Social mobility and redistributive politics. *The Quarterly Journal of Economics*, 110(3):551–584, 1995.

Christopher Roth and Johannes Wohlfart. Experienced inequality and preferences for redistribution. *Journal of Public Economics*, 167:251–262, 2018.

Arthur Sakamoto, Jason Rarick, Hyeyoung Woo, and Sharron X Wang. What underlies the great gatsby curve? psychological micro-foundations of the "vicious circle" of poverty. *Mind & Society*, 13(2):195–211, 2014.

Azim F Shariff, Dylan Wiwad, and Lara B Aknin. Income mobility breeds tolerance for income inequality: Cross-national and experimental evidence. *Perspectives on Psychological Science*, 11(3):373–380, 2016.

Meir Yaish and Robert Andersen. Social mobility in 20 modern societies: The role of economic and political context. *Social Science Research*, 41(3):527–538, 2012.

VI Appendix

VI.1 Data and Methodology

VI.1.1 Perception of Social Mobility Indicator

To capture the perception of social mobility of respondents as accurately as possible, I generated an indicator based on individuals' answers to three separate questions focused on different aspects of mobility within society using principle component analysis (PCA). Indicator V8 in the cumulative dataset of the ISSP Social Inequality Module asks respondents "How important is coming from a wealthy family for getting ahead in life?". Respondents can respond with either "Essential", "Very important", "Fairly important", "Not very important" or "Not important at all".

Table 3.1a: Distribution of Components of the Perceived Social Mobility Index

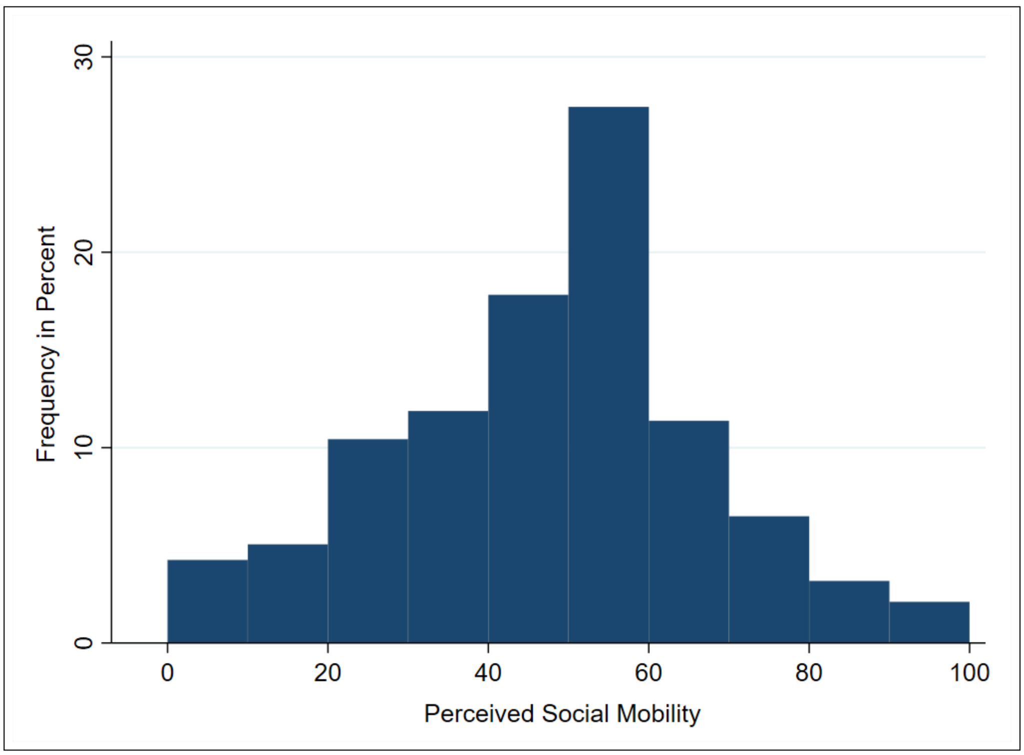
	Essential	Very Important	Fairly Important	Not very important	Not important at all	Total
Indicator						
How important is coming from a wealthy family for getting ahead in life?	8.59%	20.50%	30.91%	27.29%	12.72%	100%
How important is having well-educated parents for getting ahead in life?	7.96%	27.78%	36.60%	20.28%	7.37%	100%
How important is knowing the right people for getting ahead in life?	16.52%	34.20%	33.62%	12.36%	3.29%	100%

Indicator V9 asks respondents "How important is having well-educated parents for getting ahead in life?". Respondents can again respond with either "Essential", "Very important", "Fairly important", "Not very important" or "Not important at all". Finally, indicator V14 asks respondents "How important is knowing the right people for getting ahead in life?". Respondents can again respond with either "Essential", "Very important", "Fairly important", "Not very important" or "Not important at all". The distribution of responses to all three indicators is reported in table 3.1a. The correlation between indicator V8 and

V9 is 0.46, between V8 and V14 0.36 and between V9 and V14 0.24. These three indicators each ask about a different aspect of social mobility – parental wealth, parental education and personal connections – which correspond to the three forms of capital as defined by Bourdieu (1986).

To combine the three questions into one indicator, I have used principal component analysis (PCA) following Esarey et al. (2012) who also use PCA to generate an index of individual-level 'conservatism' based on survey data. This method allows me to isolate the underlying common component of perceived social mobility in individuals' responses to these three separate questions.

Figure 3.1a: Distribution of Perceived Social Mobility Index



The first principal component has by far the largest Eigenvalue of all three potential components and is the only component that is correlated with all three indicators in the correct direction. The compositions of the different components can be found in table 3.1b.

Table 3.1b: Principal Components

	Component 1	Component 2	Component 3
Variable			
V8	0.6308	-0.1546	-0.7604
V9	0.5822	-0.5537	0.5954
V14	0.5130	0.8183	0.2592

To make the interpretation of the values more intuitive I normalised the index and multiplied each value by 100. The resulting index then ranges from 0 to 100 with a higher value indicating a higher level of perceived social mobility. Figure 3.1a shows the distribution of the generated index in percent. An overview of country-level mean values of the generated index by waves can be found in table 3.1c.

VI.1.2 Matching Procedure for the socio-economic index of social mobility

The ISEI is available for 533 of the individual ISCO88 occupation types (Ganzeboom and Treiman 1996, appendix A 221-37). Respondents in the ISSP dataset indicated a total of 566 different ISCO88 occupation types, leading to a total of 670 respondents for which no status score is available based on the ISEI. On top of that 1,059 ISEI values are missing for fathers of respondents and 25 values are missing for respondents' mothers. Most of these respondents are armed forces personnel (347 of respondents, 802 of respondents' fathers and 15 of respondents' mothers) which the ISEI treats differently depending on the role of the individual within the armed forces (Ganzeboom and Treiman 1996, 209). For example, an ordinary soldier has an ISEI score of 40 whilst a non-commissioned officer has a score of 56. Given that the ISSP does not provide any further information on the role of respondents within the military, no ISEI score can reasonably be included for these respondents without biasing the estimate given the large disparity of ISEI scores for different armed forces personnel. Another group of respondents which do not match directly onto the ISEI scores are middle school teachers, as these are divided into those on an academic track and those on a vocational track in their ISEI ranking. The ISEI score difference between the two groups is

Table 3.1c: Perceived Social Mobility Index by year (mean)

	1987	1992	1999	2009	Average
Country					
Australia	55.42	49.65		49.10	51.39
Austria	41.26	44.12		42.91	42.76
Bulgaria		42.93		36.73	39.83
Canada		52.95			52.95
Chile				43.73	43.73
Cyprus				47.21	47.21
Czech Republic		57.57		51.03	54.30
France				54.34	54.34
Germany (East)		48.08		40.61	44.35
Germany (West)	45.47	49.08		41.43	45.33
Hungary	49.81	49.77		42.15	47.24
Israel				41.82	41.82
Italy	40.84	41.81		42.50	41.71
Japan				59.28	59.28
Latvia				42.77	42.77
New Zealand		52.16		57.00	54.58
Norway		57.93		55.87	56.90
Philippines		37.28		42.86	40.07
Poland		38.86		35.58	37.22
Portugal				46.39	46.39
Russia		42.94		42.24	42.59
Slovak Republic		50.10		41.21	45.66
Slovenia		51.27		42.81	47.04
Spain				43.29	43.29
Sweden		53.43		53.93	53.68
Switzerland	49.08			51.07	50.08
United Kingdom	51.15	52.63		53.36	52.38
United States	48.34	48.83		44.31	47.16
Average	47.67	51.19		47.91	49.11

only four points and so I decided to match respondents and their parents with the occupation 'middle school teacher' to ISEI code 2322 which is the vocational track-subgroup. This covers all ten remaining missing values for respondents' mothers. The remaining 272 missing ISEI values for respondents and 234 missing values for respondents' fathers are all country-specific classifications from Norway and New Zealand that cannot reasonably be assigned to existing ISEI codes without any further information. These respondents are therefore also excluded from the analysis.

Table 3.1d: Luxembourg Income Waves used by Country and Year

	1987	1992	1999	2009
Country				
Austria			at00p	at10p
Canada			ca98p	
Czech Republic			cz96p	cz10p
Germany (West)	de87p	de91p	de98p	de09p
Israel				il10p
Slovak Republic				sk10p
Spain			es00p	es10p
Switzerland				ch10p
United States		us91p	us00p	us10p

VI.1.3 Luxembourg Income Study - Matching Procedure

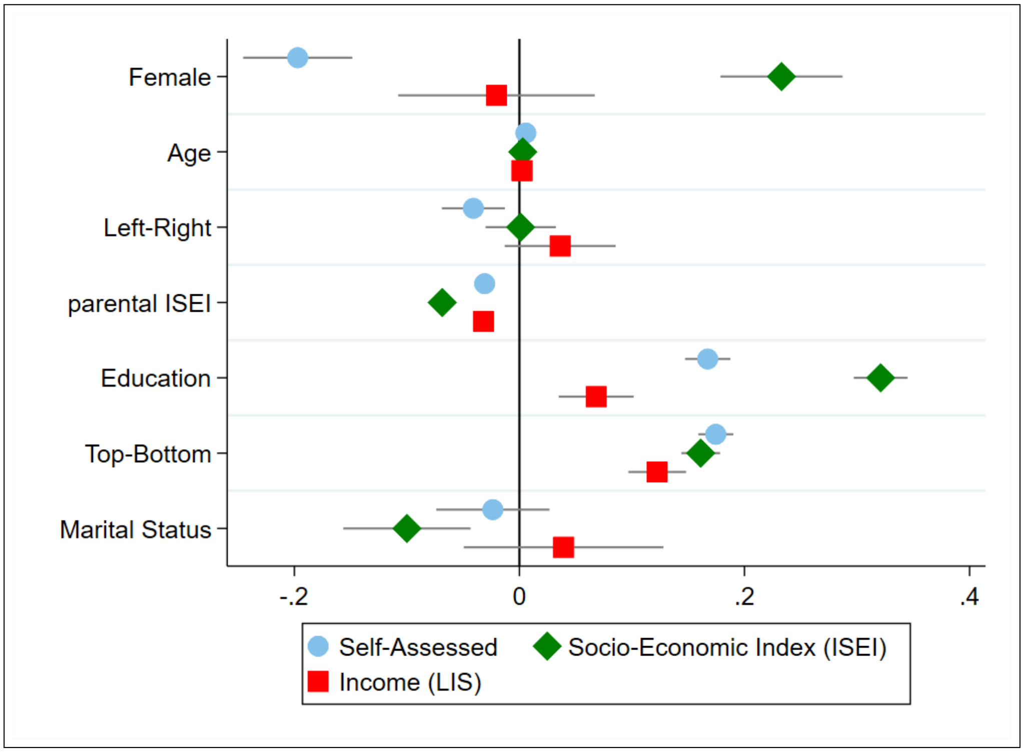
To match the Luxembourg Income Data to ISSP respondents' occupations I retrieved average gross hourly wages for people between the ages of 25 and 55 by ISCO88 occupation type, country and year. Where average gross hourly wages were not available, I used average net hourly wages. Table 3.1d lists the individual LIS waves used by country and year.

VI.2 Additional Analysis of Descriptive Data

VI.2.1 Likelihood of experiencing upward social mobility

Figure 3.2a reports the likelihood of having experienced positive social mobility by basic demographic characteristics, using the three available measures of social mobility experience: self-assessed, socio-economic and income mobility.

Figure 3.2a: Likelihood of a positive social mobility experience by demographic characteristics



As Figure 3.2a illustrates, there are some differences between the three alternative indicators. The three demographic factors which have a uniform and significant relationship with respondents' likelihood of having experienced positive mobility are parental ISEI scores, which are negatively associated with a positive mobility experience, as well as education and self-placement on the income distribution, which are both positively associated with experiencing upward mobility. The direction of the correlation between social mobility experience and these three factors is not surprising.

Interestingly, women are significantly more likely to have experienced upward mobility when using the socio-economic scale but assess themselves to have experienced more negative or stagnating mobility than men. Age and marital status mostly do not appear to matter significantly to the likelihood of having experienced upward mobility.

Political orientation does not significantly differ between those who experienced positive and those who experienced negative or stagnating mobility when looking at the two objective measures. There is a slight but significant negative relationship between the self-assessed measure and political orientation which suggests that those who believe themselves to have experienced negative or stagnating mobility are slightly more left-wing. However, this effect is minimal and only exists for one of the three indicators. This is encouraging for the interpretation of the effect of the left-right indicator on perceptions of social mobility. The other factors which differ between the two groups are controlled for in the main estimation.

VI.2.2 Balance Test of Misinformation

Table 3.2a reports mean values of individual-level characteristics for respondents who are misinformed and correctly informed about the direction of their own mobility experience in the ISSP Cumulative, as well as t-statistics for differences in means. Unsurprisingly, there are a lot of significant differences between the two groups. Given that this dataset is not used to make any causal claims, this is however not a significant issue. The same balance test for the experimental data reported in part C.2 illustrates that almost none of these differences can be found in the experimental data. Additionally, there are no significant differences in perceived social mobility between those who are misinformed and correctly informed about the direction of their own mobility experience in the ISSP Cumulative.

Table 3.2a shows that those who are misinformed are both, more likely to overestimate their own mobility experience with the father and to have a lower ISEI mobility score themselves. While there are some significant differences on the main preference variables of interest, these do not point into a consistent direction - those who misperceive their own mobility are more supportive of redistribution in general but less likely to support more spending on the poor. The demographics show that misperception is not driven by parental ISEI scores but by own ISEI scores - those who are more likely to misperceive have a somewhat lower ISEI score than those who perceive the direction of their mobility experience correctly. Interestingly, there are no party differences but those who misperceive are somewhat less educated and more likely to be women.

Table 3.2a: Balance Test by Misinformation

	ISEI Mobility		
	Correct	Misinformed	t-statistic
Mobility Experience			
Self-assessed	0.47	0.55	-8.43***
	(1.07)	(0.96)	(45,398)
ISEI	8.39	5.42	16.85***
	(20.33)	(15.54)	(45,398)
Income	0.03	0.02	1.05
	(0.73)	(0.62)	(16,430)
Beliefs and Preferences			
SfR (binary)	0.66	0.68	-4.27***
	(0.47)	(0.47)	(44,043)
SfR (ordered)	3.71	3.76	-5.00***
	(1.18)	(1.16)	(44,043)
Inc. Diff too large	0.82	0.83	-2.92***
	(0.39)	(0.38)	(44,494)
More on poor	0.69	0.67	3.91***
	(0.46)	(0.47)	(22,772)
UBI	0.67	0.67	-0.66
	(0.47)	(0.47)	(12,665)
Higher Tax on rich	4.00	4.01	-0.84
	(0.76)	(0.77)	(43,395)
Overall Mobility	47.26	47.15	0.45
	(19.77)	(20.24)	(31,151)
Demographics			
ISEI score	45.75	42.59	20.09***
	(17.09)	(15.72)	(45,398)
Parents' ISEI score	37.36	37.17	1.29
	(16.48)	(14.86)	(45,398)
Party affiliation	2.91	2.91	0.05
	(0.88)	(0.86)	(22,180)
Education	2.86	2.67	14.27***
	(1.44)	(1.41)	(45,091)
Gender	0.49	0.51	-2.90***
	(0.50)	(0.50)	(45,345)
Age	46.38	46.56	-1.20
	(15.60)	(15.73)	(45,237)

Notes: Table reports the mean values for respondents based on whether they perceived the direction of their own mobility experience correctly or not. Definitions of the variables are identical to table 3.3 in the main text. Asterisks indicate significant differences in mean values between samples from a Wald test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

VI.2.3 Alternative Definitions of Mobility Experience

VI.2.3.1. High Mobility Experiences

Tables 3.2b and 3.2c report main results of the ISSP Cumulative survey data for respondents who experienced very high or very low mobility on the self-assessed measure (table 3.2b) and the ISEI mobility score (table 3.2c). As the self-assessed measure ranges from -2 to 2 with higher values indicating more upward mobility relative to the father, table 3.2b simply reports results for those with values of 2 or -2 relative to those with no mobility (a score of 0).

The ISEI mobility score ranges from -72 to 72 with higher values also indicating more positive mobility. For comparison these values are reduced into 5 groups with values also ranging from -2 to 2. ISEI mobility scores within 10% of 0 are labelled as no mobility, values between 10% and 25% are labelled as upward or downward mobility and anything above the 25% threshold is labelled as high downward or upward mobility. Table 3.2c reports results for only those respondents who are above the 25% threshold. In other words, respondents with an ISEI mobility score above 18 or below -18 compared to those with an ISEI mobility score between 7.2 and -7.2.

The results in table 3.2b are consistent with the main results reported in table 3.3: Those who experienced very negative mobility express more support for redistribution on all measures and also perceive social mobility within society as significantly more negative. Those who experienced very positive mobility show no increase in support for redistribution except for one measure - a higher tax share for the rich. This is surprising as this is not the case in the main models.

In the first panel of table 3.2c, none of the coefficients are significant. Neither those who experienced very high upward nor those who experienced very high downward mobility adjust their preferences for redistribution compared to those with no mobility experience. This is mostly consistent with the main results as mobility experience, when not accounting for those who misperceive their own mobility, does not have a consistent significant effect on preferences or beliefs. The second panel of table 3.2c reports the effect of very high downward and upward mobility experience for those who are aware of the direction of their mobility

Table 3.2b: Support for Redistribution - High Self-assessed Mobility

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich	Perception of Social mobility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Self-reported mobility experience							
Very Negative	0.081*	0.173**	0.354***	0.317***	0.332**	0.204**	-3.672***
	(0.083)	(0.069)	(0.074)	(0.099)	(0.142)	(0.086)	(0.799)
	[0.050]	[0.013]	[0.001]	[0.002]	[0.013]	[0.013]	[0.001]
Very Positive	-0.012	0.005	0.003	0.146	-0.007	0.151***	-0.542
	(0.074)	(0.062)	(0.080)	(0.073)	(0.089)	(0.047)	(0.718)
	[1.000]	[1.000]	[1.000]	[0.161]	[1.000]	[0.008]	[1.000]
Control	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Country Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Observations	12,826	12,826	12,960	6,333	3,591	12,821	8,681

Notes: Estimates come from logistic (models (1), (3), (4) and (5)), ordered logit (models (2) and 6)) and linear (model (7)) regressions. Robust standard errors clustered on a country-year level are presented in parentheses. Adjusted p-values for multiple hypothesis testing (Anderson 2008) are presented in brackets. Very positive mobility is a score of 2 and very negative mobility is a score of -2 on the self-assessed mobility scale. All models are relative to respondents with a score of 0 on the self-assessed mobility scale. Controls include the personal ISEI score, the parental score, political orientation, education, gender and age. *** p<0.01, ** p<0.05 , * p<0.1.

Table 3.2c: Support for Redistribution - High ISEI Mobility

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich	Perception of Social mobility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ISEI mobility experience							
Very Negative	0.182	0.158	0.064	-0.095	0.131	0.107	-1.010
	(0.096)	(0.079)	(0.111)	(0.113)	(0.155)	(0.070)	(1.206)
	[0.261]	[0.261]	[0.477]	[0.412]	[0.412]	[0.264]	[0.412]
Very Positive	-0.129	-0.038	0.069	0.056	-0.032	-0.094	1.015
	(0.088)	(0.070)	(0.109)	(0.119)	(0.122)	(0.083)	(1.028)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Control	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Country Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Observations	17,744	17,744	17,938	8,622	4,575	17,717	11,630
ISEI mobility experience (if aware of direction)							
Very Negative	0.341**	0.243**	0.326**	0.037	0.609**	0.327**	-3.275**
	(0.111)	(0.099)	(0.186)	(0.162)	(0.215)	(0.130)	(1.578)
	[0.015]	[0.020]	[0.042]	[0.133]	[0.016]	[0.020]	[0.031]
Very Positive	-0.116	-0.017	-0.063	0.085	-0.256	-0.114	0.516
	(0.114)	(0.097)	(0.169)	(0.155)	(0.217)	(0.102)	(1.470)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Control	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Country Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Observations	8,774	8,774	8,850	4,181	2,402	8,768	5,723

Notes: Estimates come from logistic (models (1), (3), (4) and (5)), ordered logit (models (2) and 6)) and linear (model (7)) regressions. Robust standard errors clustered on a country-year level are presented in parentheses. Adjusted p-values for multiple hypothesis testing (Anderson 2008) are presented in brackets. Very positive mobility is a score of 2 and very negative mobility is a score of -2 on the reduced ISEI mobility scale. All models are relative to respondents with a score of 0 on the reduced ISEI mobility scale. Controls include the personal ISEI score, the parental score, political orientation, education, gender and age. *** p<0.01, ** p<0.05 , * p<0.1.

experience. Consistent with table 3.3 in the main text, mobility experience now has a significant effect on all reported preferences except for more spending on the poor, as well as a significant and negative effect on social mobility perceptions.

VI.2.3.2. Different Definition of ISEI Mobility Experience

While the income mobility measure based on the LIS dataset reported in table 3.3 of the main text did not show any significant effects on preferences or beliefs, this could be due to how the income mobility groups are calculated.

Income mobility is defined as the difference in standardised average earnings between the respondent and the parent with the highest standardised average earnings. The income mobility measure then defines no mobility as being within +/- 5% of the mean standardised average earnings difference. A difference above that as defined as upward mobility and a difference below that is defined as downward mobility. Given that this threshold of +/- 5% of the mean is somewhat arbitrary, table 3.2d reports income mobility models based on the LIS dataset using a +/- 10% threshold to define no mobility. This, effectively, increases the number of respondents who are defined as having experienced no mobility and increases the threshold to define a respondent as having experienced upward or downward mobility.

Again, consistent with the findings in table 3.3 of the main text, upward or downward income mobility has no effect on preferences and beliefs in table 3.2d. Neither those who experienced upward income mobility nor those who experienced downward income mobility show any significant difference in distributive preferences and mobility beliefs compared to those who experienced no income mobility based on this measure.

Table 3.2d: Support for Redistribution - LIS Income Measure

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich	Perception of Social mobility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income mobility experience							
Negative	-0.008	-0.057	-0.011	0.028	0.069	-0.016	0.102
	(0.075)	(0.066)	(0.089)	(0.068)	(0.123)	(0.076)	(0.835)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Positive	0.100	0.079	0.024	-0.090	0.455	0.014	-0.533
	(0.111)	(0.106)	(0.133)	(0.070)	(0.266)	(0.074)	(0.638)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Control	✓	✓	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Country Fixed Effects	✓	✓	✓	✓	✓	✓	✓
Observations	26,056	26,056	26,360	12,823	6,902	26,027	17,400

Notes: Estimates come from logistic (models (1), (3), (4) and (5)), ordered logit (models (2) and 6)) and linear (model (7)) regressions. Robust standard errors clustered on a country-year level are presented in parentheses. Adjusted p-values for multiple hypothesis testing (Anderson 2008) are presented in brackets. The income mobility measure defines no mobility as being within +/- 10% of the mean standardised average earnings difference. Controls include the personal ISEI score, the parental score, political orientation, education, gender and age. *** p<0.01, ** p<0.05 , * p<0.1.

VI.3 Additional Analysis of Experimental Data

VI.3.1 Placebo Test

To account for the possibility that simply under- or overestimating something causes some negative reaction that might affect preferences, subjects in the control group of the survey experiment were given a placebo treatment. During the demographics part of the experiment, subjects were asked how long they believe the difference in length between the longest river in North America, the Missouri river, and the fifth longest river in North America, the Arkansas river to be. If randomly assigned to the placebo treatment, subjects were then told whether you objectively under-, over-, or correctly estimated the difference in length between the two rivers. Overall, there were 797 subjects who underestimated, 303 who overestimated and 19 who correctly estimated the difference in length of the two rivers, allowing for a margin of error of +/- 15 miles. After a first initial pilot using two different rivers, I changed the placebo to the Missouri and the Arkansas river as I wanted to ensure that a significant enough number of subjects would underestimate the difference (given that the group I am primarily interested in for the main estimations are the under-estimators).

Table 3.3a and table 3.3b report the results for the placebo group. As none of the coefficients reach conventional levels of significance, the placebo test suggests that simply under- or overestimating something does not affect the outcome variables of interest.

Table 3.3a: Placebo: Support for Redistribution

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo Treatment						
Underestimate	-0.199	-0.159*	-0.012	0.128	-0.122	-0.040
	(0.147)	(0.091)	(0.080)	(0.079)	(0.098)	(0.028)
Overestimate	-0.041	0.096	0.041	-0.003	0.106	-0.015
	(0.210)	(0.125)	(0.117)	(0.108)	(0.129)	(0.039)
Controls	✓	✓	✓	✓	✓	✓
Observations	895	895	896	873	890	897

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are placebo treatment effects relative to comparable subjects in the treatment group. Robust standard errors are presented in parentheses. The analysis is restricted to subjects who indicated that they believed the provided information. *** p<0.01, ** p<0.05 , * p<0.1.

Table 3.3b: Placebo: Mobility and Societal Perceptions

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
Underestimate	-0.120	-0.059	0.044	0.035
	(0.087)	(0.097)	(0.092)	(0.215)
Overestimate	-0.027	0.094	0.179	-0.016
	(0.119)	(0.129)	(0.125)	(0.301)
Controls	✓	✓	✓	✓
Observations	886	887	886	842

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are placebo treatment effects relative to comparable subjects in the treatment group. The analysis is restricted to subjects who indicated that they believed the provided information. *** p<0.01, ** p<0.05 , * p<0.1.

VI.3.2 Individual-level characteristics by mobility experience

Table 3.3c reports mean values of individual-level characteristics for subjects who experienced an effective downward mobility shock as opposed to an upward mobility shock during the survey experiment, as well as t-statistics for differences in means. Importantly, these mean values include subjects who were in the treatment as well as in the control groups. In other words, the values include subjects who were informed and not informed of the respective shocks. The two measures reported are ISEI and income mobility shocks as defined in table 3.4 in the main text.

While there are obvious differences in the variables used to generate the mobility measures between those who experienced upward and downward shocks, there is a striking lack of significant differences in any of the other variables. The only variable not used for the generation of the measures with significant differences is the personal benefit variable. Here, those who experienced an upward mobility shock on both measures are more likely to think that they personally benefit from redistribution. This may be due to the fact that those in control and treatment group are included in the mean values and those who experienced an upward shock are more likely to believe themselves to be worse off relative to others prior to receiving the treatment.

The differences in relative and household income between those who experienced an upward- and downward-mobility shock, while seemingly counter-intuitive at first, merely reflect that those that considered themselves to be worse off prior to receiving the treatment are more likely to experience an upward mobility shock when being informed of their objective mobility experience.

Table 3.3d reports mean values of individual-level characteristics for subjects who experienced any mobility shock as opposed to no mobility shock during the survey experiment, as well as t-statistics for differences in means. Importantly, these mean values include subjects who were in the treatment as well as in the control groups. In other words, the values include subjects who were informed and not informed of the respective shocks. The two measures reported are ISEI and income mobility shocks as defined in table 3.4 in the main text.

Table 3.3c: Balance Test by Mobility Shock

	ISEI Mobility			Income Mobility		
	Upward Shock	Downward Shock	t-statistic	Upward Shock	Downward Shock	t-statistic
ISEI Data						
Treatment assignment	0.50	0.50	0.00	0.53	0.54	-0.23
	(0.51)	(0.51)	(762)	(0.50)	(0.50)	(426)
Mobility experience	9.49	-13.67	16.87***	16.38	-12.75	21.04***
	(20.92)	(15.11)	(755)	(12.55)	(15.82)	(419)
ISEI	55.76	45.44	7.67***	62.55	45.23	11.23***
	(19.86)	(15.96)	(755)	(13.28)	(18.35)	(419)
father's ISEI	40.23	50.19	-6.12***	40.68	48.48	-3.61***
	(22.69)	(21.38)	(755)	(21.30)	(22.67)	(419)
mother's ISEI	29.33	40.74	-6.59***	29.77	38.15	-3.63***
	(23.10)	(24.14)	(755)	(22.49)	(24.57)	(419)
Perceptions						
seSM with father	-0.58	0.41	-12.24***	0.00	0.21	-1.73*
	(1.09)	(1.02)	(687)	(1.11)	(1.20)	(391)
seSM with mother	0.258	0.435	-1.93*	0.57	0.52	0.39
	(1.18)	(1.13)	(646)	(1.11)	(1.15)	(373)
Relative past income	-0.01	-0.07	0.88	0.27	-0.51	10.25***
	(0.91)	(0.85)	(762)	(0.78)	(0.76)	(426)
Relative income	-0.16	0.15	-4.50***	-0.24	0.70	-11.66***
	(0.96)	(0.86)	(762)	(0.89)	(0.72)	(426)
Demographics						
Age	35.65	35.48	0.32	36.11	35.17	1.30
	(7.62)	(7.19)	(762)	(7.51)	(7.26)	(426)
Gender	0.49	0.52	-0.81	0.56	0.44	2.31**
	(0.50)	(0.50)	(753)	(0.50)	(0.50)	(422)
Education	2.97	2.94	0.42	3.19	2.92	3.03***
	(0.92)	(0.89)	(762)	(0.92)	(0.92)	(426)
Household Income	6.28	6.79	-2.60***	6.57	7.47	-3.59***
	(2.71)	(2.59)	(749)	(2.47)	(2.58)	(418)
Party affiliation	1.22	1.21	0.23	1.24	1.22	0.31
	(0.41)	(0.41)	(609)	(0.43)	(0.42)	(328)
Beliefs and Preferences						
SfR (binary)	0.56	0.55	0.20	0.54	0.54	-0.01
	(0.50)	(0.50)	(757)	(0.50)	(0.50)	(424)
SfR (ordered)	0.40	0.29	1.09	0.30	0.30	-0.06
	(1.30)	(1.29)	(757)	(1.25)	(1.33)	(424)
Inc. Diff too large	1.14	1.17	-0.34	1.06	1.16	-0.91
	(1.16)	(1.05)	(759)	(1.16)	(1.05)	(424)
More on poor	1.21	1.17	0.43	1.21	1.16	0.43
	(1.08)	(1.08)	(739)	(1.02)	(1.08)	(419)
UBI	0.52	0.46	0.66	0.42	0.30	0.83
	(1.37)	(1.35)	(755)	(1.35)	(1.39)	(424)
Higher Tax on rich	1.20	1.15	0.90	1.12	1.05	0.92
	(0.76)	(0.78)	(752)	(0.75)	(0.89)	(418)
Mob. lowest quintile	-0.57	-0.57	0.02	-0.58	-0.51	-0.59
	(1.21)	(1.17)	(752)	(1.15)	(1.18)	(422)
Overall mobility	-0.33	-0.31	-0.15	-0.24	-0.25	0.14
	(1.30)	(1.29)	(750)	(1.32)	(1.32)	(421)
Diff. due to effort	-0.60	-0.59	-0.07	-0.59	-0.60	0.07
	(1.26)	(1.21)	(752)	(1.21)	(1.22)	(423)
Personal benefit	3.87	3.29	2.72***	3.91	3.08	2.97***
	(2.90)	(2.70)	(704)	(2.79)	(2.65)	(392)

Notes: The table reports mean values for subjects based on their experienced mobility shock during the experiment, irrespective of treatment assignment. Definitions of the variables are identical to tables 3.4 and 3.5 in the main text. Asterisks indicate significant differences in mean values between samples from a Wald test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Table 3.3d: Balance Test by Mobility Shock vs. No Shock

	ISEI Mobility			Income Mobility		
	Shock	No Shock	t-statistic	Shock	No Shock	t-statistic
ISEI Data						
Treatment assignment	0.50	0.50	0.09	0.51	0.47	-1.05
	(0.50)	(0.50)	(1,102)	(0.50)	(0.50)	(1,102)
Mobility experience	-0.36	-5.38	-3.41***	-2.27	-0.69	0.96
	(21.89)	(24.01)	(1,095)	(23.73)	(18.51)	(1,095)
ISEI	51.37	51.24	-0.11	50.89	52.88	1.48
	(19.00)	(17.67)	(1,095)	(18.82)	(17.68)	(1,095)
father's ISEI	44.46	51.75	5.05***	46.55	47.32	0.47
	(22.67)	(20.88)	(1,095)	(22.65)	(21.42)	(1,095)
mother's ISEI	34.18	34.46	0.18	33.98	35.27	0.73
	(24.20)	(24.54)	(1,095)	(24.13)	(24.88)	(1,095)
Perceptions						
seSM with father	-0.12	-0.26	-1.66*	-0.18	-0.10	0.83
	(1.17)	(1.43)	(1,029)	(1.25)	(1.30)	(1,029)
seSM with mother	0.34	0.33	-0.16	0.33	0.36	0.30
	(1.16)	(1.35)	(944)	(1.23)	(1.22)	(944)
Relative past income	-0.04	0.10	2.36**	-0.01	0.07	1.31
	(0.88)	(0.92)	(1,104)	(0.89)	(0.93)	(1,104)
Relative income	-0.03	0.01	0.61	-0.026	0.012	0.55
	(0.93)	(1.00)	(1,104)	(0.96)	(0.93)	(1,104)
Demographics						
Age	35.58	35.74	0.34	35.50	36.09	1.09
	(7.44)	(7.37)	(1,104)	(7.42)	(7.38)	(1,104)
Gender	0.50	0.46	-1.27	0.49	0.49	0.00
	(0.50)	(0.50)	(1,090)	(0.50)	(0.50)	(1,090)
Education	2.95	3.07	2.05**	2.98	3.02	0.67
	(0.91)	(0.85)	(1,104)	(0.90)	(0.86)	(1,104)
Household Income	6.49	6.47	-0.11	6.48	6.52	0.23
	(2.67)	(2.63)	(1,085)	(2.64)	(2.71)	(1,085)
Party affiliation	1.22	1.23	0.39	1.21	1.24	0.66
	(0.41)	(0.42)	(881)	(0.41)	(0.43)	(881)
Beliefs and Preferences						
SfR (binary)	0.56	0.53	-0.69	0.56	0.53	-0.79
	(0.50)	(0.50)	(1,098)	(0.50)	(0.50)	(1,098)
SfR (ordered)	0.35	0.27	-1.02	0.33	0.30	-0.40
	(1.30)	(1.23)	(1,098)	(1.27)	(1.29)	(1,098)
Inc. Diff too large	1.16	1.16	0.04	1.17	1.11	-0.67
	(1.11)	(1.09)	(1,101)	(1.10)	(1.12)	(1,101)
More on poor	1.19	1.17	0.37	1.22	1.05	2.10**
	(1.08)	(1.06)	(1,072)	(1.05)	(1.14)	(1,072)
UBI	0.49	0.41	-0.91	0.47	0.46	-0.17
	(1.36)	(1.32)	(1,093)	(1.35)	(1.34)	(1,093)
Higher Tax on rich	1.18	1.11	-1.50	1.16	1.16	0.08
	(0.77)	(0.77)	(1,083)	(0.77)	(0.76)	(1,083)
Mob. lowest quintile	-0.57	-0.47	1.27	-0.58	-0.40	2.04**
	(1.19)	(1.16)	(1,091)	(1.16)	(1.25)	(1,091)
Overall mobility	-0.32	-0.24	0.98	-0.34	-0.14	2.11**
	(1.30)	(1.30)	(1,091)	(1.29)	(1.31)	(1,091)
Diff. due to effort	-0.60	-0.59	0.08	-0.63	-0.47	1.71*
	(1.24)	(1.24)	(1,090)	(1.24)	(1.23)	(1,090)
Personal benefit	3.61	3.99	1.96*	3.74	3.70	-0.22
	(2.82)	(2.92)	(1,027)	(2.83)	(2.94)	(1,027)

Notes: The table reports mean values for subjects based on whether they received a mobility shock during the experiment or not. Definitions of the variables are identical to tables 3.4 and 3.5 in the main text. Asterisks indicate significant differences in mean values between samples from a Wald test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

As in table 3.3c most of the variables where significant differences in means can be observed are those variables used to generate the shock variables. Using the ISEI mobility measure, none of the other variables show significant differences between those who experienced no shock and those who experienced either a positive or negative shock. A few more differences are observable when using the income mobility measure. Specifically, those who experienced no mobility shock are significantly less likely to support less government spending on the poor, have a slightly more negative view of the mobility of the lowest quintile in society and slightly lower overall perceived mobility. These differences make the main finding, that those who experience a negative mobility shock and are informed of that shock reduce their mobility perception and certain preferences, even more striking.

VI.3.3 Balance Test by Treatment Assignment

Table 3.3e reports mean values of individual-level characteristics by random treatment assignment, as well as t-statistics for differences in means. None of the variables reported show significant differences between the treatment and control group, suggesting that the random assignment was successful.

VI.3.4 Information Provision Tests

Table 3.3f reports mean values of individual-level characteristics grouped by whether the subject believed the information provided during the experiment or not, as well as t-statistics for differences in means. There is, strikingly, only one variable which shows significant differences in means: treatment assignment. Maybe somewhat unsurprisingly, subjects who were randomly assigned to be in the treatment as opposed to the control group are more likely to state that they do not believe the information provided. Given that the treatment information directly relates to the personal experiences of subjects while the placebo information does not, this is a reasonable difference. Given that there are no other significant differences between the two groups, there do not appear to be fundamental differences between those who believed the information and those who did not (which are excluded in the main analysis).

Table 3.3e: Balance Test by Treatment Assignment

	Treatment Group	Control Group	t-statistic
ISEI Data			
Mobility experience	-1.72	-2.10	-0.28
	(23.16)	(22.24)	(1,093)
ISEI	51.39	51.22	-0.15
	(19.01)	(18.18)	(1,093)
father's ISEI	46.36	47.00	0.47
	(22.65)	(22.11)	(1,093)
mother's ISEI	35.45	33.21	-1.52
	(23.61)	(24.89)	(1,093)
Perceptions			
seSM with father	-0.152	-0.174	-0.28
	(1.29)	(1.24)	(1,027)
seSM with mother	0.33	0.34	0.09
	(1.21)	(1.25)	(944)
Relative past income	0.02	-0.01	-0.64
	(0.88)	(0.91)	(1,102)
Relative income	0.01	-0.04	-0.79
	(0.96)	(0.94)	(1,102)
Demographics			
Age	35.37	35.88	1.13
	(7.11)	(7.71)	(1,102)
Gender	0.51	0.47	-1.52
	(0.50)	(0.50)	(1,088)
Education	3.00	2.98	-0.30
	(0.87)	(0.92)	(1,102)
Household Income	6.54	6.43	-0.65
	(2.68)	(2.63)	(1,083)
Party affiliation	1.23	1.21	-0.59
	(0.42)	(0.41)	(880)
Beliefs and Preferences			
SfR (binary)	0.57	0.53	-1.45
	(0.50)	(0.50)	(1,096)
SfR (ordered)	0.37	0.28	-1.29
	(1.28)	(1.27)	(1,096)
Inc. Diff too large	1.18	1.13	-0.75
	(1.05)	(1.15)	(1,099)
More on poor	1.23	1.14	1.40
	(1.05)	(1.10)	(1,070)
UBI	0.50	0.44	-0.64
	(1.35)	(1.34)	(1,091)
Higher Tax on rich	1.18	1.14	-0.90
	(0.75)	(0.79)	(1,082)
Mob. lowest quintile	-0.59	-0.49	1.47
	(1.16)	(1.20)	(1,089)
Overall mobility	-0.32	-0.28	0.50
	(1.30)	(1.30)	(1,089)
Diff. due to effort	-0.57	-0.62	-0.59
	(1.27)	(1.20)	(1,088)
Personal benefit	3.75	3.71	-0.19
	(2.89)	(2.83)	(1,025)

Notes: Table reports the mean values for subjects based on their treatment assignment. Definitions of the variables are identical to tables 3.4 and 3.5 in the main text. Asterisks indicate significant differences in mean values between samples from a Wald test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3.3f: Balance Test by Info Belief

	Yes	No	t-statistic
ISEI Data			
Treatment assignment	0.50	0.73	3.24**
	(0.50)	(0.45)	(944)
Mobility experience	-2.03	-0.12	0.57
	(22.95)	(18.06)	(937)
ISEI	51.84	52.33	0.18
	(18.59)	(18.01)	(937)
father's ISEI	47.15	47.22	0.02
	(22.30)	(26.02)	(937)
mother's ISEI	34.63	36.14	0.42
	(24.24)	(26.57)	(937)
Perceptions			
seSM with father	-0.16	0.02	0.91
	(1.26)	(1.26)	(881)
seSM with mother	0.34	0.46	0.61
	(1.22)	(1.31)	(810)
Relative past income	0.02	0.00	-0.16
	(0.89)	(1.00)	(945)
Relative income	-0.01	0.31	2.32
	(0.94)	(0.94)	(945)
Demographics			
Age	35.46	37.02	1.44
	(7.39)	(7.35)	(945)
Gender	0.50	0.38	-1.59
	(0.50)	(0.49)	(931)
Education	3.00	3.08	0.59
	(0.90)	(1.06)	(945)
Household Income	6.49	7.04	1.41
	(2.64)	(2.87)	(932)
Party affiliation	1.21	1.25	0.55
	(0.41)	(0.44)	(772)
Beliefs and Preferences			
SfR (binary)	0.57	0.49	-1.05
	(0.50)	(0.51)	(941)
SfR (ordered)	0.39	0.13	-1.37
	(1.25)	(1.48)	(941)
Inc. Diff too large	1.18	1.11	-0.45
	(1.09)	(1.17)	(942)
More on poor	1.23	1.02	1.36
	(1.05)	(1.31)	(920)
UBI	0.51	0.50	-0.04
	(1.32)	(1.47)	(937)
Higher Tax on rich	1.17	1.10	-0.58
	(0.76)	(0.95)	(929)
Mob. lowest quintile	-0.54	-0.50	0.21
	(1.18)	(1.43)	(933)
Overall mobility	-0.29	-0.27	0.10
	(1.30)	(1.65)	(934)
Diff. due to effort	-0.61	-0.44	0.95
	(1.22)	(1.50)	(933)
Personal benefit	3.79	3.56	-0.54
	(2.83)	(2.94)	(886)

Notes: Table reports the mean values for subjects based on whether subjects believed the information provided or not. Definitions of the variables are identical to tables 3.4 and 3.5 in the main text. Asterisks indicate significant differences in mean values between samples from a Wald test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Apart from whether or not subjects believed the provided information, how careful the information was read may also influence treatment effects. While this cannot directly be measured, I can measure the time subjects spent on the treatment and placebo screens as a proxy for attention paid to the provided information. The average time spent was 14.5 seconds. Table 3.4 in the main text reports main treatment effects for subjects who spent more than 8 seconds on the treatment and placebo screens. While this length is somewhat arbitrary, it is roughly enough time to carefully read through the provided paragraph. Table 3.3g below reports treatment effects for all subjects, irrespective of the time spent on the treatment and placebo screens. The results are consistent with the findings reported in table 3.4 although more noisy. On both measures included in the table, those who experienced a downward mobility shock show again significantly more support for more spending on the poor and for higher taxes on the rich.

Table 3.3g: Experiment: Support for Redistribution

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
Upward mobility	0.305	0.255*	0.042	0.110	0.209*	0.056
	(0.215)	(0.100)	(0.089)	(0.084)	(0.111)	(0.031)
	[0.169]	[0.071]	[0.294]	[0.169]	[0.137]	[0.137]
No mobility	0.289	0.060	0.074	0.101	0.022	0.005
	(0.219)	(0.104)	(0.085)	(0.097)	(0.110)	(0.037)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Downward mobility	0.321	0.215	0.158	0.220**	0.036	0.094**
	(0.244)	(0.125)	(0.094)	(0.091)	(0.131)	(0.035)
	[0.163]	[0.103]	[0.103]	[0.045]	[0.294]	[0.038]
Controls	✓	✓	✓	✓	✓	✓
Observations	817	808	809	787	804	810
Treatment						
Upward income mobility	-0.012	0.066	0.031	0.093	0.087	0.030
	(0.284)	(0.131)	(0.120)	(0.109)	(0.144)	(0.046)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
No income mobility	0.017	0.063	-0.158	-0.133	0.179	0.069
	(0.545)	(0.240)	(0.183)	(0.226)	(0.259)	(0.082)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Downward income mobility	0.340	0.307*	0.219*	0.334**	0.082	0.164***
	(0.331)	(0.167)	(0.127)	(0.126)	(0.185)	(0.050)
	[0.149]	[0.095]	[0.095]	[0.021]	[0.282]	[0.007]
Controls	✓	✓	✓	✓	✓	✓
Observations	369	369	369	362	366	369

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Robust standard errors are presented in parentheses. Adjusted p-values are not reported for these estimations as the sample size is quite small compared to the descriptive data. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information. *** p<0.01, ** p<0.05 , * p<0.1.

Table 3.3h reports main treatment effects on the mobility perception variables for subjects irrespective of the time spent on the treatment and placebo screens. Here, the results are also consistent with those reported in 3.5 of the main text where subjects who spent less than 8 seconds on the treatment and placebo screens are excluded. There is a significant negative effect on overall perceived mobility for those who experienced a downward mobility shock on both measures. Additionally, those who experienced a downward mobility shock are now also perceiving the mobility of the lowest quintile significantly more negatively.

Table 3.3h: Experiment: Mobility and societal perceptions

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
Upward mobility	-0.105	-0.037	-0.013	-0.019
	(0.101)	(0.113)	(0.104)	(0.261)
No mobility	-0.082	-0.014	0.056	0.439
	(0.095)	(0.107)	(0.107)	(0.278)
Downward mobility	-0.330***	-0.338***	-0.096	0.232
	(0.121)	(0.129)	(0.129)	(0.298)
Controls	✓	✓	✓	✓
Observations	808	800	799	763
Treatment				
Upward income mobility	-0.101	-0.045	0.076	0.001
	(0.131)	(0.144)	(0.135)	(0.333)
No income mobility	0.242	0.147	0.088	0.418
	(0.279)	(0.232)	(0.231)	(0.633)
Downward income mobility	-0.193	-0.477***	-0.055	-0.292
	(0.159)	(0.182)	(0.163)	(0.416)
Controls	✓	✓	✓	✓
Observations	365	365	366	347

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information. *** p<0.01, ** p<0.05 , * p<0.1.

VI.3.5 Main analysis with alternative definitions of Mobility Experience

VI.3.5.1. Extreme mobility shocks

Table 3.3i reports main treatment effects for subjects who experienced extreme mobility shocks during the experiment. This is defined as either a mobility shock score of -2 and less or +2 and more. Given that restricting the models to subjects with such extreme values significantly reduces the sample size, unsurprisingly, almost none of the results remain when accounting for multiple hypothesis testing. Only on the income mobility measure do the main results survive, although only with weak significance: those who experienced a downward shock significantly increase their support for more spending on the poor and higher taxes on the rich.

Table 3.3i: Experiment: Support for Redistribution

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
Upward mobility	0.511	0.281	0.099	0.276	0.206	0.020
	(0.353)	(0.167)	(0.151)	(0.130)	(0.191)	(0.051)
	[0.325]	[0.308]	[0.445]	[0.257]	[0.393]	[0.534]
No mobility	0.467	0.089	0.082	0.163	0.078	0.001
	(0.240)	(0.114)	(0.093)	(0.106)	(0.122)	(0.041)
	[0.441]	[0.725]	[0.725]	[0.455]	[0.725]	[1.000]
Downward mobility	0.495	0.319	0.324	0.233	0.115	0.053
	(0.393)	(0.197)	(0.149)	(0.157)	(0.198)	(0.065)
	[0.350]	[0.302]	[0.220]	[0.302]	[0.390]	[0.386]
Controls	✓	✓	✓	✓	✓	✓
Observations	475	468	468	457	465	469
Treatment						
Upward income mobility	0.187	0.175	0.041	0.065	0.092	0.142
	(0.464)	(0.218)	(0.210)	(0.196)	(0.238)	(0.063)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[0.194]
No income mobility	0.100	0.052	-0.189	-0.083	0.162	0.067
	(0.566)	(0.249)	(0.195)	(0.237)	(0.266)	(0.087)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Downward income mobility	0.064	0.149	0.102	0.373*	-0.061	0.156*
	(0.426)	(0.222)	(0.169)	(0.162)	(0.251)	(0.061)
	[1.000]	[1.000]	[1.000]	[0.071]	[1.000]	[0.071]
Controls	✓	✓	✓	✓	✓	✓
Observations	193	193	193	190	190	193

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Robust standard errors are presented in parentheses. Adjusted p-values are not reported for these estimations as the sample size is quite small compared to the descriptive data. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** p<0.01, ** p<0.05 , * p<0.1.

The fact that the results for those who experienced extreme mobility shocks during the experiment are not vastly different to the main results, merely somewhat less significant, is however encouraging.

Table 3.3j reports main treatment effects on the mobility perception variables for subjects who experienced extreme mobility shocks during the experiment. The results are very similar to those reported in 3.5 of the main text: On both measures, those who experienced downward mobility shocks have a significantly more negative overall perception of mobility. In addition, on the first measure, those subjects also have a significantly more negative perception of the mobility of the lowest quintile.

Table 3.3j: Experiment: Mobility and societal perceptions

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
Upward mobility	0.041	0.072	0.099	-0.315
	(0.167)	(0.182)	(0.170)	(0.403)
No mobility	-0.110	-0.089	0.034	0.262
	(0.107)	(0.119)	(0.119)	(0.303)
Downward mobility	-0.466***	-0.573***	-0.270	-0.416
	(0.174)	(0.198)	(0.206)	(0.469)
Controls	✓	✓	✓	✓
Observations	469	466	463	448
Treatment				
Upward income mobility	-0.053	0.011	0.267	-0.417
	(0.217)	(0.250)	(0.239)	(0.503)
No income mobility	0.368	0.273	0.124	0.209
	(0.300)	(0.260)	(0.239)	(0.657)
Downward income mobility	-0.196	-0.631**	0.066	-0.271
	(0.210)	(0.251)	(0.235)	(0.558)
Controls	✓	✓	✓	✓
Observations	181	178	179	173

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** p<0.01, ** p<0.05 , * p<0.1.

VI.3.5.2. Main treatment effects using mobility with parents

A strong assumption made for the first measure of the models reported in table 3.4 and 3.5 in the main text is that subjects will use their self-reported mobility with their father as a pre-treatment reference point. All the models using the first measure in the main text therefore report mobility shocks based on item Q1 from the survey instrument. It is however not unlikely that subjects will use a combination of Q1 and Q2 (mobility relative to the mother) to assess their self-assessed mobility. As self-assessed mobility with the mother was not asked during the ISSP Cumulative, I have only used mobility with the father in the main text. Tables 3.3k to 3.3n below report the main treatment effects using two alternative measures of self-assessed mobility to calculate the mobility shock: First, the maximum of Q1 and Q2 and second the minimum of Q1 and Q2. In other words, tables 3.3k and 3.3l report the results using the parent who the subject believes themselves to have experienced the most mobility in comparison to and tables 3.3m and 3.3n report the results using the parent who the subject believes themselves to have experienced the least mobility in comparison to.

Table 3.3k: Support for Redistribution - Maximum Mobility Measure

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
Upward mobility	0.570*	0.422**	0.082	0.179*	0.391**	0.047
	(0.307)	(0.136)	(0.122)	(0.102)	(0.147)	(0.046)
	[0.087]	[0.013]	[0.201]	[0.087]	[0.021]	[0.138]
No mobility	-0.146	0.020	0.159	0.195	0.078	0.005
	(0.311)	(0.151)	(0.119)	(0.126)	(0.152)	(0.052)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Downward mobility	0.341	0.178	0.142	0.112	-0.018	0.085
	(0.241)	(0.119)	(0.091)	(0.100)	(0.129)	(0.035)
	[0.245]	[0.245]	[0.245]	[0.309]	[0.421]	[0.107]
Controls	✓	✓	✓	✓	✓	✓
Observations	596	592	593	582	590	593

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Robust standard errors are presented in parentheses. Adjusted p-values are not reported for these estimations as the sample size is quite small compared to the descriptive data. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds.. *** p<0.01, ** p<0.05 , * p<0.1.

Looking first at the models using the maximum mobility value of both parents, the only significant effects that remain after accounting for multiple hypothesis testing are that those who experienced an upward mobility shock based on this alternative specification are more likely to support redistribution and Universal Basic Income. This is inconsistent with any of the other results reported in the main text and appendix. Interestingly, these findings from table 3.3k also do not match the null-results in table 3.3l. In other words, it seems that these effects are not driven by particular changes in mobility perceptions or perceived personal benefit from redistribution. It may be worth noting that this alternative measure is more likely to pick up on perceived mobility with the mother as many mothers of subjects in the experiment did not work (26%). It may therefore pick up on a difference between subjects whose mothers are or were part of the workforce and those who were/are not. What drives these effects when using the maximum mobility score of the parents may therefore be an avenue for future research.

Table 3.3l: Mobility and societal perceptions - Maximum Mobility Measure

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
Upward mobility	-0.153	-0.055	-0.046	0.414
	(0.149)	(0.159)	(0.155)	(0.383)
No mobility	-0.092	-0.108	0.077	0.618
	(0.140)	(0.157)	(0.159)	(0.387)
Downward mobility	-0.164	-0.127	-0.018	0.107
	(0.118)	(0.126)	(0.123)	(0.297)
Controls	✓	✓	✓	✓
Observations	590	588	586	558

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** p<0.01, ** p<0.05 , * p<0.1.

Tables 3.3m and 3.3n report models using the minimum mobility value of both parents. One of the two main results survives here (those who experienced a downward mobility shock are more likely to support higher taxes on the rich) and no other coefficients are significant. This is consistent with the main results in table 3.4. Equally, a downward mobility shock significantly reduces the perception of overall mobility in society, although only weakly using this measure. Again, these results are consistent with those in table 3.5 in the main text.

Table 3.3m: Support for Redistribution - Minimum Mobility Measure

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
Upward mobility	0.116	0.237	0.124	0.150	0.159	0.033
	(0.259)	(0.125)	(0.102)	(0.094)	(0.132)	(0.040)
	[0.527]	[0.513]	[0.513]	[0.513]	[0.513]	[0.513]
No mobility	0.201	0.124	0.162	0.119	0.072	0.026
	(0.268)	(0.123)	(0.102)	(0.115)	(0.130)	(0.043)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Downward mobility	0.643	0.276	0.089	0.201	0.119	0.132**
	(0.321)	(0.168)	(0.120)	(0.119)	(0.181)	(0.045)
	[0.127]	[0.145]	[0.206]	[0.145]	[0.206]	[0.019]
Controls	✓	✓	✓	✓	✓	✓
Observations	596	592	593	582	590	593

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Robust standard errors are presented in parentheses. Adjusted p-values are not reported for these estimations as the sample size is quite small compared to the descriptive data. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds.. *** p<0.01, ** p<0.05 , * p<0.1.

Table 3.3n: Mobility and societal perceptions - Minimum Mobility Measure

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
Upward mobility	-0.094	-0.005	-0.062	0.430
	(0.127)	(0.134)	(0.126)	(0.329)
No mobility	-0.132	-0.077	0.064	0.341
	(0.116)	(0.134)	(0.141)	(0.335)
Downward mobility	-0.236	-0.297*	-0.002	0.117
	(0.170)	(0.172)	(0.169)	(0.380)
Controls	✓	✓	✓	✓
Observations	590	588	586	558

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** p<0.01, ** p<0.05 , * p<0.1.

VI.3.5.3. Main treatment effects using mobility information

Table 3.3o: Support for Redistribution - Information Effects

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
Upward mobility	0.264	0.207	0.130	0.076	0.181	0.017
	(0.263)	(0.124)	(0.109)	(0.102)	(0.134)	(0.042)
	[0.873]	[0.873]	[0.873]	[0.873]	[0.873]	[0.873]
No mobility	0.484	0.347	0.133	0.173	0.074	0.046
	(0.431)	(0.183)	(0.134)	(0.170)	(0.201)	(0.060)
	[0.671]	[0.534]	[0.671]	[0.671]	[0.929]	[0.799]
Downward mobility	0.209	0.151	0.128	0.211*	0.078	0.091*
	(0.243)	(0.125)	(0.094)	(0.094)	(0.129)	(0.037)
	[0.457]	[0.294]	[0.294]	[0.085]	[0.517]	[0.085]
Controls	✓	✓	✓	✓	✓	✓
Observations	596	592	593	582	590	593

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Robust standard errors are presented in parentheses. Adjusted p-values are not reported for these estimations as the sample size is quite small compared to the descriptive data. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds.. *** p<0.01, ** p<0.05 , * p<0.1.

Table 3.3p: Mobility and societal perceptions - Information Effects

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
Upward mobility	0.030	0.074	0.056	0.405
	(0.126)	(0.131)	(0.133)	(0.313)
No mobility	-0.003	-0.054	-0.101	0.165
	(0.194)	(0.202)	(0.168)	(0.475)
Downward mobility	-0.349***	-0.277**	-0.017	0.303
	(0.117)	(0.131)	(0.134)	(0.320)
Controls	✓	✓	✓	✓
Observations	590	588	586	558

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** p<0.01, ** p<0.05 , * p<0.1.

While the models used to report the main treatment effects in tables 3.4 and 3.5 of the main text look at the difference between pre-treatment beliefs about own mobility experience and treatment information, another way of measuring treatment effects is by simply looking at the effects of receiving a certain treatment information on its own. This is what I report in tables 3.3o to 3.3r. Tables 3.3o and 3.3p look at overall information effects of receiving a positive, negative or neutral information about one's own mobility experience compared to someone in the control group with the same mobility experience. Tables 3.3q and 3.3r report the same but only for subjects who received very positive or very negative mobility information.

Table 3.3q: Support for Redistribution - High Information Effects

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
Upward mobility	0.192	0.237	0.093	0.117	0.233	0.052
	(0.394)	(0.191)	(0.195)	(0.162)	(0.214)	(0.064)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
No mobility	0.414	0.359	0.135	0.188	0.053	0.044
	(0.431)	(0.187)	(0.145)	(0.179)	(0.210)	(0.064)
	[0.790]	[0.507]	[0.790]	[0.790]	[1.000]	[0.957]
Downward mobility	0.469	0.186	0.039	0.252	0.108	0.021
	(0.352)	(0.187)	(0.133)	(0.147)	(0.201)	(0.058)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Controls	✓	✓	✓	✓	✓	✓
Observations	327	324	324	319	321	324

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Robust standard errors are presented in parentheses. Adjusted p-values are not reported for these estimations as the sample size is quite small compared to the descriptive data. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds.. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The results in table 3.3o are entirely consistent with those in table 3.4 of the main text but are only weakly significant when accounting for multiple hypothesis testing (the two main variables of interest reach standard levels of significance when not using adjusted p-values). Those who received a negative mobility information significantly increased their support for more spending on the poor and higher taxes on the rich compared to subjects with the same mobility experience but no information provision in the control group. Neither those who received neutral nor those who received positive information adjusted their preferences in any significant way.

Table 3.3r: Mobility and societal perceptions - High Information Effects

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
Upward mobility	0.020	0.031	0.240	0.587
	(0.208)	(0.217)	(0.228)	(0.491)
No mobility	0.026	-0.021	-0.103	0.076
	(0.205)	(0.217)	(0.180)	(0.497)
Downward mobility	-0.302*	-0.174	-0.079	0.500
	(0.171)	(0.200)	(0.187)	(0.493)
Controls	✓	✓	✓	✓
Observations	323	320	321	302

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The results in table 3.3p are also entirely consistent with those in table 3.5 of the main text. Those who received a negative mobility information significantly decreased their perception of overall mobility in society and their perception of the mobility of the lowest quintile.

Overall, these information treatment effects are encouragingly consistent with the main results using the shock measures. Unsurprisingly, when taking pre-treatment beliefs into account, as I do in the main models in tables 3.4 and 3.5, the effects are somewhat stronger and more robust (at least for the models looking at preferences rather than beliefs).

When looking at only those subjects who were told that they experienced very high or very low mobility during treatment, as I do in tables 3.3q and 3.3r, almost none of the results survive. This is most likely due to the much larger sample size as the signs of the coefficients remain largely the same.

VI.3.5.4. Main treatment effects using a continuous mobility measure

Rather than splitting up subjects in those with negative, neutral or positive experiences, tables 3.3s and 3.3s report treatment effects using a continuous shock measure.

Table 3.3s: Support for Redistribution - Continuous Shock

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
Mobility	-0.085	0.010	0.001	0.013	0.071	-0.010
	(0.115)	(0.059)	(0.046)	(0.041)	(0.060)	(0.017)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Controls	✓	✓	✓	✓	✓	✓
Observations	596	592	593	582	590	593

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Robust standard errors are presented in parentheses. Adjusted p-values are not reported for these estimations as the sample size is quite small compared to the descriptive data. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds.. *** p<0.01, ** p<0.05 , * p<0.1.

Table 3.3t: Mobility and societal perceptions - Continuous Shock

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
Mobility	0.086	0.124**	0.025	0.028
	(0.053)	(0.057)	(0.056)	(0.136)
Controls	✓	✓	✓	✓
Observations	590	588	586	558

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** p<0.01, ** p<0.05 , * p<0.1.

The coefficients in these two tables indicate the effect of a one-point increase in the experienced mobility shock in the treatment compared to the control group. A larger value on the shock measure is hereby associated with a more positive shock. As is evident, none of the preferences are affected by the shock in a continuous way. This is unsurprising given the earlier discussion about a lacking relationship between mobility experience and preferences in the main text. Only one variable is significantly affected by the continuous measure in table 3.3s: Overall mobility. Here, a more positive mobility shock is associated with an increase in perceived societal mobility. Again, this is somewhat unsurprising given the existing literature discussed earlier. Overall, the results in tables 3.3s and 3.3s suggest that continuous models are not helpful in understanding the effects of mobility experience on redistributive preferences.

VI.3.6 Main analysis assuming positive mobility expectations

A possibility not previous discussed is that subjects may expect to have experienced some upward mobility and that the reference point should therefore not be those who experienced no mobility but those who experienced weakly positive mobility. This would suggest that subjects who are told that they experienced no mobility similarly adjust their preferences and beliefs as those who are told that they experienced downward mobility. Tables 3.3u and 3.3v test this possibility for both, preferences and beliefs. Evidently, as none of the coefficients reach any level of significance when accounting for multiple hypothesis testing, there is no evidence for this explanation. It does not seem to be the case that subjects expect to be better off than their parents but instead, that no mobility, is a reasonable reference point to use.

Table 3.3u: Support for Redistribution - Upward Expectations

	Support for Redistribution (binary)	Support for Redistribution (ordered)	Income Differences too large	More spending on Poor	Universal Basic Income	Higher Tax Share for Rich
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment						
High upward mobility	0.140	0.308	0.244	0.335	0.300	0.041
	(0.469)	(0.217)	(0.191)	(0.147)	(0.244)	(0.060)
	[0.490]	[0.377]	[0.377]	[0.161]	[0.377]	[0.490]
Upward mobility	0.132	0.223	0.080	0.046	0.146	0.064
	(0.321)	(0.155)	(0.116)	(0.122)	(0.162)	(0.047)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]
Downward & no mobility	0.345	0.180	0.126	0.154	0.075	0.054
	(0.215)	(0.104)	(0.085)	(0.088)	(0.110)	(0.034)
	[0.197]	[0.197]	[0.197]	[0.197]	[0.197]	[0.197]
Controls	✓	✓	✓	✓	✓	✓
Observations	596	592	593	582	590	593

Notes: Estimates come from logit (model (1)) and linear regressions. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Robust standard errors are presented in parentheses. Adjusted p-values are not reported for these estimations as the sample size is quite small compared to the descriptive data. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds.. *** p<0.01, ** p<0.05 , * p<0.1.

Table 3.3v: Mobility and societal perceptions - Upward Expectations

	Mobility of lowest quintile	Overall mobility	Differences due to effort	Personal benefit
	(1)	(2)	(3)	(4)
Treatment				
High upward mobility	0.068	0.084	-0.064	-0.196
	(0.205)	(0.215)	(0.210)	(0.541)
Upward mobility	-0.128	0.022	-0.055	0.393
	(0.161)	(0.173)	(0.163)	(0.394)
Downward & no mobility	-0.186	-0.180	0.028	0.395
	(0.104)	(0.112)	(0.111)	(0.270)
Controls	✓	✓	✓	✓
Observations	590	588	586	558

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. The effects are treatment effects relative to subjects with the same experimental mobility score in the control group. Controls include self-assessed mobility experience, household income and political party affiliation. The analysis is restricted to subjects who indicated that they believed the provided information and remained on the treatment and placebo screen for more than 8 seconds. *** p<0.01, ** p<0.05 , * p<0.1.

VI.4 Survey Instrument

Thank you for participating in this study. In the following, you will be asked a series of questions about your own social mobility experience. Please read the questions very carefully and answer honestly.

Part I: Demographics

D1: How old are you?

D2: What is your gender?

- Female
- Male
- Other
- Prefer not to say

D3: How many children do you have?

- I do not have children
- 1
- 2
- 3
- 4
- 5 or more

D4: Please indicate your marital status:

- Single
- Married
- Cohabiting with a partner
- Other

D5: What is your highest level of educational attainment?

- No formal qualification
- Primary education
- Secondary education
- Undergraduate degree or equivalent (e.g. bachelor's degree)
- Graduate degree or equivalent (e.g. master's degree)
- Doctoral Degree (e.g. PhD)

D6: What is your father's highest level of educational attainment?

- No formal qualification
- Primary education

- Secondary education
- Undergraduate degree or equivalent (e.g. bachelor's degree)
- Graduate degree or equivalent (e.g. master's degree)
- Doctoral Degree (e.g. PhD)

D7: What is your mother's highest level of educational attainment?

- No formal qualification
- Primary education
- Secondary education
- Undergraduate degree or equivalent (e.g. bachelor's degree)
- Graduate degree or equivalent (e.g. master's degree)
- Doctoral Degree (e.g. PhD)

D8: What is your total household income before tax?

- Under $10,000
- $10,000 - $20,000
- $20,001 - $30,000
- $30,001 - $40,000
- $40,001 - $50,000
- $50,001 - $60,000
- $60,001 - $80,000
- $80,001 - $100,000
- $100,001 - $150,000

- $150,001 - $200,000
- Above $200,000
- Don't know
- Prefer not to answer

D9: To the best of your knowledge, what was your family's household income when growing up (not accounting for inflation)?

- Under $10,000
- $10,000 - $20,000
- $20,001 - $30,000
- $30,001 - $40,000
- $40,001 - $50,000
- $50,001 - $60,000
- $60,001 - $80,000
- $80,001 - $100,000
- $100,001 - $150,000
- $150,001 - $200,000
- Above $200,000
- Don't know
- Prefer not to answer

D10: What is your current employment status?

- Full-time employee

- Part-time employee
- Self-employed or small business owner
- Unemployed and looking for work
- Student
- Not in labour force (for example: retired, or full-time parent)

D11: To which of these groups do you consider you belong? You can choose more than one group.

- American Indian or Alaska Native
- Asian
- Black or African-American
- Native Hawaiian or other Pacific Islander
- Spanish, Hispanic or Latino
- White
- Other group
- Prefer not to answer

D12: How much of the time do you think you can trust the government to do what is right?

- Never
- Only some of the time
- Most of the time
- Always

D13: In politics people sometimes talk of left and right. Where would you place yourself on the following scale? (Scale from 0 - left to 10 - right)

D14: Please select your current job from the below dropdown menu (or your last one if you don't have one now).

D19: Which party do you feel closest to?[26]

- Democratic Party
- Republican Party
- Other
- Don't know

D20: Who did you vote for in the recent 2020 Presidential Election?

- Joe Biden
- Donald Trump
- Other candidate
- Didn't vote
- Don't remember
- Prefer not to say

D15: Please select your fathers' job when you were about 14 years old from the below dropdown menu.

D16: Please select your mothers' job when you were about 14 years old from the below dropdown menu.

D17: What do you think is the difference in length in miles between the longest river in North America, the Missouri river, and the fifth longest river in North America, the Arkansas

[26] Item D19 and item D20 were not included in the pre-analysis plan but added prior to running the main study.

river?[27]

D18: Before proceeding to the next set of questions, we want to ask for your feedback about the responses you provided so far. It is vital to our study that we only include responses from people who devoted their full attention to this study. This will not affect in any way the payment you will receive for taking this survey. In your honest opinion, should we use your responses, or should we discard your responses since you did not devote your full attention to the questions so far?

- Yes, I have devoted full attention to the questions so far and I think you should use my responses for your study.
- No, I have not devoted full attention to the questions so far and I think you should not use my responses for your study.

Part II: Pre-treatment experience of Social Mobility

Q1: Please think about your present job (or your last one if you don't have one now). If you compare this job to the job your father had when you were growing up, would you say that the status of your job is:

- Much higher than your father's
- Higher
- About equal
- Lower
- Much lower than your father's
- I never had a job
- My father did not have a job while I was growing up

[27] In the original pre-analysis plan this question asked about the Missouri and the Mississippi river. After an initial pilot I changed the Mississippi to the Arkansas river as too many people had underestimated the difference between the other two rivers and so a placebo analysis based on the original question would not have been useful.

- I don't know

Q2: Please now think again about your present job (or your last one if you don't have one now). If you compare this job to the job your mother had when you were growing up, would you say that the status of your job is:

- Much higher than your mother's
- Higher
- About equal
- Lower
- Much lower than your mother's
- I never had a job
- My mother did not have a job while I was growing up
- I don't know

Q3: When you were growing up, compared with other families back then, would you say your family income was:

- Far below average
- Below average
- Average
- Above average
- Far above average

Q4: Right now, compared with other households, would you say your household income is:

- Far below average
- Below average

- Average
- Above average
- Far above average

Part III: Treatment & Control

Treatment:

We will now tell you, based on the information you gave us earlier about your own job and the jobs your parents had when you grew up, whether you have objectively experienced upward, downward or no social mobility.

The data we use is based on the International Standard Classification of Occupations (ISCO88) and the International Standard of Occupational Status (ISEI).

Based on the information you gave us, you experienced high upward mobility/upward mobility/no mobility/ downward mobility/high downward mobility.

Control:

We will now tell you, based on the answer you gave us earlier about the two rivers in North America, the Missouri and the Arkansas river, whether you have objectively overestimated, underestimated or correctly estimated the difference in length between the two rivers.

The data we use is based on the book "Rivers of North America" by Arthur C. Benke and Colbert E. Cushing.

Based on the response you gave us, you overestimated/correctly estimated/underestimated the difference in length between the two rivers.

Part IV: Post-treatment preferences for redistribution and beliefs

Q5: Please indicate to what extent you agree or disagree with the following statements:

1. It is the responsibility of the government to reduce the differences in income between people with high incomes and those with low incomes.
2. Differences in income in your country are too large.

3. The government should spend less on benefits for the poor.
4. The government should provide basic income for all.
5. In your country, a person born into the lowest income quintile has a good chance of improving their standard of living as an adult.
6. In your country, everybody has a chance to make it and be economically successful.
7. In your country, income differences are the result of differences in effort rather than luck.

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree
- Can't choose

Q6: Please tick one box for each of these to show how important you think it is for getting ahead in life...

1. How important is coming from a wealthy family?
2. How important is having well-educated parents?
3. How important is knowing the right people?

- Essential
- Very important
- Fairly important
- Not very important

- Not important at all
- Can't choose

Q7: Do you think people with high incomes should pay a larger share of their income in taxes than those with low incomes, the same share, or a smaller share?

- Much larger share
- Larger
- The same share
- Smaller
- Much smaller share
- Can't choose

Q8: To what extent do you believe that income differences arise from luck and to what extent from differences in efforts and skills? (Scale from 0 - 'from luck' to 10 - 'from effort and skills')

Q9: To what extent do you think it is acceptable for income differences to exist if they arise from luck? (Scale from 0 - 'not acceptable at all' to 10 - 'completely acceptable')

Q10: Do you think you personally benefit from redistribution by the government? (Scale from 0 - 'Not at all to my benefit' to 10 - 'Completely to my benefit')

Part V: End

C1: Do you feel that this survey was biased?

- Yes, left-wing bias
- Yes, right-wing bias
- No, it did not feel biased

C2: Did you find the information we provided you with believable?

C3: Do you have any feedback or impressions regarding this survey?

VI.5 Pre-Analysis Plan

VI.5.1 Motivation

This study seeks to understand, with the aid of a behavioural insight, how individuals' personal experience of intergenerational social mobility shapes their distributive preferences.

It addresses, with an experiment, a puzzle in the literature on distribution preferences. There is considerable evidence in this literature that individuals' perceptions of societal social mobility affect their support for redistribution (Corneo and Grüner 2002; Davidai and Gilovich 2015; Bjørnskov et al. 2013; Shariff et al. 2016; Alesina et al. 2018). However, the evidence on how ones' own experience of mobility affects preferences is mixed (Corneo and Grüner 2002; Alesina and Angeletos 2005; Clark et al. 2010; Guillaud 2013). This is surprising, given that personal mobility experience seems to be an obvious, and previously proposed (Piketty 1995), candidate for explaining distributive preferences.

In the experiment, I consider two potential mechanisms through which the experience of social mobility could affect distributive preferences. The first is a simple application of the (selfish) rational choice model: social mobility causes real changes in benefits to an individual from redistribution. The second mechanism exploits the well-documented connection between perceptions of social mobility and distributive preferences. It argues social mobility experience affects perceptions of societal mobility and, in turn, distributive preferences.

While the first mechanism suggests a linear relationship between experience of social mobility and support for redistribution, I argue, contrary to previous studies, that the second mechanism points to an asymmetric relationship between the two variables. This is because of a behavioural insight, the self-serving bias, which states that people blame external circumstances for their failures and take excessive personal credit for successes (Gilovich et al. 2002; Campbell and Sedikides 1999). Applying this bias to the case of social mobility experience suggests that people who have experienced upward mobility may believe that they "beat the odds" and do not extrapolate from their own experience onto society at large. On the other hand, those who experienced downward mobility may "blame the system" and therefore extrapolate from their experience onto society. Such an asymmetric relationship between mobility experience and distributive preferences would explain the mixed evidence in the existing literature.

Understanding how the experience of social mobility affects support for redistribution is important for three main reasons:

First, the new mechanism considered here may explain why despite the growth of inequality and the fall in social mobility, especially in the United States (Chetty et al. 2014a), there has been no significant strengthening of redistributive preferences (Kenworthy and McCall 2008). That is because a decrease in absolute mobility would mean that there are both, less people with upward and less people with downward mobility experience, ceteris paribus, leading to less demand for redistribution overall.

Second, the proposed relationship between people's experience of social mobility and distributive preferences allows to make predictions about changes in distributive preferences across time. This is more difficult when only looking at people's perceptions of societal mobility, given that little is known about how these perceptions are formed or affected by real-world events.

Third, economic policy makers focus increasingly on equalising opportunities by increasing social mobility rather than equalising outcomes by redistributing earned income (Augoustinos and De Garis 2012; Friedman 2016). This is a good political strategy only in so far as one is a substitute for the other. This is what the literature on perceived societal social mobility suggests – as the perception of mobility in a country improves, demand for redistribution decreases. If the experience of mobility is however asymmetrically related to distributive preferences, then improvements to social mobility will not actually translate into less demand for redistribution but potentially even increase that demand.

This study will also add to the growing number of information provision experiments in social sciences and thereby help advance this new methodology (Haaland et al. 2020).

VI.5.2 Research Design

Using the ISSP Social Inequality Cumulative (ISSP 2014) matched with the ISEI Index (Ganzeboom et al. 1992; Ganzeboom and Treiman 1996), I have already been able to find evidence for an asymmetric relationship between social mobility experience and distributive preferences (as well as societal mobility perceptions) in the aggregate. The aim of this proposed study is to test the causality of the relationship and to differentiate between the two potential mechanisms outlined above.

While I cannot change the real intergenerational mobility experience of respondents, I can make use of the fact that about 50% of respondents in the ISSP survey were consistently misinformed about the direction of their own mobility experience. This fact allows me to test the relationship through an information provision experiment. In particular, I will provide respondents with information on their personal intergenerational mobility experience and test how this information, if contradictory to their previously held beliefs, changes their support for redistribution. In other words, I will provide respondents with an effective personal mobility shock.

The experiment will be coded in Qualtrics with a representative subject pool of the United States population recruited via Prolific Academic. To avoid deception and ensure that the information provided is believable, I will use respondents' real experienced social mobility to tailor the information provision conditional on the actual experience, as opposed to randomly giving people potentially false information about their own experience. The basic structure of the experiment is outlined below:

Part I: Respondents state which ISCO88 code best describes their own job title and that of their parents when they were growing up. They also state how they personally assess their social mobility experience to date relative to their parents.

Part II: Respondents are divided into control and treatment group. The treatment group is given a short paragraph describing the person's mobility experience and the data used to calculate the mobility experience.[28] The control group is given similarly framed information

[28]To provide respondents with information about their mobility experience, the ISEI value of the parent with the highest ISEI score will be subtracted from the respondent's ISEI score. ISEI scores for each ISCO88

about the differences in length of the two longest rivers in the US, the Missouri and the Mississippi river.

Part III: Post-treatment questions about distributive preferences and beliefs about social mobility.

To provide subjects with the information treatment, the experiment is coded to automatically calculate the mobility experience based on subjects' responses and to then display a text based on the calculated value. These values are grouped into 5 categories: high upward mobility/upward mobility/no mobility/ downward mobility/high downward mobility. Subjects are also told explicitly within the paragraph which datasets were used to generate this information.

I will test the following main research question and hypotheses:

Do changes in perceived mobility experience asymmetrically affect distributive preferences?

H1: Respondents who experience an effective negative personal mobility shock will express more support for redistribution post-treatment than respondents in the control group.

H2: Respondents who experience an effective positive personal mobility shock will express the same level of support for redistribution post-treatment as respondents in the control group.[29]

If both, H1 and H2, can be confirmed, this is preliminary evidence for the proposed relationship between experience of social mobility and support for redistribution mediated by perceptions of societal mobility. However, it would still be possible that differences in distributive preferences between treatment and control group could be explained by beliefs about marginal benefits from taxation. For example, due to loss aversion, subjects who experienced an effective negative mobility shock may adjust their distributive preferences to a significantly larger extent than those who experienced an effective positive mobility shock because they believe their marginal benefits from redistribution to be larger.

code are available in Ganzeboom (2010).

[29]Respondents who are already aware of the direction of their mobility experience may also display a stronger asymmetric relationship between their mobility experience and distributive preferences once their experience becomes more salient post-treatment.

To further test the underlying mechanism, I will therefore ask subjects in both, treatment and control group, about their perceptions of societal mobility in the United States and about the extent to which they believe themselves to benefit from governmental redistribution. While an asymmetric change in the treatment group, relative to the control group, in the first variable would support the self-serving bias explanation, an asymmetric change in the second variable would support the rational choice explanation for the relationship. This is to answer my secondary research question:

What explains changes in distributive preferences?

H3: Effective personal mobility shocks asymmetrically affect perceptions of societal mobility.

H4: Effective personal mobility shocks asymmetrically affect perceived marginal gains from redistribution.

In a pilot experiment ran with 100 respondents from the United States in October 2020, I found that the information treatment had a highly significant effect (p¡0.01) on respondents who experienced an effective negative mobility shock, but no effect on those who experienced an effective positive mobility shock, supporting both H1 and H2. I was not yet able to test H3 and H4 during the pilot experiment.

Based on a power analysis conducted using the data from my pilot study, I will require a total of n=3,200 respondents, with 1,600 respondents in the treatment and control group, respectively, to confidently find a significant treatment effect for H1 and to confidently be able to confirm H2.

VI.5.3 Empirical Strategy

To answer my primary research question, I will estimate the following model, whereby PfR is individual i's support for governmental redistribution measured by a series of survey questions, eSM is the effective social mobility shock individual i experiences during the experiment, measured as the difference between self-assessed social mobility and objective social mobility, D is the treatment assignment, γ is a vector of controls and ϵ is the error term:

$$PfR_i = eSM_i x D_i + \gamma_i + \epsilon_i \tag{3.7}$$

Hypothesis 1 would predict that the interaction term $eSM_i x D_i$ is positive for those who have a negative eSM value and were assigned to the treatment as opposed to the control group, while hypothesis 2 would predict this to not be the case for those with a positive value of eSM.

To answer the secondary research question, I will regress the same set of explanatory variables on both, perceived societal mobility in the United States, and perceived personal gain from governmental redistribution as measured on a 10-point scale.

$$Y_i = eSM_i x D_i + \gamma_i + \epsilon_i \tag{3.8}$$

Hypothesis 3 would predict the interaction term to be positive for those who have a negative eSM value and were assigned to the treatment as opposed to the control group when perceived societal mobility is the outcome variable. Hypothesis 4 would predict the interaction term to be positive for the same group when perceived personal gain from governmental redistribution is the outcome variable.

VI.5.4 Robustness

Believability: A key concern in information provision experiments is that subjects do not believe the information they are provided with (Haaland et al. 2020). I have therefore added questions at the end of the study asking whether subjects believe the researchers to have a political bias and whether they found the information provided believable.

Social mobility and family structures: An important issue in social mobility research is that the mothers of a lot of people currently in the workforce were not actively in the workforce themselves when the respondents were growing up (Beller 2009). This can lead to different estimates of personal mobility experience for women and men if men tend to compare themselves to their fathers and women to their mothers. The subjective mobility experience of a female respondent with a mother who was not in the workforce may then reflect their difference to their mother while the objective mobility experience measure would only take the father into account. This would however only be an issue if it affected the believability of the information treatment for women more so than for men. To be able to test whether this is indeed the case, I measure self-assessed social mobility threefold: relative to the father, relative to the mother and to the parental household overall.

Chapter 4

Prosocial Risk-Taking: Distributive preferences in the presence of externalities

I Introduction

Many decisions people make have the potential to increase overall welfare by creating positive externalities. This is especially the case for decisions that are personally risky: Starting a business, developing new technologies, or investing into new ventures, are all decisions that involve uncertainty about personal gains and losses, but have potential positive externalities for wider society. For example, an entrepreneur may create knowledge spillovers through her products and services, while taking on personal risk over her own earnings. In this paper, I will refer to these decisions as prosocial risk-taking (Do et al. 2017). Prosocial risk-taking may be especially important in addressing societal challenges, such as economic inequalities or the climate emergency (Embry et al. 2019), through social entrepreneurship or by creating climate-friendly technologies. The likelihood of a person taking such prosocial risk is however affected by whether and how societies reward this decision. A growing literature on the optimal taxation of innovation is therefore concerned with how to ideally incentivise and reward innovators (e.g. Djankov et al. 2010; Da Rin et al. 2011; Stantcheva 2021a; Akcigit and Stantcheva 2020; Akcigit et al. 2022). While this literature is concerned with the optimal supply of such policies, there is so far no research on the demand for policies aimed at rewarding prosocial risk-takers. Specifically, how potential positive externalities affect distributive preferences is an open question. In this paper, I provide a simple theoretical model and a first experimental test of how distributive preferences are affected by the potential positive externalities of prosocial risk-taking.

The potential positive externalities of prosocial risk-taking can affect distributive preferences in a couple of ways which have so far not been taken into account in the literature and the experiment is designed with this in mind: The first and obvious one is that distributive preferences over income earned from risky choices could be sensitive to those potential positive externalities. While distributive preferences over income earned from risky decisions have been examined by Cappelen et al. (2013a) and others, it is plausible that these preferences would be affected if risk-taking also creates a potential benefit for others. This possibility has, however, not been examined yet. The second and less obvious consequence of the presence of positive externalities is that they create windfall gains for the individuals who receive them. Those individuals enjoy benefits which have nothing to do with their own choice of actions. Much less is known about distributive preferences with respect to windfall gains than with respect to outcomes for individuals that can be directly traced to decisions made by those individuals. For this reason, I focus on the allocation of windfall gains in my experimental design.

A large existing literature in economics is concerned with understanding the determinants of individuals' distributive preferences (Fehr and Schmidt 1999; Corneo and Grüner 2002; Falk et al. 2003; Alesina and Angeletos 2005; Klor and Shayo 2010; Cappelen et al. 2010; Alesina and Giuliano 2011; Luttmer and Singhal 2011; Cappelen et al. 2013b; Durante et al. 2014; Kuziemko et al. 2015; Alesina et al. 2018a,b; Stantcheva 2021b). The main finding, especially of the experimental strand of this literature, is that income earned through effort is redistributed less than income earned through luck (Cherry et al. 2002; Cherry and Shogren 2008; Krawczyk 2010; Cappelen et al. 2013a; Lefgren et al. 2016; Gee et al. 2017; Rey-Biel et al. 2018; Almås et al. 2020). Here, income is usually distributed by a third-party spectator with no stake in the outcomes, to isolate people's prosocial preferences. Some papers have also looked at how income from risky choices is redistributed, finding that most redistribute little even if outcomes were unlucky (Cappelen et al. 2013a; Mollerstrom et al. 2015; de Oliveira et al. 2017). This suggests that people tend to be held accountable for the choices they make when it comes to distributive preferences. While these studies have developed a good understanding of the effects of different earnings mechanisms on distributive preferences, how these preferences are affected when choices can also create externalities remains an open question. Most closely related, a small literature has looked at how a surplus is divided if subjects contributed unequal amounts to this surplus (Ruffle 1998; Rodriguez-Lara and Moreno-Garrido 2012). Here, subjects who contribute more to the surplus also tend to receive a larger payoff.

Almost all of the existing experimental studies on distributive preferences focus on how income earned through these different mechanisms is redistributed, mostly by third-party spectators, but how genuine windfalls are distributed is so far not well understood. Somewhat related, Chowdhury and Jeon (2014) study the effects of increased common show-up payments on dictator game behaviour and find that dictators distribute more if the show-up payment is larger. Importantly however, an increase in the show-up payment is not directly comparable to a windfall gain as it is still earned through participation in the experiment. Heap et al. (2021) use actual windfalls to study group identification in dictator games. How windfall gains are distributed by third-party spectators is however also an open question.

To provide a first experimental test of how potential externalities affect distributive preferences, I run an experiment consisting of decision makers and spectators. Decision makers have the option to choose a lottery or a safe payoff. In a control group, both choices only affect the own income of the decision maker. In the treatment group, decision makers face an identical choice between a lottery and a safe payoff; however, the lottery might also produce positive externalities for other, anonymous participants. Whether these potential positive externalities realise is unknown ex ante and the likelihood of them realising is independent of the outcome of the lottery.

After decision makers have made their choices, spectators have to allocate a windfall bonus between randomly selected pairs of decision makers. Pairs are designed to always be composed of one decision maker who chose the safe option and one who chose the lottery. In a within-subject design, spectators make 20 such allocation decisions across three different conditions which are randomised in order. One of their allocation decisions is ultimately selected at random to determine the payment of a decision-maker pair. In the control condition, pairs are selected from the control group of decision makers and so no potential externalities are present. Spectators therefore make their allocation decisions knowing only whether the decision makers in each pair selected the safe or risky option and their respective income. In the ex ante treatment, decision makers are selected from the treatment group. As in the control condition, spectators know which choices decision makers made and their respective income, but are not informed about whether the potential externalities actually generated. In the ex post treatment, spectators are fully informed about decision makers choices, income, and whether the externality realised.

The results of the experiment mostly support the expectations of my theoretical model and suggest that spectators reward decision makers for prosocial risk-taking. On average, the share of the windfall bonus allocated to the risk-taker increases from 49% to 53% if risk-taking can result in potential externalities. This result is independent of the risk-taker's own income and, on average, identical for the ex post and ex ante treatment. As predicted by the theoretical model, this effect also increases as the opportunity cost of choosing the risky option increases. Specifically, as the value of the safe option increases by $1, the reward allocated to the risk-taker increases by about 28-30ct in both treatments.

While these main treatment effects suggest that individuals have a significant preference for rewarding prosocial risk-taking, the results also indicate outcome bias: Spectators primarily reward risk-takers if the externality actually realised ex post. In fact, the treatment effect disappears when focusing only on the scenarios in the ex post treatment condition where the externality did not realise. That is, despite risk-takers having no agency over the likelihood of the externality realising. Additionally, and contrary to the theoretical prediction, spectators do not compensate risk-takers more in either treatment if the personal lottery was unsuccessful. Taken together, these two additional findings indicate that while spectators reward prosocial risk-taking if outcomes are unknown, they only reward successful risk-takers ex post and do not compensate prosocial risk-taking if unsuccessful. These outcome-dependent findings are in line with a growing literature on attribution bias in distributive preferences (Gurdal et al. 2013; Brownback and Kuhn 2019; Cappelen et al. 2016; Andre 2021; König-Kersting et al. 2021).

The implications of these main findings are primarily relevant for the stability of policies aimed at rewarding prosocial risk-takers, such as entrepreneurs and innovators. While people have a preference to reward prosocial risk-takers, they only wish to do so if the prosocial risk-taker was successful in generating positive externalities, even if this outcome is the result of pure luck. In practice, this finding suggests that policies aimed at rewarding prosocial risk-takers may receive support ex ante, but will likely lose this support ex post, if externalities did not realise. For example, monetary rewards for potential innovators may be in line with people's preferences to reward prosocial risk-taking ex ante. However, the support for such a policy is likely unstable and highly dependent on the success of those innovators ex post. This would be the case, even if the successful or unsuccessful outcome could not have been predicted ex ante. An important follow-up question for future research is, therefore, how this instability of policy support, if such policies are implemented, might affect the likelihood of subsequent prosocial risk-taking by decision makers.

As further analysis of spectator beliefs and choices indicates, spectators appear to trade-off their personal fairness criteria over earned income with a desire to reward prosocial risk-taking. In practice, this means that while a spectator might be a choice egalitarian in the control condition, this preference is traded-off against a desire to increase the reward for the risk-taking subject when risk-taking can generate externalities. Theoretically, this finding suggests that the desire to reward externality-generating choices can be added as an additional term in existing fairness models (Cappelen et al. 2007; Almås et al. 2020). The additional analysis, however, also reveals substantial heterogeneity in distributive decision making over prosocial risk-taking and identifies five distinct reward types among spectators.

Finally, the results also indicate that decision makers take substantially more risk when risk-taking can create potential positive externalities for others. This is the case even though reputational concerns most likely do not affect choices, as all participants are anonymous and unlikely to ever interact in person. In a belief elicitation, risk takers indicate, however, that they expect to receive a larger share of the reward in the treatment compared to the control condition. This indicates that a potential motive for prosocial risk-taking might be the expected subsequently larger share of the windfall bonus. What motivates prosocial risk-taking in the first place can, however, not be identified conclusively within the given experimental design of this paper. Spectators also underestimate how strongly decision-makers react to the possibility to generate positive externalities for others. They believe the average first-stage treatment effect to be only half its actual size. They also underestimate the importance of opportunity costs for prosocial risk-takers. Another potentially interesting question for future research would therefore be whether spectator preferences change if these incorrect beliefs about prosocial risk-taking are corrected.

The outline of this paper is as follows. Section II develops a simple and generalisable theoretical framework that incorporates externalities into distributive decision making and outlines the main hypotheses following from this. Section III describes the experimental design and section IV reports the results and further analysis of spectator beliefs. Section V concludes.

II Theoretical Framework

To provide a generalisable framework to analyse distributive choices in the presence of externalities, I propose a theoretical model that takes the size of any potential externalities and the expected costs and benefits of exposing oneself to risk into account. My basic model builds on the decision model proposed in Cappelen et al. (2007, 2013a, 2016). This model assumes that in a situation in which an impartial spectator decides on the allocation of money between two individuals, each spectator holds a specific fairness criterion which specifies a fair allocation of money to each of the two individuals. While a stakeholder would trade off their fairness criterion with their own self-interest, a spectator is exclusively motivated by their fairness criterion as self-interest cannot affect the decision of the spectator (Cappelen et al. 2013a).

There are two possible ways in which externalities can enter this model: First, as an additional fairness criterion that affects choices only in the presence of externalities. Second, as an additional term in the utility function that spectators weigh against their fairness criterion over decision makers' own income. I choose the second option in this paper, for the following reason: In most of the scenarios in which externalities are present, a desire to reward risk takers and common fairness criteria contradict each other. For example, a risk taker's own lottery might be successful and the risk taker becomes the person with the higher income in the pair. An inequality averse spectator who also wants to reward risk taking would then have to weigh her desire to reward the risk taker against her desire to decrease inequality within the pair. Depending on where she positions herself on this trade-off, she may then either allocate the bonus to equalise incomes, allocate it to the risk taker as a reward, or decide on a combination of these two options. I therefore assume that spectators' fairness criteria are independent of their desire to reward risk takers.

For a given subject, a spectator may then choose an allocation of payoff y based on the below equation. I assume for now that there are only two subjects i and j and that both subjects face the same choice set with options $a = 1$ and $a = 0$:

$$V(y_i) = -\beta \frac{(y_i - FE_i)^2}{2X} + \delta(e_i - d_i)y_i \tag{4.1}$$

Here, as in Cappelen et al. (2007, 2013a, 2016), β is the weight attached to fairness over earned income, FE the share of the bonus the subject should receive based on the spectator's fairness criterion, and X the total available bonus. This first term of the equation is almost identical to the model proposed in Cappelen et al. (2013a). The second term of the equation is where externalities enter the model. It captures the weight δ attached to the unallocated externalities e_i generated by choosing option $a = 1$ over $a = 0$. These externalities are assumed to be "unallocated" as the spectator has no information on the recipient of the externalities when making their choice. Spectators are also assumed to account for expected benefits in the payoffs subjects receive when choosing option $a = 1$ over $a = 0$ by reducing the reward and, vice versa, to increase the reward as the expected cost of choosing the option that creates the externality rises. This is captured by the term d_i.

Given the well-documented presence of outcome bias in distributive decision making (e.g. Brownback and Kuhn 2019; Cappelen et al. 2016; Andre 2021; König-Kersting et al. 2021), the value of e_i and d_i is likely going to vary depending on whether outcomes have already been realised or not. Spectators may also differ in their tendency to display outcome bias. Specifically, e_i is then given as:

$$e_i = \begin{cases} E[\pi(a_i = 1) - \pi(a_i = 0)] & \text{if } \pi \text{ is unknown} \\ \rho[\pi(a_i = 1) - \pi(a_i = 0)] + (1 - \rho)E[\pi(a_i = 1) - \pi(a_i = 0)] & \text{if } \pi \text{ is known} \end{cases} \tag{4.2}$$

Here, π is the size of the total unallocated externalities. If π is unknown, e_i is equal to the difference in the expected size of the total unallocated externality π when subject i chooses option $a_i = 1$ as opposed to $a_i = 0$. If, however, outcomes are already realised and spectators know the value of π, then they may display some degree of outcome bias. This is captured by a a value of $\rho > 0$. A spectator who displays complete outcome bias, i.e. does not take ex ante options into account at all, has a value of $\rho = 1$. Equally, d_i is given as:

$$d_i = \begin{cases} E[y_i(a_i = 1) - y_i(a_i = 0)] & \text{if } y_i \text{ is unknown} \\ \rho[y_i(a_i = 1) - y_i(a_i = 0)] + (1 - \rho)E[y_i(a_i = 1) - y_i(a_i = 0)] & \text{if } y_i \text{ is known} \end{cases} \tag{4.3}$$

Similarly, if the actual payoff of the risk taking subject y_r is yet unknown, d_i is equal to the difference in expected payoffs y_i for subject i when choosing option $a = 1$ as opposed to $a = 0$. If, however, spectators know the outcome of the personal lottery of the risk taking subject, then they may display again some degree of outcome bias as captured by a value of $\rho > 0$. In practice, this suggests that spectators assess the cost of the risky decision as lower (higher), if the personal lottery was successful (unsuccessful), and reduce (increase) their reward accordingly.

If spectators are motivated by a particular fairness view over the own income of the decision makers, they may follow a choice egalitarian or ex ante fairness criterion (Cappelen et al. 2013a) and decide on an even split of the bonus. Alternatively, spectators may follow an inequality averse or ex post criterion (Fehr and Schmidt 1999) and decide to equalise outcomes.

To ensure that spectators can follow either of these allocation rules while also rewarding risk-takers, the size of the windfall bonus exceeds the maximum possible level of inequality within the pair, which equals \$3.5, as well as the maximum value for $e_i - d_i$, which is \$2.5.[1] Given this, spectators have a large enough choice set so that I can observe interior solutions of the proposed decision model for all subjects. The optimal choice of y for a spectator is then given by:

$$y^* = FE + \frac{\delta}{\beta}(e_i - d_i) \tag{4.4}$$

In practice, this predicts that a spectator who is motivated both, by a particular fairness criterion and by potential externalities, would allocate to subject i the amount the given fairness criterion suggests, as well as an externality reward depending on the relative importance of δ over β. I can thus estimate the reward parameter δ by comparing spectator allocations between treatment and control for given values of e_i and d_i.

II.1 Hypotheses

The theoretical model outlined in section II allows me to make several predictions about spectator decisions. First, if $e_i - d_i$ is positive, which is the case for all decisions in which the safe option A is equal to \$1.50 or larger, then the model predicts that a spectator who has a positive reward parameter δ will increase y_r, the share allocated to the subject choosing the lottery, in the presence of potential positive externalities in the ex ante treatment condition. Outcome bias cannot influence spectator choices in this treatment as the realised values of $\pi(a_i = 1) - \pi(a_i = 0)$ and $y_r(a_i = 1) - y_r(a_i = 0)$ are unknown and therefore only the expected values can influence e_i and d_i. These theoretical predictions lead to H1a:

Hypothesis 1a: y_r, the share of the bonus allocated to the subject choosing the lottery, is higher in the ex ante treatment than in the control condition.[2]

[1] The externality has an expected value of \$1 as the \$2 only realise in 50% of cases irrespective of the outcome of the lottery. If a spectator has to redistribute between two subjects, whereby one chose the lottery and one took \$3.5, the maximum safe option available, then this implies $d = -\$1.5$ and so $e_i - d_i = \$1 + \$1.5$.

[2] That is, if the difference between the expected size of the externality and the expected personal benefit from choosing the risky option is positive or zero. Therefore, H1 will primarily be tested for decisions where option A > \$1.

The model further predicts that for larger values of the safe option A, the value of $e_i - d_i$ increases. This means that spectators who have a positive reward parameter δ will reward risk takers more in the presence of potential positive externalities if the expected value of the lottery chosen decreases relative to the available safe option.[3] H1b follows:

Hypothesis 1b: The treatment effect in H1a increases as the value of the safe option A increases.

As outlined in section II, an extensive literature on financial decision making as well as distributive preferences for choices that are affected by luck, points to the importance of outcome bias in spectator (or principal) decision making (e.g. Brownback and Kuhn 2019; Cappelen et al. 2016; Andre 2021; König-Kersting et al. 2021). Outcome bias is captured in the theoretical model by a value of $\rho > 0$. Assuming that spectators display some degree of outcome bias in their decision making, the model predicts a larger treatment effect on y_r when the externality realised than when it did not. That is, because spectators may not take into account the ex-ante choice set available to decision makers, but focus mostly on the ex-post monetary outcomes. H2a follows:

Hypothesis 2a: The ex post treatment effects in H1a and H1b are larger when the externality realised than when it did not.

Given that the probabilities of the externality realising and the lottery being successful for the risk-taking subject are independent in the proposed experimental design, I am also able to test whether the outcome of the personal lottery, irrespective of whether the externality realised, affects y_r, the share of the bonus allocated to the subject choosing the lottery. The theoretical model here predicts that if $\rho > 0$, the positive treatment effect on y_r, predicted in H1a, will be larger if the personal lottery was unsuccessful compared to when it was successful. That is, because for a positive value of ρ, the value of $e_i - d_i$ increases if the lottery was unsuccessful and decreases if it was successful. H2b follows:

[3] This effect could be nonlinear as it may be particularly strong in those cases where the expected monetary value (EMV) of the lottery is smaller than the EMV of the safe option (when the safe option is > \$2). That is, because risk takers incur an actual cost to their own expected payoff in those instances, which spectators may wish to compensate. For simplicity, such non-linearity is not included in the theoretical model.

Hypothesis 2b: The ex post treatment effects in H1a and H1b are larger when the personal lottery was unsuccessful than when it was successful.

While not central to the main research question of this study, the first stage of the experimental design also allows me to test whether the choices made by decision makers are affected by the potential positive externalities of the risky option. This is, in fact, a topic which has received very little attention in the experimental literature so far. To the best of my knowledge, only one paper has tested the effects of externalities on risky choices (de Oliveira 2021). She finds that while subjects are less willing to make risky choices in the presence of negative externalities, they are not more willing to expose themselves to risk to create positive externalities for others. Given that the externalities in her experimental design are certain, this would suggest that subjects who are faced with only potential positive externalities, as is the case in my design, may be even less likely to respond to these externalities with more risk seeking behaviour. Importantly, the decision makers in this study are, however, aware of the third stage of the experiment, and know that a spectator will judge their choices before awarding a windfall bonus. Even though they might therefore not be motivated to choose the risky option due to altruism (de Oliveira 2021), they may nonetheless expect spectators to judge the risk taking behaviour favourably when positive externalities could result from their choice. H3 follows:

Hypothesis 3: Decision makers are more risk seeking in the treatment than in the control group.

H3 also provides a further justification for providing spectators with hypothetical as well as real decision scenarios, given that the distribution of choices is likely to be different in the treatment as opposed to the control group.

III Experimental Design

My basic experimental design follows in particular Cappelen et al. (2013a) by asking an impartial spectator to decide on a fair allocation of a monetary bonus between two decision makers. The impartial spectator frame hereby allows me to isolate distributive preferences by eliminating the possibility of any selfish considerations affecting the decision making. Prior to spectators making their distributive choice, decision makers have to choose between a lottery and a safe income. In the treatment condition, choosing the lottery does not just yield a potentially high reward for the decision maker if successful, but might also result in positive externalities for the participant pool. The probability of the externality realising and the personal lottery being successful are, however, independent of each other for two reasons: First, this aspect of the design allows for a clean estimation of the parameters of my decision model outlined in II. Second, risky choices may also have positive externalities even if the outcome of the choice itself is unsuccessful. In the case of entrepreneurship, an example of this are knowledge spillovers (Acs et al. 2009).

Importantly, this positive externality will not benefit the two decision makers in the pair but will be used to reward randomly chosen participants from other, unrelated studies. This is a crucial aspect of the design as it allows me to hold both, expected payoffs and the degree of inequality within pairs, constant. It also ensures that there are no strategic reasons for decision makers to expose themselves to risk to generate the externality beyond a potential desire to be received as altruistic by the spectators. In the experimental instructions, I also do not refer to "externalities" at any point to avoid any positive or negative connotation subjects may have with the word.

Finally, the expected monetary value (EMV) of the risky option is, in some decisions, lower than the EMV of the safe option. While these risky options are strictly dominated for a risk averse or risk neutral subject when no potential externalities are present, this aspect of the design allows me to test whether decision makers become more risk seeking in the presence of externalities; and, how spectators value such choices.

I depart from most existing studies by asking spectators not to redistribute the earned income of the pair but to allocate an additional monetary bonus between the two decision makers. This addresses a concern arising from the literature on distributive preferences over income earned from different choices: Given that a substantial proportion of spectators can be classified as acting according to a choice egalitarian or ex ante fairness criterion (e.g. Cappelen et al. 2013a), one would likely observe little baseline redistribution in a design that asks spectators to redistribute existing earnings as opposed to distributing an additional bonus. By asking spectators to allocate an additional monetary amount, I am therefore able to observe the distributive preferences of all spectators. Additionally, while most of the experimental literature on distributive preferences focuses on subjects' choices over endowments, preferences over windfall gains have been studied less. As the existence of positive externalities means that many people's income will, at least partly, consist of what from the perceptive of that individual are windfalls, I focus on spectators' distributive choices over a windfall gain.

Stage 1: Participants recruited via Prolific Academic participate in either the control or treatment condition that determines individual payoffs.

> **Control:** Subjects are asked to decide between the following two options seven times, one choice being randomly selected for payment, each time with a slightly different value for option 1:
>
> *Option 1:* \$0.5/\$1/\$1.5/\$2/\$2.5/\$3/\$3.5
>
> *Option 2:* A 50% chance to receive \$4 and a 50% chance to receive \$0.

> **Treatment:** Subjects are asked to decide between the following two options seven times, one choice being randomly selected for payment, each time with a slightly different value for option 1:
>
> *Option 1:* \$0.5/\$1/\$1.5/\$2/\$2.5/\$3/\$3.5
>
> *Option 2:* A 50% chance to receive \$4 and a 50% chance to receive \$0. Irrespective of the outcome of the lottery, there is a 50% chance that an externality of \$2 will generate which will be used to reward two randomly chosen participants from other studies.

Stage 2: Participants are paired based on the procedure outlined in IV.2 and one of the seven decisions subjects made is randomly chosen to determine payoffs. Each pair consists of a subject that chose the lottery and a subject that chose the safe option for the randomly selected decision. This is however unknown to subjects prior to making their choice, to ensure that subjects cannot make strategic decisions about the likely composition of the pair. Importantly, only subjects within the same treatment condition are matched.

Stage 3: Impartial Spectators, who have not participated in the first two stages of the experiment, are asked to allocate a windfall bonus of $4 between pairs of decision makers in one control and two treatment conditions. Each spectator makes allocation decisions in all three conditions, but the order in which spectators see these conditions is randomized:

Control condition: Spectators allocate the $4 between pairs of decision makers from the control group in stage 1. They receive full information of the choices made and the resulting earnings of each decision maker.

Ex ante treatment condition: Spectators allocate the $4 between pairs of decision makers from the treatment group in stage 1. They receive full information of the choices made and on the resulting earnings of the decision makers but no information on whether the externality realised or not.

Ex post treatment condition: Spectators allocate the $4 between pairs of decision makers from the treatment group in stage 1. They receive full information of the choices made, the resulting earnings of the decision makers, and on whether the externality realised or not.

Figure 4.1 shows example scenarios spectators might face during the experiment. Importantly, spectators only compare the decisions of two subjects for the same choice set, i.e., the choices made when faced with the same safe and risky option. They also do not have the option to communicate with the participants.

Overall, spectators make 20 such allocation decisions - 5 in the control condition, 5 in the ex ante treatment condition, and 10 in the ex post treatment condition - with 19 being hypothetical and one resulting in actual payoffs for a participant pair. Spectators however do not know which of the allocation decisions will result in actual payoffs when making their decisions.[4]

[4]This method is commonly used in experimental designs to increase the number of observed choices without affecting the behaviour of the subjects (Charness et al. 2016).

Figure 4.1: Example Spectator Scenarios in Control & Treatments

Notes: These figures illustrate example screens spectators face during the experiment. The top left panel shows a control condition scenario, the top right panel shows an ex ante treatment condition scenario, and the two bottom panels show ex post treatment scenarios with different outcomes for the lottery over the externality. For comparability, the personal lottery decision makers faced and its outcome is identical and positive in all four cases.

IV Results

The main experiment was conducted via Prolific Academic in July 2022 with a total sample size of 180 spectators and 360 decision makers. The experimental design and all tests conducted in the following analysis were pre-registered via the American Economic Association registry for Randomized Controlled Trials with RCT ID AEARCTR-0009701.[5] The average time subjects took to complete the experiment was 15.64 minutes. The average earnings were \$5.52, which corresponds to an average hourly rate of \$21.18.

IV.1 Prosocial Risk-Taking

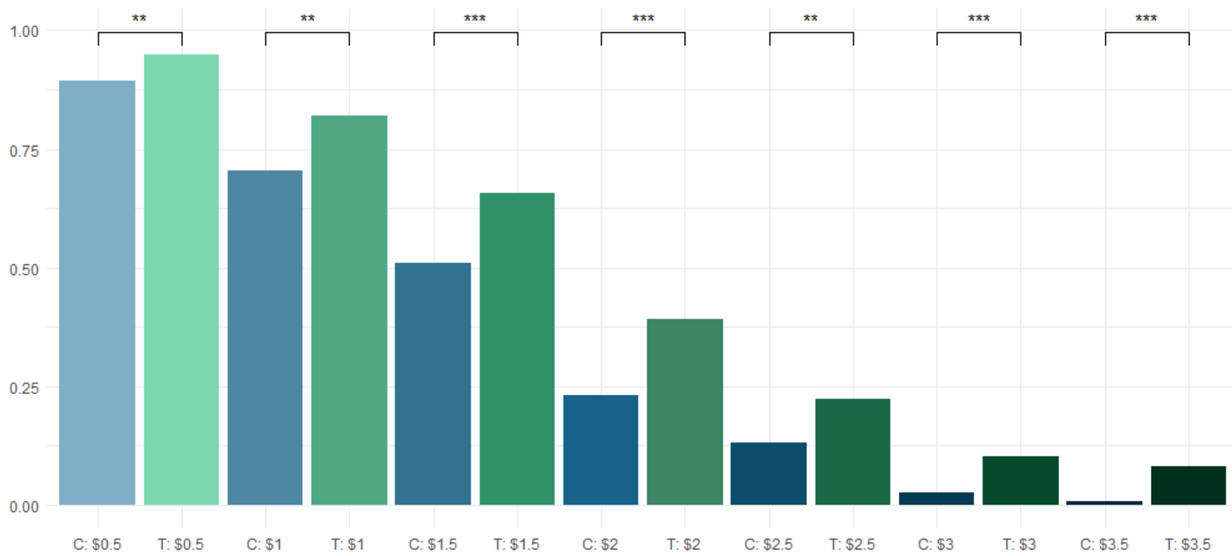
Before analyzing spectator choices, I first turn to the results of the decision maker stage. Figure 4.2 reports the proportion of decision makers choosing the risky option for a given value of the safe option. The blue bars indicate the amount of risk taking in the control conditions and the green bars the corresponding amount of risk taking in the treatment condition. In both conditions, decision makers become decreasingly likely to choose the risky option as the value of the safe option increases. This preliminary result is encouraging as it indicates decision makers understand the choices they are asked to make and rationally decrease their risk taking as the relative benefit of doing so also decreases. A share of 11.61% of decision makers in the control condition always choose the safe option, while none choose the risky option throughout. On average, decision makers in the control group are risk averse.

As figure 4.2 illustrates, the proportion of decision makers choosing the risky option in the treatment condition increases significantly for all values of the safe option compared to the control condition. This difference is also statistically significant in the aggregate ($p<0.001$) in a Pearson's chi-squared test. The results of the first stage of the experiment therefore provide strong support for hypothesis 3:

Result 1: Decision makers are significantly more likely to take risk in the treatment than in the control condition.

[5]The pre-analysis plan can be accessed at the AEA registry and in appendix VI.4.

Figure 4.2: Proportion of decision makers choosing the risky option (EMV=$2)



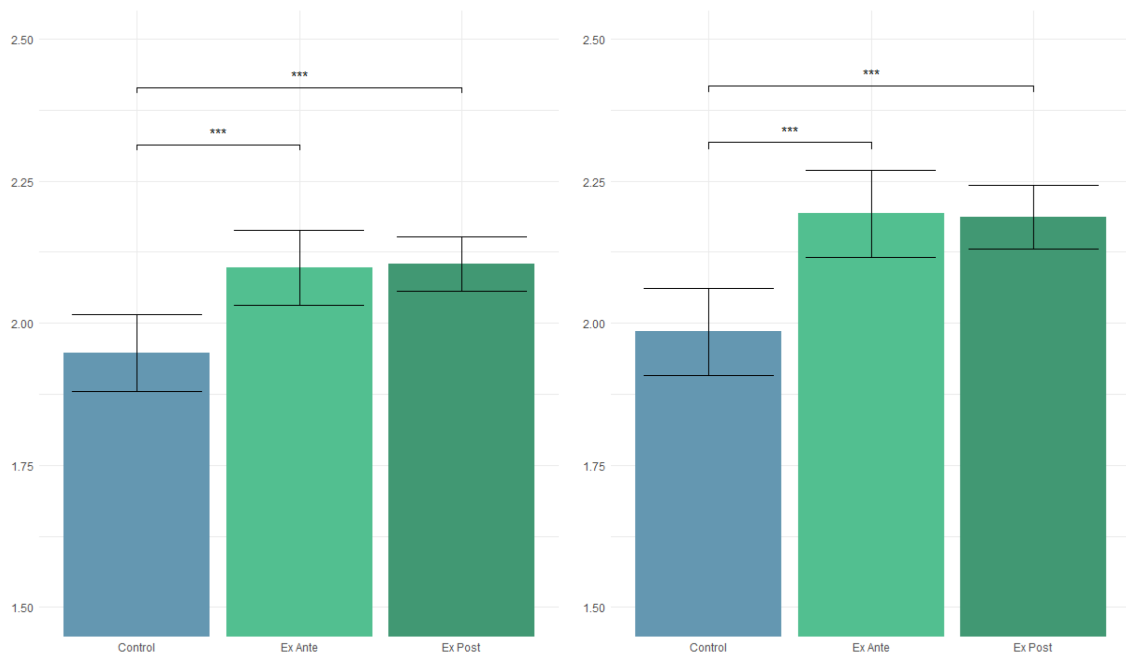
Notes: The figure reports the proportion of decision makers choosing the risky option with an expected value of $2, given the amount of the safe option and treatment condition. For example, "C: $1" reports the proportion of decision makers choosing the risky option with EMV=$2 if the safe option is equal to $1 and there is no possibility to generate externalities. Significance levels indicate the results of simple t-test. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4.1a in appendix section VI.1 reports a balance test by random treatment assignment of decision makers. As not all variables are perfectly balanced between the two samples, table 4.1b also reports treatment effects controlling for those unbalanced variables. Result 1 is robust to all of the conducted tests. In fact, the unbalanced variables actually make it less likely that a treatment effect would be observed. Therefore, the reported result is likely an underestimate of the true first stage treatment effect.

IV.2 Spectator Allocation Decisions

Figure 4.3 summarises the mean share of the bonus allocated to the risk-taking subject by treatment condition. While spectators allocate, on average, $1.95 to the risk-taker in the control condition, this value increases to $2.10 in the ex ante and ex post treatment conditions. The differences between the control and respective treatment conditions are highly statistically significant. The right part of figure 4.3 restricts the same analysis to those scenarios where the safe option was larger than $1. As outlined in section II.1, the theoret-

Figure 4.3: Share allocated to risk taker by treatment



Notes: The figure reports the share of the total bonus of $4 allocated to the subject who chose option B by treatment condition. The figure on the left includes all scenarios while the figure on the right includes only those scenarios where the safe option was larger than $1. Significance levels indicate the results of non-parametric Wilcoxon rank-sum tests. *** p<0.01, ** p<0.05 , * p<0.1.

ical framework predicts a positive treatment effect primarily if the difference between the expected size of the externality, which equals $2, and the expected personal benefit from choosing the risky option, is positive or zero. That is, because choosing the lottery in scenarios where the safe option is equal to $1 or less is a strictly dominated choice for a risk neutral decision maker even if the potential externalities are taken into account. Spectators may therefore not consider such decisions as prosocial risk-taking.

In line with this prediction, the treatment effect on the mean share of the bonus allocated to the risk-taker is also larger in the right panel of figure 4.3. Here, spectators allocate, on average, $1.98 to the risk-taker in the control condition, $2.19 to those in the ex ante treatment condition, and equally $2.19 to those in the ex post treatment condition. On average, risk takers therefore receive more than half of the bonus in the treatment condition in scenarios where the safe option is larger than $1. These treatment effects are again highly statistically significant and provide support for Hypothesis 1a:

Result 2: The share of the bonus allocated to the subject choosing the lottery is significantly higher in both treatment conditions than in the control condition.

While hypothesis 1a only focused on the comparison between the control and ex ante treatment conditions, the main result also holds for the ex post treatment condition. In fact, the difference between the ex post and ex ante treatment conditions is insignificant in both panels of figure 4.3. This main finding is also robust to various additional tests reported in section VI.2 of the appendix. Specifically, the treatment effect is stronger if spectators are more confident in their allocation decision and if they performed better when answering the understanding questions in the beginning of the experiment. They are also robust to including a large number of control variables as reported in table 4.2d of the appendix. Although it is not entirely obvious that the order in which spectators make their allocation decisions should matter, table 4.1c in appendix section VI.1 reports a balance test by the first treatment condition block spectators were randomly assigned to see. As not all variables are perfectly balanced between those three groups of spectators, table 4.1d also reports treatment effects controlling for those unbalanced variables. The results are robust to all of these additional specifications.

As outlined in the theoretical framework, hypothesis 1b predicts that the treatment effect increases as the value of the safe option increases. That is, because the opportunity cost of choosing the lottery also increases for the risk-taking subject. Table 4.1 reports treatment interaction effects with the level of the safe option. Given that the theoretical framework predicts that the treatment effect should primarily be observed as the value of the safe option increases above \$1, the results are reported for the full set of scenarios and for those with the safe option above \$1 specifically. All the coefficients in table 4.1 are highly significant and positive, suggesting that the treatment effect increases as the value of the safe option increases. The coefficients are somewhat smaller when focusing only on those scenarios where the value of the safe option is above \$1, but this difference is not statistically significant. Using the smaller estimates, these results suggest that as the value of the safe option increases by \$0.50, the share of the bonus allocated to the risk taker increases, on average, by \$0.14. In other words, 28% of the increase in the safe option is, on average, given as a bonus to the

Table 4.1: Treatment interactions with level of safe option

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
Treatment x Level	0.157***	0.154***	0.137***	0.133***
	(0.022)	(0.019)	(0.031)	(0.029)
Constant	1.788***	1.788***	1.848***	1.848***
	(0.039)	(0.039)	(0.072)	(0.072)
Observations	3,700	3,700	2,641	2,641

Notes: Estimates come from linear regressions. A one unit increase in the level of the safe option equals an increase of $0.50. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

risk taker.

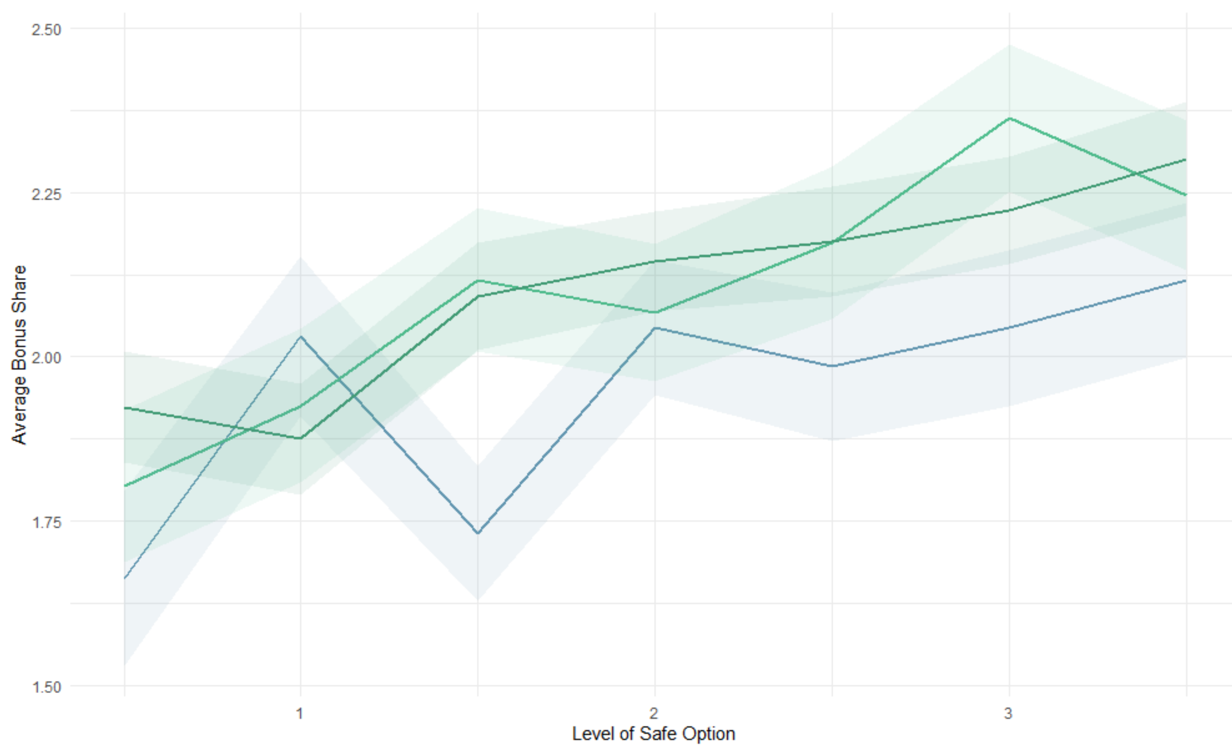
Result 3: The increase in the share of the bonus given to the risk taker in the treatment conditions increases as the value of the safe option A increases.

Figure 4.4 provides a further, and more detailed, visual test of this hypothesis. The blue line indicates the average share of the bonus given to the risk-taking subject by different values of the safe option A. The light green line does the same for the ex ante treatment and the dark green line for the ex post treatment. Shaded areas indicate 95% confidence intervals. While the two treatment condition lines overlap quite closely for all values of the safe option A, the control condition values are lower for almost all values of the safe option. Interestingly, there does not appear to be a linear increase in the difference between the control and treatment values. The values appear mostly indistinguishable until the safe option equals $1.5. At this point, the share allocated to the risk taker in the control condition drops significantly compared to the treatment conditions. The difference then becomes insignificant again at $2 but remains weakly significant and mostly stable until the safe option equals $3.5. Most of the interaction effect reported in table 4.1, therefore, appears to be driven by the increase in the treatment effect after the safe option increases to $1.5.

IV.2.1 Outcome Bias over Lottery and Externalities

As outlined in section II, my theoretical framework predicts that outcome bias will show up twofold in spectators' allocation decisions: First, as outcome bias over externalities, and

Figure 4.4: Increase in share allocated to risk taker by treatment



Notes: The figure reports the share of the total bonus of $4 allocated to the subject who chose option B by treatment condition for each value of the safe option A. As in all figures, the blue line corresponds to the control condition, the light green line corresponds to the ex ante treatment condition and the dark green line corresponds to the ex post treatment condition.

second, as outcome bias over personal lottery outcomes. Specifically, hypothesis 2a predicts that the treatment effect found previously will be larger if the positive externalities realised compared to if they did not. Hypothesis 2b predicts that the effect will also be larger if the personal lottery of the risk-taking subject was unsuccessful as opposed to successful. In other words, while spectators may reward prosocial risk-taking in itself, this reward is expected to be larger if the positive externalities actually realised and if the risk-taking subject was unlucky in their personal lottery. Table 4.2 provides a first test of these predictions focusing only on the ex post treatment, where outcomes are known. The coefficients are interaction effects of an observation being in the ex post treatment condition as opposed to the control and a particular outcome. Focusing first on whether the externality realised or not, the coefficients show significant outcome bias in the share allocated to the risk taker. If the externality realised, risk takers are rewarded with, on average, $0.23 more in the ex

Table 4.2: Ex Post treatment effect interactions with outcomes

	Amount allocated to Risk-Taker			
	Externality		Lottery	
	✓	x	✓	x
Treatment x Outcome	0.225***	0.049*	0.165***	0.156***
	(0.049)	(0.049)	(0.056)	(0.059)
Constant	1.947***	1.947***	1.656***	2.233***
	(0.034)	(0.034)	(0.045)	(0.049)
Observations	2,775	2,775	2,775	2,775

Notes: Estimates come from linear regressions. Treatment x Outcome reports coefficients for interactions between the ex post treatment condition and particular outcomes, such as the externality realising (✓) or not (x), and the personal lottery of the risk-taker being successful (✓) or not (x). Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

post treatment condition compared to the control condition. This is a 12% increase on the reward risk-takers receive in the control condition. If the externality however did not realise, there is no conventionally significant increase in the reward allocated to the risk-taker in the ex post treatment condition compared to the control condition. This means that a decision maker who took a personal risk that could have created positive externalities but by chance did not, is rewarded the same as a decision maker who took a personal risk with no possible benefit for others. This finding provides strong support for H2a:

Result 4: The treatment effect is significantly larger and only statistically significant if the externality realised ex post.

Moving to the interaction effects between the ex post treatment and the outcome of the personal lottery, there is no evidence of outcome bias. While there is a positive treatment effect for both potential outcomes of the personal lottery, this effect is not significantly larger or different in the scenarios where the personal lottery did not realise. In fact, if anything, the treatment effect is somewhat larger if the personal lottery did realise; however, the two coefficients are not significantly different from each other. This result suggests that, contrary to the predictions of the theoretical framework, a decision maker who took a personal risk that could have created positive externalities for others, does not get compensated more if the personal risk did not pay off, relative to a decision maker who took a personal risk with no possible benefit for others.[6] Hypothesis 2b can therefore not be supported:

Result 5: The treatment effect is not significantly larger if the personal lottery was unsuccessful ex post; if anything, it is somewhat smaller.

In summary, the main results suggest that the possibility to generate positive externalities significantly increases prosocial risk taking and that, on average, spectators reward prosocial risk-takers by increasing the reward they allocate to them by 8%. This reward also increases, as the personal cost of prosocial risk-taking increases. However, if outcomes are already known, spectators only reward those risk-takers who, by chance, actually generated positive benefits for others. Prosocial risk-takers who did not generate externalities by chance but faced the same choice set ex ante, get rewarded the same as risk-takers in the control condition where externalities could not be generated at all. Spectators also, on average, do not compensate prosocial risk-takers more than risk-takers in the control group if the personal risk did not pay off. In the following, I will test how to best organise the data at the individual-level and whether beliefs of spectators can explain the distributive choices made.

IV.3 Reward Types and Fairness Criteria

The within-subject design of the experiment allows me to estimate spectator-level treatment effects and reward types to gain a better understanding of how potential externalities affect distributive decision making. Figure 4.5 reports distributions of these individual treatment effects for the ex ante as well as the ex post treatment, divided into scenarios where the externality realised and those where it did not. Each value is calculated by comparing a spectator's choices between scenarios in the control and treatment conditions which are identical in the level of the safe option and the outcome of the lottery for the risk-taking subject. The difference in the share given to the risk-taking subject between the respec-

[6] It should be noted that the comparison of the *treatment* × *outcome* interactions is not identical here to that for the outcomes over the externality. That is, as there was no possibility for externalities to generate in the control condition and so the comparison in the first two columns is each to the control condition where no externalities are present. On the other hand, column three and four each report the *treatment* × *outcome* interaction effect compared to the same outcome in the control condition. This technical difference is however not relevant for the interpretation of the results.

tive treatment and control scenarios is then averaged separately for each spectator and the resulting individual-level treatment effects are reported in the figure. The red dotted line reports the average individual-level treatment effect.

The distributions of individual-level treatment effects show a similar pattern for each of the three treatment scenarios. A bit less than half of all spectators have an individual-level treatment effect of zero in each treatment scenario. The remaining majority of spectators show striking heterogeneity in how they reward prosocial risk-taking, with individual-level treatment effects ranging from -\$4 to +\$4. As reported previously in the main results, there is a positive average treatment effect in the ex ante treatment and in the ex post treatment if the externality realised. There is, on average, no treatment effect in the ex post treatment if the externality did not realise. The red dotted line indicating the average is here almost exactly at \$0.

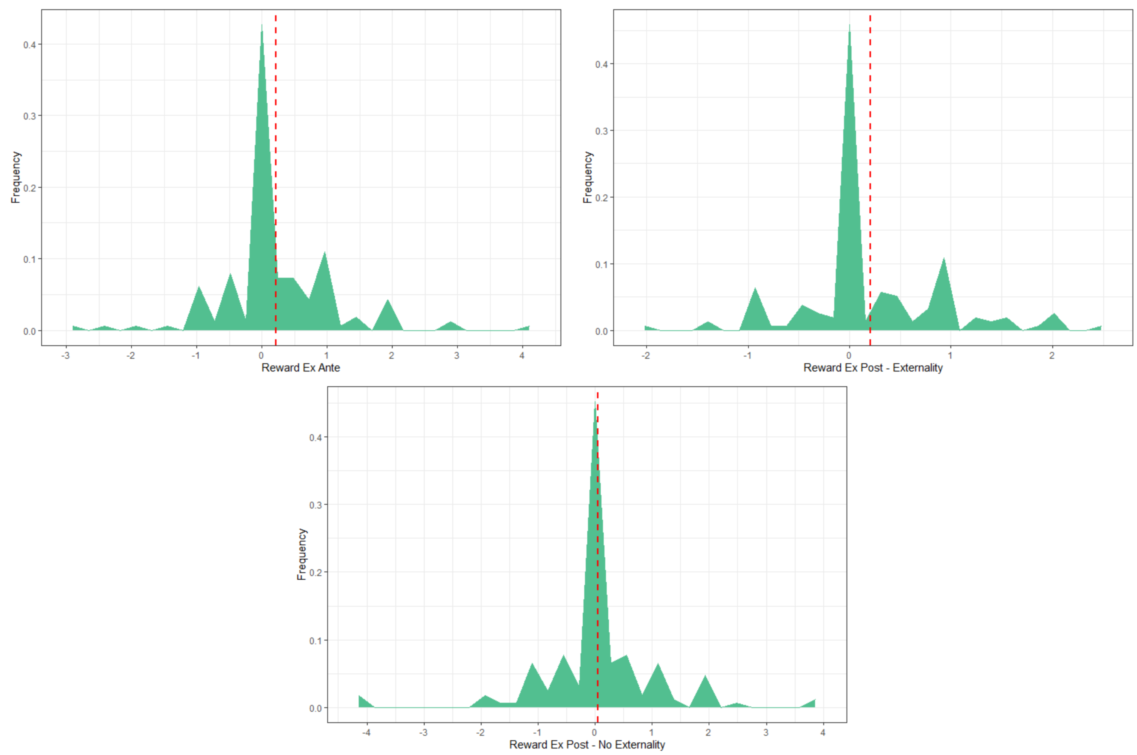
IV.3.1 Reward Types

The individual-level treatment effects for each spectator further allow me to organise spectators into reward types. Given that this paper is a first experimental test of how externalities affect distributive choices, this is also the first attempt at specifying such reward types. Each reward type proposed below can be modeled using the theoretical framework outlined previously by adjusting the reward parameter δ and the outcome bias parameter ρ.

Table 4.3 reports the share of spectators who can be classified as a particular reward type and the average bonus share given to the risk-taker by treatment condition for each type. The table also reports the share of spectators within each group of reward types who can be classified as choice egalitarian or inequality averse based on their control decisions. Finally, the table reports the average level of confidence spectators of a particular type indicated to have when making their distributive choices in the treatment scenarios.

Consistent with the heterogeneity in individual-level treatment effects reported in figure 4.5, there is also substantial heterogeneity in reward types. At least five distinct types can be identified with a substantial share of spectators within each category. About 13% of spectators do not reward prosocial risk-taking at all and have an individual-level treatment effect

Figure 4.5: Distribution of Individual-level Treatment Effects



Notes: Each figure displays a density plot of individual-level treatment effects of all spectators by treatment condition. The ex post treatment condition is divided into scenarios where the externality realised and those, where it did not. Spectator-specific treatment effects are equal to the average difference in the absolute share of the bonus given to the risk-taker comparing control and treatment conditions where the safe option and the lottery outcome are identical.

of zero across all treatment conditions. A larger share, about 18% of spectators, always reward prosocial risk-taking, irrespective of whether the decision is made ex ante or ex post and whether the externality actually realised ex post. 13.50% of spectators only reward prosocial risk-taking ex ante. This group of spectators even has a slightly negative average treatment effect in the ex post scenarios. The largest share of spectators, about 20.50%, can be classified as rewarding success as they have a positive individual treatment effect in the ex post scenarios where the externality realised. This group also has a negative average treatment effect for scenarios where the externality did not realise, suggesting that these spectators even punish prosocial risk-takers who were unsuccessful in generating the externality. The size of this negative average treatment effect is more than twice as large as the main treatment effect reported in figure 4.3. About 15% of spectators behave in exactly the

Table 4.3: Reward Types

	Spectator Share	Avg. Bonus Ex Ante	Avg. Bonus Ex Post 0	Avg. Bonus Ex Post 1	Choice Egalitarian	Inequality Averse	Confidence
Always Zero	12.97%	0.00	0.00	0.00	54.17%	16.67%	8.34
Always Positive	18.38%	1.15	1.10	1.15	2.94%	29.41%	7.42
Ex Ante Only	13.51%	0.84	-0.69	-0.22	12%	16%	7.93
Success Reward	20.54%	0.17	-0.43	0.76	13.16%	26.32%	7.36
Failed Reward	12.97%	0.01	0.73	-0.31	8.33%	45.83%	7.08
No Type	21.62%	-0.46	-0.02	-0.07	10%	40%	7.66

Notes: Reward Types are based on individual-level treatment effects for the ex ante condition, ex post condition where the externality realised, and the ex post condition where the externality did not realise. Choice Egalitarian and Inequality Averse report the share of spectators that follow either fairness criteria with 80% consistency in the control. Here, Inequality Averse refers to weak inequality aversion as specified in table 4.4. Confidence reports the average level of self-reported confidence in the distribution decisions within all treatment conditions.

opposite way and reward prosocial risk-takers ex post if the externality did not generate, but do not reward them if it did generate. Finally, a bit less than 22% of spectators cannot be categorised into any of the previously mentioned reward types. This group of spectators appears to reward prosocial risk-takers less in the ex ante than in the control condition, but appear to have an individual-level treatment effect of around 0 in the ex post conditions. While, on average, prosocial risk-taking is therefore rewarded primarily in the ex ante and ex post treatment if the externality generated, these main results conceal significant heterogeneity in reward types across spectators.

Table 4.3 also reports interesting heterogeneity in fairness criteria over earned income across reward types. Specifically, the majority of spectators who never reward prosocial risk-taking are also choice egalitarians in the control condition. On the other hand, almost half of all spectators who reward prosocial risk-takers if the externality did not realise are inequality averse in the control condition. The largest group of spectators who can be classified as a reward type, those who reward prosocial risk-taking if the externality realised, do not show a distinct pattern in their fairness criteria. If anything, these spectators seem to be less likely to be inequality averse than most of the others.

IV.3.2 Consistency of Fairness Criteria

As discussed in section II, spectators are assumed to follow a personal fairness criterion over earned income in the control condition which they trade-off against a desire to reward prosocial risk-takers in the treatment conditions. This suggests that the share of spectators who can be classified as following particular fairness criteria over earned income should be lower in the treatment as opposed to the control scenarios, as spectators are motivated by more than their personal criteria in the treatment conditions.

Table 4.4 reports the number of subjects who follow a particular fairness criterion in the control condition as well as the share of those subjects who are consistent in following this criterion across treatments. A spectator is classified as following a particular fairness criterion if at least 80% or 60% of the decisions within a treatment condition are made in line with the particular criterion. This translates to at least four or three of five decisions in the control and ex ante condition and at least eight or six of ten decisions in the ex post condition, respectively. Out of a total of 185 spectators, 28 spectators can be classified as being choice egalitarians in 80% of control condition scenarios. This number increases to 52 if the consistency requirement is reduced to 60%. Much fewer spectators can be classified as strictly inequality averse, meaning that they equalise outcomes. 14 and 18 spectators follow this criterion under the 80% and 60% consistency, respectively. As strict inequality aversion is somewhat difficult to implement correctly, given that spectators have to calculate how much of the $4 bonus each decision makers would have to receive, I also test for the consistency of weak inequality aversion. This preference is specified as simply giving more to the decision maker who initially has lower earnings and is therefore much easier to implement. Here, 41 and 87 spectators can be classified as following weak inequality aversion in their decision making under the 80% and 60% consistency requirement, respectively.
As table 4.4 illustrates, on average, the consistency across all three criteria decreases from control to the ex ante and then to the ex post treatment condition. In most cases, less than half of those who follow a particular criterion in the control condition are consistent in following this criterion across all treatment conditions.

Table 4.1e in the appendix provides an analysis of demographic variables associated with holding particular fairness criteria in the control group and table 4.2c estimates treatment effects by fairness criteria in the control condition. This second analysis suggests that only spectators who can already be classified as holding a fairness criterion in the control condition react to the treatment by increasing their reward given to the risk taker. This provides

Table 4.4: Consistency of fairness criteria

	Fairness criterion		
	Choice Egalitarian	Strict Inequality Aversion	Weak Inequality Aversion
80% consistency			
Ex Ante	64.29%	50.00%	54.55%
Ex Post	60.71%	50.00%	43.64%
Full	53.57%	28.57%	27.27%
Control	28	14	41
60% consistency			
Ex Ante	59.62%	50.00%	66.67%
Ex Post	53.85%	55.56%	72.41%
Full	44.23%	38.89%	51.72%
Control	52	18	87

Notes: The percentages indicate the share of spectators who followed a particular criterion in the control condition and also followed that criteria in one or both of the treatment conditions. Strict inequality aversion means that spectators equalised ultimate payoffs while weak inequality aversion means spectators gave more to the decision maker who had a lower initial payoff.

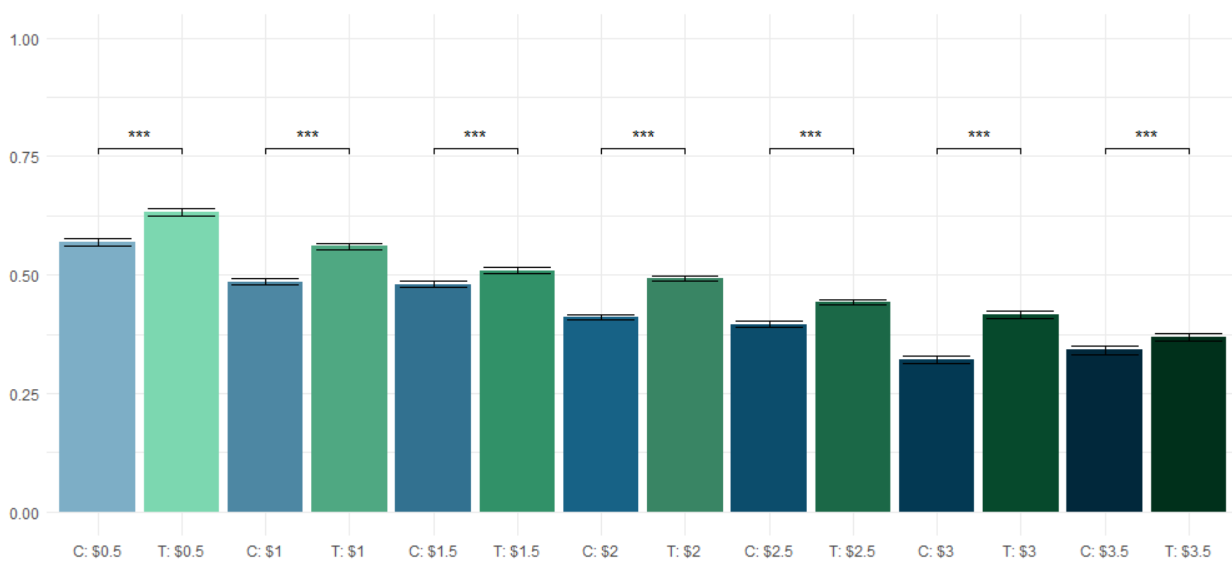
further support for a theoretical model in which spectators, on average, trade-off a fixed fairness criterion with a positive reward parameter.

IV.4 Spectator Beliefs

To better understand the main treatment effects, spectators were also asked why they believed decision makers chose the safe or risky option, and what they believe to be the effects of the generated externality for the participant pool. Additionally, they were asked to provide incentivised estimates of the number of decision makers who chose the risky option for each given value of the safe option. Appendix VI.2.3 reports the full results of these belief elicitations.

Figure 4.6 reports incentivised spectator beliefs about the share of decision makers who chose the risky option for each given level of the safe option by treatment condition. Although spectators believe decision makers were significantly more likely to choose the risky option in the treatment than control condition, these differences are, in fact, much smaller than the actual differences in decision maker choices reported in figure 4.2. Most strikingly however, when comparing these two figures, is that spectators vastly underestimate how much decision makers respond to the opportunity costs of risk-taking at the different levels of the safe option A. While the difference in the share of decision makers who chose the risky option in the scenario with a safe option of \$0.5 compared to the scenario with the safe option of \$3.5 is 88.52% in the control and 86.98% in the treatment condition, spectators estimate these same differences to only be equal to 22.83% and 26.54%, respectively. Spectators underestimate the likelihood of decision makers taking risk when the opportunity cost of doing so is low but also overestimate the likelihood of decision makers taking risk when the opportunity cost is high. Importantly, this is the case for the control as well as treatment conditions. On average, spectators estimate the average first-stage treatment effect to be an increase in risk-taking of 5.98 percentage points, while the actual increase in risk-taking is 10.26 percentage points, almost twice as large.

Figure 4.6: Spectator Beliefs about share of risk-taking decision makers by treatment condition

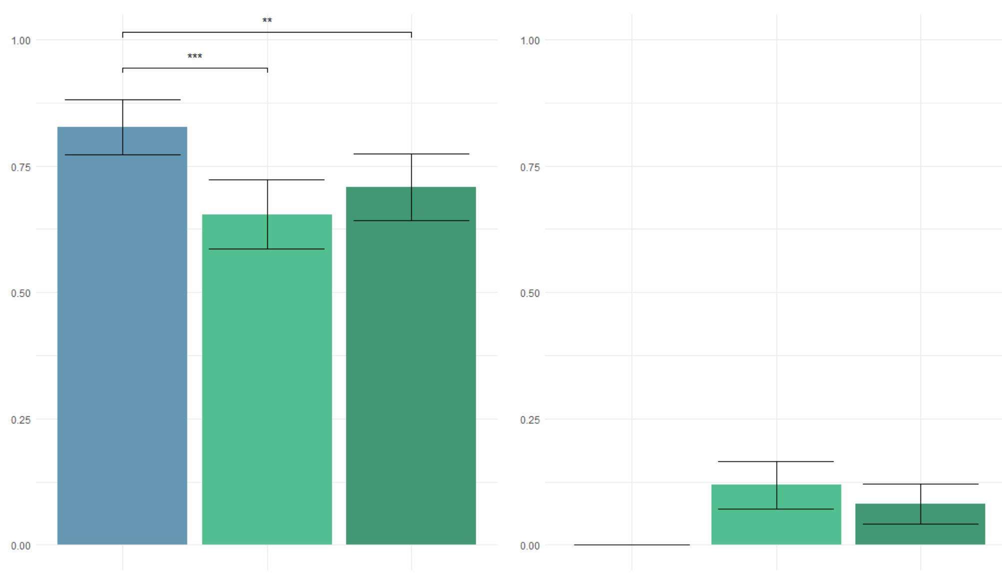


Notes: The figure reports the share of decision makers spectators believe chose the risky option B, given the amount of the safe option and treatment condition. For example, "C: \$1" reports the share of decision makers spectators belief chose the risky option if the safe option was equal to \$1 and there was no possibility to generate externalities. Significance levels indicate the results of simple t-test. *** p<0.01, ** p<0.05 , * p<0.1.

Taken together, these findings suggest that spectators (1) underestimate how strongly decision makers take opportunity costs into account and (2) underestimate how strongly decision makers react to the possibility to generate positive externalities for others. This implies that the main results likely underestimate the treatment effect one would observe if spectators were well-informed about decision makers' choices. Given that spectators do not personally benefit from allocating less of the bonus to risk-takers and receive additional rewards if the provided estimates are correct, it is also unlikely that these beliefs are motivated in some way.

A question naturally following from the previous findings is what spectators believe to be decision-makers' motives in choosing the risky option. Here, incentivising responses is, of course, not possible as there is no way to identify decision-makers true motives. Figure 4.7, therefore, provides the results of a non-incentivised belief elicitation. Spectators were asked to state why they believe some decision makers chose the risky option B. For comparability, they were asked to select one of multiple options they most agreed with. The left panel of figure 4.7 reports the share of spectators who selected the option "To maximise their own payoff" by treatment condition. The right figure does the same for the option "To generate the additional $2 for the other participants". These two statements

Figure 4.7: Spectator Beliefs about why decision makers took risk



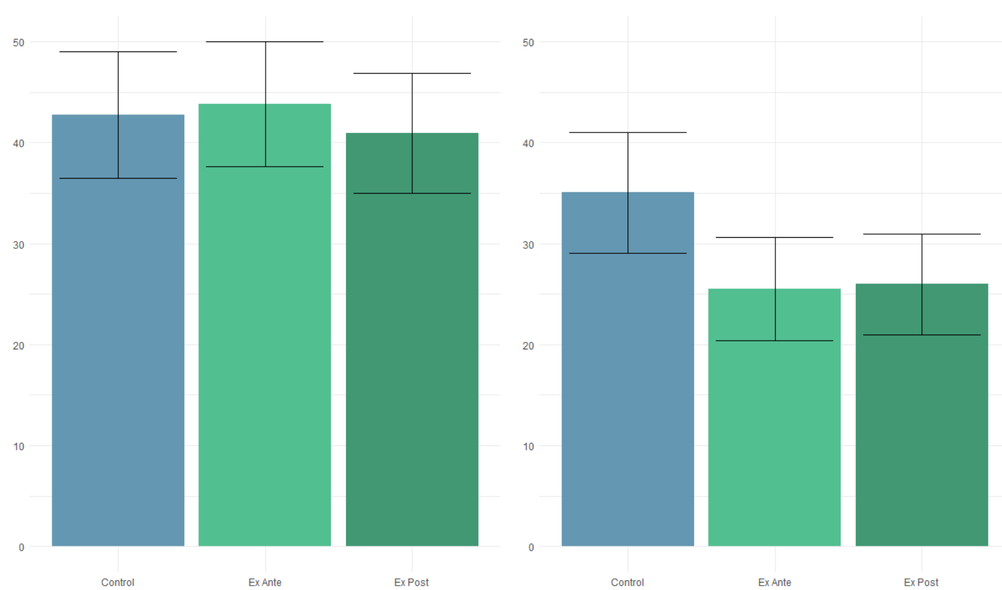
Notes: Spectators were asked why they think some participants chose the risky option B. The left figure reports the share of spectators who selected the option "To maximise their own payoff" by treatment condition. The right figure does the same for the option "To generate the additional $2 for the other participants", which was however only an option in the treatment conditions.

are the only statements on which spectator beliefs differ significantly between control and treatment conditions.[7]

As figure 4.7 illustrates, spectators are significantly less likely to believe decision makers want to maximise their own income by taking the risky option in the treatment as opposed to the control condition. This seems to be explained by some spectators believing that the main reason decision makers took the risky option was to generate the externality for others in the treatment conditions. This suggests that the treatment effect, on average, might be driven by a change in the motives spectators believe risk-takers to have when choosing the risky option. While not statistically significant, it is also interesting to note that this difference on both measures is larger between the control and ex ante conditions as opposed to the control and ex post conditions.

Finally, spectators were also asked which attributes of the scenarios they took into account when making their allocation decisions. Here, they were asked to allocate a total of 100

Figure 4.8: What mattered to your allocation decision? - 100 points



Notes: Spectators were asked to allocate 100 points between a selection of attributes depending on the significance of the attributes for the spectator's decision maker in each treatment condition. The left figure reports the average points allocated to the attribute "Inequality between participants" by treatment condition while the right figure does the same for the attribute "The choices participants made".

[7] The full results of this belief elicitation can be found in appendix VI.2.3.

points across multiple options. The full results of this elicitation can be found in appendix VI.2.3.

Figure 4.8 reports the average points allocated to two options: The left panel of the figure reports the results for the attribute "Inequality between participants" while the right panel does the same for the attribute "The choices participants made". These two options were specifically designed to test for inequality aversion and choice egalitarianism in spectator reasoning. While there appears to be a drop in the points allocated to "The choices participants made" between the control and treatment conditions, this difference is not statistically significant. The points allocated to "Inequality between participants" are strikingly consistent across treatments.

In line with the discussion in section III, these findings provide further support for the theoretical assumption that spectators trade-off their fairness criterion with their desire to reward prosocial risk-taking.

V Conclusion

Rewarding prosocial risk-taking may be a crucial strategy to address societal challenges such as the climate emergency. The aim of this paper has been to provide a theoretical framework and first experimental test of how prosocial risk-taking, and potential externalities more broadly, affect distributive decision making. The results of this paper suggest that individuals do, in fact, reward prosocial risk-taking. On average, spectators increase the share of the bonus allocated to the risk-taker from 49% to 53% if risk-taking potentially generates positive externalities for others. The treatment effect also increases as the opportunity cost of taking risk increases.

This main treatment effect however masks outcome bias and substantial heterogeneity in distributive decision-making. While individuals reward prosocial risk-taking if outcomes are unknown; they only reward successful prosocial risk-takers ex post and do not compensate unlucky risk-takers who took on prosocial risk. That is, despite the fact that individuals are more likely to believe that risk-takers have altruistic motives when there are potential positive externalities resulting from taking the risk. At least five distinct reward types of spectators can also be identified, the largest group consisting of spectators who only reward prosocial risk-taking if successful.

These results have potential implications for both, policy making and economic theory. If prosocial risk-taking is a behaviour societies wish to reward, then the results of this paper suggest that support for such policies will likely be unstable. While a large share of people may have a preference to reward prosocial risk-taking ex ante, on average, a policy in line with this preference will likely lose support ex post, if the potential positive externalities did not realise. Important questions for further research are therefore whether this instability in policy support has effects on the prevalence of prosocial risk-taking itself and whether there are ways to reduce the outcome bias in people's preferences over rewarding prosocial risk-takers.

Finally, the results of this paper demonstrate the importance of a dimension of distributive decision-making which has, so far, not received much attention: Externalities. This paper has looked at prosocial risk-taking, which focuses on *stochastic* choices and *positive* externalities. Often, externalities, however, also have negative implications for overall welfare and can affect distributive decision-making in deterministic settings. The size and type of externalities might also be important to distributive choices and their relevance might vary across different reward types. There is, therefore, a lot of scope for future research on the role of externalities in distributive decision making. The aim of this paper has been to provide a theoretical framework and an experimental design which can be adapted to study many more aspects of externalities in distributive decision-making.

Bibliography

Zoltan J Acs, Pontus Braunerhjelm, David B Audretsch, and Bo Carlsson. The knowledge spillover theory of entrepreneurship. *Small Business Economics*, 32(1):15–30, 2009.

Ufuk Akcigit and Stefanie Stantcheva. Taxation and innovation: What do we know? Technical report, National Bureau of Economic Research, 2020.

Ufuk Akcigit, Douglas Hanley, and Stefanie Stantcheva. Optimal taxation and r&d policies. *Econometrica*, 90(2):645–684, 2022.

Alberto Alesina and George-Marios Angeletos. Fairness and redistribution. *American Economic Review*, 95(4):960–980, 2005.

Alberto Alesina and Paola Giuliano. Preferences for redistribution. In *Handbook of Social Economics*, volume 1, pages 93–131. Elsevier, 2011.

Alberto Alesina, Armando Miano, and Stefanie Stantcheva. Immigration and redistribution. Technical report, National Bureau of Economic Research, 2018a.

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

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

Peter Andre. Shallow meritocracy: An experiment on fairness views. *Available at SSRN 3916303*, 2021.

Andy Brownback and Michael A Kuhn. Understanding outcome bias. *Games and Economic Behavior*, 117:342–360, 2019.

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

Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden. Responsibility for what? fairness and individual responsibility. *European Economic Review*, 54(3):429–441, 2010.

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

Alexander W Cappelen, Karl O Moene, Erik Ø Sørensen, and Bertil Tungodden. Needs versus entitlements—an international fairness experiment. *Journal of the European Economic Association*, 11(3):574–598, 2013b.

Alexander W Cappelen, Sebastian Fest, Erik Ø Sørensen, and Bertil Tungodden. Choice and personal responsibility: What is a morally relevant choice? *The Review of Economics and Statistics*, pages 1–35, 2016.

Elisa Cavatorta and David Schröder. Measuring ambiguity preferences: A new ambiguity preference survey module. *Journal of Risk and Uncertainty*, 58(1):71–100, 2019.

Gary Charness, Uri Gneezy, and Brianna Halladay. Experimental methods: Pay one or pay all. *Journal of Economic Behavior & Organization*, 131:141–150, 2016.

Todd L Cherry and Jason F Shogren. Self-interest, sympathy and the origin of endowments. *Economics Letters*, 101(1):69–72, 2008.

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

Subhasish M Chowdhury and Joo Young Jeon. Impure altruism or inequality aversion?: An experimental investigation based on income effects. *Journal of Public Economics*, 118: 143–150, 2014.

Giacomo Corneo and Hans Peter Grüner. Individual preferences for political redistribution. *Journal of Public Economics*, 83(1):83–107, 2002.

Marco Da Rin, Marina Di Giacomo, and Alessandro Sembenelli. Entrepreneurship, firm entry, and the taxation of corporate income: Evidence from europe. *Journal of Public Economics*, 95(9-10):1048–1066, 2011.

Angela de Oliveira. When risky decisions generate externalities. *Journal of Risk and Uncertainty*, 63(1):59–79, 2021.

Angela CM de Oliveira, Alexander Smith, and John Spraggon. Reward the lucky? an experimental investigation of the impact of agency and luck on bonuses. *Journal of Economic Psychology*, 62:87–97, 2017.

Simeon Djankov, Tim Ganser, Caralee McLiesh, Rita Ramalho, and Andrei Shleifer. The effect of corporate taxes on investment and entrepreneurship. *American Economic Journal: Macroeconomics*, 2(3):31–64, 2010.

Kathy T Do, João F Guassi Moreira, and Eva H Telzer. But is helping you worth the risk? defining prosocial risk taking in adolescence. *Developmental Cognitive Neuroscience*, 25: 260–271, 2017.

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

Elizabeth Embry, Jessica Jones, and Jeffrey G York. Climate change and entrepreneurship. In *Handbook of Inclusive Innovation*, pages 377–393. Edward Elgar Publishing, 2019.

Armin Falk, Ernst Fehr, and Urs Fischbacher. On the nature of fair behavior. *Economic Inquiry*, 41(1):20–26, 2003.

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

Laura K Gee, Marco Migueis, and Sahar Parsa. Redistributive choices and increasing income inequality: experimental evidence for income as a signal of deservingness. *Experimental Economics*, 20(4):894–923, 2017.

Mehmet Y Gurdal, Joshua B Miller, and Aldo Rustichini. Why blame? *Journal of Political Economy*, 121(6):1205–1247, 2013.

Shaun P Hargreaves Heap, Eugenio Levi, and Abhijit Ramalingam. Does group identification promote giving with respect to windfall gains in a dictator game? 2021.

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

Christian König-Kersting, Monique Pollmann, Jan Potters, and Stefan T Trautmann. Good decision vs. good results: Outcome bias in the evaluation of financial agents. *Theory and Decision*, 90(1):31–61, 2021.

Michał Krawczyk. A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. *Journal of Public Economics*, 94(1-2):131–141, 2010.

Ilyana Kuziemko, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva. How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, 105(4):1478–1508, 2015.

Lars J Lefgren, David P Sims, and Olga B Stoddard. Effort, luck, and voting for redistribution. *Journal of Public Economics*, 143:89–97, 2016.

Erzo FP Luttmer and Monica Singhal. Culture, context, and the taste for redistribution. *American Economic Journal: Economic Policy*, 3(1):157–79, 2011.

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

Pedro Rey-Biel, Roman Sheremeta, and Neslihan Uler. *When income depends on performance and luck: The effects of culture and information on giving*. Emerald Publishing Limited, 2018.

Ismael Rodriguez-Lara and Luis Moreno-Garrido. Self-interest and fairness: self-serving choices of justice principles. *Experimental Economics*, 15(1):158–175, 2012.

Bradley J Ruffle. More is better, but fair is fair: Tipping in dictator and ultimatum games. *Games and Economic Behavior*, 23(2):247–265, 1998.

Stefanie Stantcheva. Perceptions and preferences for redistribution. Technical report, National Bureau of Economic Research, 2021a.

Stefanie Stantcheva. Understanding tax policy: How do people reason? *The Quarterly Journal of Economics*, 136(4):2309–2369, 2021b.

Ana Venâncio, Victor Barros, and Clara C Raposo. Corporate taxes and high-quality entrepreneurship: evidence from a tax reform. *Available at SSRN 3523371*, 2020.

VI Appendix

VI.1 Descriptive Analysis of Experimental Data

VI.1.1 Balance Test of First Stage Treatment Assignment

Although decision makers are not the primary subjects of the experiment, table 4.1a reports a balance test for decision makers in the first stage of the experiment by random treatment assignment. Detailed descriptions of each variable can be found in appendix VI.3. Interestingly, quite a few variables are not properly balanced in the decision maker sample despite the random assignment. Specifically, decision makers in the treatment group are more likely to be female, less likely to have studied economics at university, somewhat less likely to be black or African-American, and display slightly less ambiguity aversion.

To test whether prosocial risk-taking in the first stage is driven by any of these underlying demographic differences between the samples, table 4.1b reports the results of simple and logit regressions with the switch point of decision makers as the outcome variables. The switchpoint is hereby defined as the level of the safe option at which decision makers moved from the risky to the safe option in the first two models and as a dummy variable dependent on whether the decision maker switched before or after the safe option is equal to $2. This binary specification of the switchpoint aims to capture risk-seeking behaviour within the personal lottery decision.

The results of these regressions show that the main first stage treatment effect is robust to adding the required control variables based on the balance test in table 4.1a. Decision makers in the treatment group are significantly more likely to switch to the safe option at a higher level of the safe option A, and are also substantially more likely to display risk neutral or risk seeking behaviour by switching at a level of the safe option above $1.50. On average, the highly significant coefficient of 0.862 in the full model suggests that decision makers in the treatment group switch almost one level later than decision makers in the control group, even when accounting for the unbalanced variables in the sample.

Table 4.1a: Balance Test by treatment assignment of decision makers

	Control	Treatment	t-statistic
Demographics			
Age	40.37	40.37	-0.009
	(13.78)	(14.09)	(2,518)
Gender	0.61	0.45	7.83***
	(0.49)	(0.50)	(2,469)
Education	4.40	4.47	-1.07
	(1.47)	(1.47)	(2,497)
Income	54,960	58,665	-1.81*
	(50,040)	(47,334)	(2,469)
Employment	0.70	0.66	2.09**
	(0.46)	(0.47)	(2,518)
Economics	0.47	0.35	5.59***
	(0.02)	(0.01)	(2,518)
Ethnicity			
American Indian or Alaska Native	0.00	0.00	-1.90*
	(0.00)	(0.00)	(2,518)
Asian	0.10	0.12	-1.76*
	(0.30)	(0.33)	(2,518)
Black or African-American	0.16	0.11	3.35***
	(0.36)	(0.31)	(2,518)
Native Hawaiian or Pacific Islander	0.00	0.00	-1.90*
	(0.00)	(0.06)	(2,518)
Spanish, Hispanic or Latino	0.07	0.05	1.90*
	(0.26)	(0.23)	(2,518)
White	0.72	0.73	-0.52
	(0.45)	(0.44)	(2,518)
Preferences			
Risk seeking	4.78	4.72	0.60
	(2.18)	(2.48)	(2,518)
Ambiguity aversion	16.63	15.66	4.10***
	(5.72)	(5.56)	(2,518)
Left-Right Placement	3.65	3.55	0.78
	(2.76)	(2.83)	(2,462)
Party affiliation			
Democrats	0.61	0.63	-0.76
	(0.49)	(0.48)	(2,518)
Republicans	0.18	0.20	-1.28
	(0.38)	(0.40)	(2,518)
Other	0.11	0.12	-0.21
	(0.32)	(0.32)	(2,518)
2020 Vote			
Joe Biden	0.57	0.54	1.13
	(0.50)	(0.50)	(2,518)
Donald Trump	0.17	0.16	0.80
	(0.38)	(0.37)	(2,518)
Other	0.02	0.03	-1.42
	(0.13)	(0.16)	(2,518)
Didn't vote	0.21	0.22	-0.31
	(0.41)	(0.41)	(2,518)

Notes: Table reports the mean values for decision makers based on treatment group assignment. Detailed descriptions of each variable can be found in appendix section VI.3. Asterisks indicate significant differences in mean values between samples from a simple test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

When looking at the treatment effects interacted with each of the unbalanced variables, some interesting patterns emerge. Although male decision makers react more strongly than female decision makers to the treatment by switching later in the individual interactions model, this is not robust to any of the other specifications. Decision makers who studied economics at university, however, are more likely to switch at a later point in the treatment group in both interaction models. They are no more likely to become risk neutral or risk seeking in the treatment group though. The opposite pattern can be observed for black and African-American decision makers. They are substantially more likely to become risk neutral or risk-seeking when risk-taking leads to potential positive externalities for others, compared to other ethnicities. Finally, ambiguity aversion does not seem to have a substantial impact on the treatment effect.

Table 4.1b: Prosocial Risk-taking of Decision Makers with Controls

	Full Results			Above $1.50		
	Indv. Interacted	Full Interacted	Controls	Indv. Interacted	Full Interacted	Controls
Treatment	-	-	0.862***	-	-	0.659***
	-	-	(0.185)	-	-	(0.249)
Male	0.516***	0.230	0.336*	0.255	0.029	0.064
	(0.195)	(0.229)	(0.187)	(0.237)	(0.268)	(0.226)
Economics	0.777***	0.670***	0.297	0.438*	0.351	0.074
	(0.221)	(0.248)	(0.191)	(0.253)	(0.288)	(0.233)
Black	0.774**	0.520*	0.131	1.201***	1.103**	0.580*
	(0.312)	(0.309)	(0.250)	(0.428)	(0.438)	(0.320)
Ambiguity	0.027***	0.008	-0.025	0.022*	0.010	-0.009
	(0.010)	(0.011)	(0.016)	(0.012)	(0.014)	(0.020)
Constant	-	3.616***	3.481***	-	-0.746***	-0.893**
	-	(0.136)	(0.332)	-	(0.174)	(0.414)
Observations	326	326	322	358	358	353

Notes: Estimates come from linear regressions. Models 1 and 4 report results from individual regressions with interactions between the specific control variable and the treatment. Models 2 and 5 report results from one model in which the treatment is interacted with each control variable separately. Models 3 and 6 report a simple model in which the variables are added as control variables. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Given that decision makers in the treatment group are more likely to be female, less likely to have studied economics, and less likely to be black or African-American, these results actually indicate that the first stage treatment effect reported in the main text is most likely an underestimate of the true treatment effect. That is, as each of these demographic groups more present in the treatment than control group also react less strongly to the treatment.

VI.1.2 Balance Test of Second Stage Treatment Assignment

As the second stage of the experiment consists of a within subject design, whereby spectators see all three treatment conditions but in randomised order, a balance test by treatment assignment cannot be conducted. However, to test whether the randomised viewing order is correlated with any demographic variables, table 4.1c reports the results of a balance test by the first treatment condition spectators were randomly assigned to. This can be conducted as spectators made allocation decisions in randomly ordered treatment blocks. In other words, spectators, for example, first made five allocation decisions in the control scenario, then five in the ex ante, and then ten in the ex post treatment condition. The table therefore reports the results of a balance test based on the first randomly shown treatment block. The significance tests are always conducted between the control and the respective treatment conditions.

Similarly to the random treatment assignment of decision makers, there are multiple demographic variables which are correlated with the randomly assigned first treatment condition spectators saw. However, only some are significantly different in both, the ex ante and ex post conditions, compared to the control. Given that the primary concern about the viewing order would be that being informed of the potential treatment externalities before completing the control condition might bias results, this aggregate comparison between seeing either of the treatment as opposed to the control condition first seems most important. Specifically, spectators who first completed either of the treatment conditions were more likely to be female, more highly educated, very slightly more likely to be asian and spanish, hispanic, or latino, somewhat more likely to place themselves further towards the right of a political left-right scale, somewhat less likely to have voted for Joe Biden, and more likely to not have voted at all in the 2020 election. Simiarly to table 4.1b, table 4.1d reports the results of simple regressions with the absolute share of the bonus allocated to the risk taking subject as the outcome variable.

The first result to note is that the main treatment effects are robust to including all the variables as controls that are unbalanced in the sample. Additionally, none of the variables show a consistent treatment effect interaction across all the reported models.

Table 4.1c: Balance Test by first treatment condition of spectators

	Control	Ex Ante	t-statistic	Ex Post	t-statistic
Demographics					
Age	39.27	40.20	-1.52	41.67	-4.20***
	(14.20)	(15.59)	(2,358)	(14.89)	(2,598)
Gender	0.46	0.32	6.91***	0.37	4.52***
	(0.50)	(0.47)	(2,318)	(0.48)	(2,598)
Education	4.06	4.62	-9.74***	4.52	-7.40***
	(1.50)	(1.23)	(2,358)	(1.65)	(2,598)
Income	48,246	49,020	-0.53	57,232	-5.94***
	(32,248)	(35,732)	(2,158)	(41,173)	(2,438)
Employment	0.67	0.65	0.62	0.67	-0.27
	(0.47)	(0.48)	(2,358)	(0.47)	(2,598)
Economics	0.38	0.51	-6.30***	0.39	-0.37
	(0.49)	(0.50)	(2,358)	(0.49)	(2,598)
Ethnicity					
American Indian or Alaska Native	-	-	-	0.03	-6.22***
	-	-	-	(0.17)	(2,598)
Asian	0.08	0.11	-2.48**	0.12	-3.41***
	(0.27)	(0.31)	(2,358)	(0.32)	(2,598)
Black or African-American	0.11	0.13	-1.21	0.10	0.55
	(0.31)	(0.33)	(2,358)	(0.31)	(2,598)
Native Hawaiian or Pacific Islander	-	-	-	-	-
	-	-	-	-	-
Spanish, Hispanic or Latino	0.08	0.20	-8.67***	0.18	-7.62***
	(0.27)	(0.40)	(2,358)	(0.38)	(2,598)
White	0.73	0.65	3.99***	0.66	4.07***
	(0.44)	(0.48)	(2,358)	(0.47)	(2,598)
Preferences					
Risk seeking	5.11	5.49	-3.89***	5.02	1.09
	(2.30)	(2.44)	(2,358)	(2.13)	(2,558)
Ambiguity aversion	15.52	15.82	-1.28	15.34	0.81
	(6.03)	(5.00)	(2,358)	(5.28)	(2,598)
Left-Right Placement	3.56	3.85	-2.78***	3.82	-2.37**
	(2.69)	(2.42)	(2,338)	(2.85)	(2,558)
Party affiliation					
Democrats	0.52	0.47	2.48**	0.61	-4.55***
	(0.50)	(0.50)	(2,358)	(0.49)	(2,598)
Republicans	0.22	0.20	1.32	0.19	1.77*
	(0.42)	(0.40)	(2,358)	(0.40)	(2,598)
Other	0.22	0.27	-2.85***	0.16	3.76***
	(0.42)	(0.45)	(2,358)	(0.37)	(2,598)
2020 Vote					
Joe Biden	0.60	0.47	6.40***	0.54	3.40***
	(0.49)	(0.50)	(2,358)	(0.50)	(2,598)
Donald Trump	0.13	0.20	-4.84***	0.15	-1.64
	(0.33)	(0.40)	(2,358)	(0.36)	(2,598)
Other	0.05	0.04	1.35	0.01	4.84***
	(0.21)	(0.19)	(2,358)	(0.12)	(2,598)
Didn't vote	0.14	0.25	-6.90***	0.27	-8.00***
	(0.35)	(0.44)	(2,358)	(0.44)	(2,598)

Notes: Table reports the mean values for spectators based on the first treatment condition block they were randomly allocated to make distribution decisions for. Asterisks indicate significant differences in mean values between samples from a simple test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Table 4.1d: Main Treatment Effects with Controls

	Indv. Interacted	Ex Ante Full Interacted	Controls	Indv. Interacted	Ex Post Full Interacted	Controls
Treatment	-	-	0.142***	-	-	0.149***
	-	-	(0.049)	-	-	(0.043)
Male	0.157**	0.103	0.066	0.066	-0.013	0.005
	(0.066)	(0.073)	(0.053)	(0.049)	(0.052)	(0.043)
Education	0.027***	0.011	-0.018	0.038***	0.048***	0.021
	(0.010)	(0.018)	(0.016)	(0.008)	(0.012)	(0.013)
Asian	0.223**	0.171	0.053	0.176**	0.183*	0.103
	(0.112)	(0.126)	(0.092)	(0.086)	(0.095)	(0.077)
Spanish, Hispanic, Latino	-0.089	-0.142*	-0.132**	0.051	0.053	-0.009
	(0.081)	(0.086)	(0.065)	(0.063)	(0.067)	(0.056)
Left-Right Placement	0.017*	0.006	-0.019	0.018**	0.001	-0.010
	(0.010)	(0.013)	(0.012)	(0.007)	(0.009)	(0.010)
Joe Biden Vote	0.122**	0.058	-0.156**	0.091**	-0.088	-0.161**
	(0.052)	(0.079)	(0.079)	(0.041)	(0.056)	(0.063)
No Vote	-0.022	-0.068	-0.136	-0.106*	-0.217***	-0.202***
	(0.075)	(0.095)	(0.083)	(0.057)	(0.068)	(0.068)
Constant	-	1.962***	2.208***	-	1.958***	2.019***
	-	(0.034)	(0.127)	-	(0.034)	(0.107)
Observations	1,820	1,820	1,800	2,730	2,730	2,700

Notes: Estimates come from linear regressions. Models 1 and 4 report results from individual regressions with interactions between the specific control variable and the treatment. Models 2 and 5 report results from one model in which the treatment is interacted with each control variable separately. Models 3 and 6 report a simple model in which the variables are added as control variables. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

The only variable showing a consistent and significant effect across all models is the ”no vote” variable in the ex post treatment condition. Those spectators who self-reported to not have voted in the 2020 election, on average, are less likely to reward the risk taker in the ex post treatment condition than in the control. They are also, however, less likely to reward the risk taker across the control and ex post conditions, as the negative and significant coefficient in the control model shows.

VI.1.3 Demographics by Fairness Criteria in Control Condition

Table 4.1e reports differences in demographics between spectators who were identified as having either a choice egalitarian or weakly inequality averse fairness criterion. For the purpose of this comparison, the 60% consistency requirement was used, meaning that spectators had to make at least three of the five control allocation decisions in line with the criterion. As fairness criteria are endogenous, the results in table 4.1e are not relevant for the robustness of any of the main treatment effects.

Spectators who have a choice egalitarian criterion in the control condition are somewhat older, more likely to be male, less educated, have a lower household income, are more likely to be employed, and are less likely to have studied economics. There are also more likely to be white and less likely to be asian and, although only weakly significanly so, less likely to be spanish, hispanic, or latino. Arguably, the sample size for American Indian or Alaska Native is too small to justify an interpretation of the significance test for this variable. Choice egalitarians are also less likely to be ambiguity averse and less likely to be Republicans than inequality averse spectators.

VI.2 Additional Results

VI.2.1 Main effects with different specifications

VI.2.1.1. Treatment effects by confidence

Table 4.2a reports interactions of treatment and levels of confidence in the allocation decisions. Spectators were asked about their level of confidence in their decision on a ten-point scale after each allocation scenario.

The results show that the treatment effect is stronger for spectators who are more confident in their allocation decisions. This result is comparable for both treatment conditions and also increases somewhat when focusing only on those scenarios where the safe option is above $1. This result is encouraging as it suggests that the treatment effect is due to a conscious choice by spectators.

VI.2.1.2. Treatment effects by understanding

Table 4.2b reports interactions of treatment assignment and the level of understanding of spectators. Prior to making their first allocation decisions, spectators were asked four questions about the scenarios they would face. After answering each question, they were provided with the correct answer to each question. The specific questions can be found in appendix VI.4.

Table 4.1e: Demographics by revealed fairness criteria of spectators

	Choice Egalitarian	Inequality Aversion	t-statistic
Demographics			
Age	40.96	39.29	-2.91***
	(15.19)	(14.35)	(2,778)
Gender	0.33	0.38	3.01***
	(0.47)	(0.49)	(2,758)
Education	4.15	4.45	4.95***
	(1.62)	(1.45)	(2,778)
Income	45,401	52,840	5.06***
	(31,206)	(39,492)	(2,618)
Employment	0.69	0.64	-2.63***
	(0.46)	(0.48)	(2,778)
Economics	0.38	0.43	2.11**
	(0.49)	(0.49)	(2,778)
Ethnicity			
American Indian or Alaska Native	0.02	0.00	-5.84***
	(0.14)	(0.00)	(2,778)
Asian	0.02	0.15	11.28***
	(0.00)	(0.01)	(2,778)
Black or African-American	0.13	0.13	-0.62
	(0.34)	(0.33)	(2,778)
Native Hawaiian or Pacific Islander	-	-	-
	-	-	-
Spanish, Hispanic or Latino	0.13	0.16	1.87*
	(0.34)	(0.37)	(2,778)
White	0.71	0.66	-3.08***
	(0.45)	(0.48)	(2,778)
Preferences			
Risk seeking	5.17	5.13	-0.53
	(2.38)	(2.28)	(2,738)
Ambiguity aversion	14.67	15.47	3.68***
	(6.03)	(5.21)	(2,778)
Left-Right Placement	3.50	3.57	0.69
	(2.40)	(2.63)	(2,718)
Party affiliation			
Democrats	0.60	0.56	-1.70*
	(0.49)	(0.50)	(2,778)
Republicans	0.13	0.18	3.39***
	(0.34)	(0.39)	(2,778)
Other	0.21	0.23	1.12
	(0.41)	(0.42)	(2,778)
2020 Vote			
Joe Biden	0.58	0.57	-0.11
	(0.49)	(0.49)	(2,778)
Donald Trump	0.12	0.14	1.71*
	(0.32)	(0.34)	(2,778)
Other	0.00	0.03	6.09***
	(0.00)	(0.18)	(2,778)
Didn't vote	0.23	0.23	-0.05
	(0.42)	(0.42)	(2,778)

Notes: Table reports the mean values for spectators based on their revealed fairness criteria in the control condition assuming 60% consistency and using the weak definition of inequality aversion. Asterisks indicate significant differences in mean values between samples from a simple test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4.2a: Treatment interactions with level of confidence

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
Treatment x Confidence	0.016***	0.015***	0.023***	0.022***
	(0.006)	(0.005)	(0.007)	(0.006)
Constant	1.963***	1.975***	2.000***	2.005***
	(0.033)	(0.033)	(0.038)	(0.037)
Observations	1,833	2,731	1,309	1,949

Notes: Estimates come from linear regressions. Confidence is self-reported and ranges from 1 to 10 with higher values indicating more confidence in the allocation decision. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

In table 4.2b, the understanding variable is coded on a five-point scale with 0 for spectators who got all questions wrong and 4 for spectators who correctly answered all questions. As the table reports, a higher value on this understanding scale increases the treatment effect for both treatments. Similarly to the confidence models in table 4.2a, this effect is even stronger when focusing only on the scenarios where the safe option is above $1. Again, these results are encouraging as they suggest that a better understanding of the decision problem led to spectators increasing their reward for prosocial risk-taking.

Table 4.2b: Treatment interactions with degree of understanding

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
Treatment x Understanding	0.061***	0.068***	0.075***	0.079***
	(0.014)	(0.012)	(0.016)	(0.014)
Constant	1.928***	1.910***	1.971***	1.957***
	(0.033)	(0.032)	(0.038)	(0.036)
Observations	1,850	2,775	1,322	1,982

Notes: Estimates come from linear regressions. Understanding ranges from 0 to 4 depending on the number of questions correctly answered. A higher value indicates a better understanding of the decision problem. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

VI.2.1.3. Treatment effects by fairness criteria

The experimental design allows for spectators to be classified by their fairness criteria in the control condition. Table 4.2c therefore reports interactions between control condition fairness criteria and treatment effects. In other words, the models reported in the table test whether certain types of spectators react more strongly to the treatment than others.

Both, choice egalitarians and weakly inequality averse spectators have a robust and highly significant positive treatment effect across both, the ex ante and the ex post treatment conditions. Interestingly, spectators who cannot be classified as either choice egalitarians or as inequality averse only have a robustly positive and significant treatment effect in the models when using a strict definition of inequality aversion. Under the 60% consistency requirement for weak inequality aversion the treatment effect for this subject group disappears entirely. This suggests that the treatment effect is entirely driven by spectators whose allocation decisions in the control group already follow a version of the two fairness criteria. This result is especially surprising considering that even under the 60% consistency requirement and the weak definition of inequality aversion, still 46 out of 185 spectators cannot be classified as following a particular fairness criterion so given that each spectator contributes a total of 20 allocation decisions, the lack of a significant treatment effect for this group is unlikely to be an issue of power. Additionally, the coefficients are substantially smaller and at times negative for this group, again suggesting that there really is no treatment effect in either the ex post or ex ante conditions for this group of spectators.

This finding also provides additional support for the proposed theoretical model assuming that spectators have a given fairness criterion in the control condition which they trade-off against a reward parameter in the treatment conditions.

Table 4.2c: Treatment interactions with fairness criteria

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
80% consistency				
Treatment x CE	0.137***	0.158***	0.165***	0.201***
	(0.050)	(0.049)	(0.061)	(0.060)
Treatment x strict IA	-0.048	0.119	0.091	0.290***
	(0.120)	(0.097)	(0.130)	(0.104)
Treatment x NF	0.174***	0.161***	0.227***	0.192***
	(0.054)	(0.046)	(0.062)	(0.052)
Treatment x CE	0.137***	0.158***	0.165***	0.201***
	(0.050)	(0.049)	(0.061)	(0.060)
Treatment x weak IA	0.269***	0.199***	0.413***	0.354***
	(0.070)	(0.059)	(0.076)	(0.066)
Treatment x NF	0.092	0.134***	0.100	0.119**
	(0.061)	(0.049)	(0.071)	(0.057)
60% consistency				
Treatment x CE	0.131**	0.130***	0.199***	0.123**
	(0.052)	(0.047)	(0.059)	(0.056)
Treatment x strict IA	0.052	0.170*	0.190	0.357***
	(0.109)	(0.090)	(0.118)	(0.097)
Treatment x NF	0.176***	0.167***	0.214***	0.212***
	(0.060)	(0.049)	(0.068)	(0.056)
Treatment x CE	0.131**	0.130***	0.199***	0.123**
	(0.052)	(0.047)	(0.059)	(0.056)
Treatment x weak IA	0.268***	0.211***	0.359***	0.328***
	(0.060)	(0.051)	(0.067)	(0.057)
Treatment x NF	-0.046	0.085	-0.086	0.041
	(0.093)	(0.068)	(0.109)	(0.080)
Observations	1,850	2,775	1,322	1,982

Notes: Estimates come from linear regressions. Strict inequality aversion means that spectators equalised ultimate payoffs while weak inequality aversion means spectators gave more to the decision maker who had a lower initial payoff. Consistency relates to the number of decisions spectators made in line with the particular criterion. NF refers to spectators who cannot be classified by either of the fairness criteria in the control condition. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

VI.2.1.4. Treatment effects with demographic controls

Although table 4.1d already reports main treatment effects with controls based on the unbalanced variables reported in table 4.1c, table 4.2d below reports main treatment effects for the full sample and for scenarios with the safe option above $1 including all available control variables in the dataset. The first result to note is again that the main treatment effects are robust to the inclusion of this battery of control variables. The coefficients are also still remarkably similar to the effect observed in the simple treatment comparison in figure 4.3 in the main text. None of the control variables have a similarly robust and significant effect on the share allocated to the risk taker as the treatment assignment. Being employed decreases the share to a similar extent as the treatment assignment in the ex post condition. A higher degree of risk seeking preferences is also somewhat robustly associated with a higher reward for the risk taker while higher ambiguity aversion is correlated with a lower reward for the risk taker. These two findings suggest that spectators judge decision makers based on their own preferences which is not surprising. Interestingly, in the ex post condition, any type of party affiliation is associated with a higher reward for the risk taker. On the other hand, a vote for Joe Biden in 2020 is associated with a lower reward.

VI.2.2 Pre-registered interaction effects

As outlined in the pre-analysis plan, table 4.2e reports estimates of the increasing treatment effect reported in table 4.1 in the main text, interacted with outcomes over the personal lottery and externalities. The coefficients can be read as the increase in the amount allocated to the risk-taker for a given outcome and treatment group, if the safe option increases by 50ct.

The reported results indicate, similarly to table 4.2, that spectators show outcome bias over the externality in the ex post treatment group. The coefficient for the scenarios in which the externality realised is somewhat larger than if the externality did not realise. Both are highly significant. This suggests that spectators increase the reward they allocate to risk-takers more as the value of the safe option rises, if the externality realised.

Table 4.2d: Treatment effects with complete demographic controls

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
Treatment	0.127**	0.136***	0.196***	0196***
	(0.050)	(0.044)	(0.057)	(0.050)
Demographics				
Age	-0.002	-0.003*	-0.001	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Gender	0.025	-0.052	-0.036	-0.060
	(0.055)	(0.045)	(0.064)	(0.051)
Education	-0.034*	0.005	-0.044**	0.012
	(0.018)	(0.014)	(0.020)	(0.017)
Income	0.007	0.012*	0.014	0.012
	(0.008)	(0.007)	(0.010)	(0.008)
Employment	-0.045	-0.140***	-0.029	-0.193***
	(0.062)	(0.051)	(0.071)	(0.058)
Economics	0.094*	0.019	0.093	-0.007
	(0.056)	(0.046)	(0.064)	(0.052)
Ethnicity				
American Indian or Alaska Native	0.210	0.384**	0.156	0.427**
	(0.268)	(0.178)	(0.313)	(0.186)
Asian	0.079	0.119	0.066	0.207**
	(0.109)	(0.087)	(0.126)	(0.097)
Black or African-American	0.042	-0.092	0.044	-0.102
	(0.110)	(0.088)	(0.131)	(0.099)
Native Hawaiian or Pacific Islander	-	-	-	-
	-	-	-	-
Spanish, Hispanic or Latino	-0.111	-0.014	-0.050	0.038
	(0.085)	(0.069)	(0.095)	(0.076)
White	0.070	0.088	0.078	0.100
	(0.090)	(0.073)	(0.101)	(0.081)
Preferences				
Risk seeking	0.031**	0.058***	0.019	0.039***
	(0.012)	(0.010)	(0.014)	(0.012)
Ambiguity aversion	-0.008	-0.012***	-0.015***	-0.017***
	(0.005)	(0.004)	(0.006)	(0.005)
Left-Right	-0.015	0.008	-0.006	0.004
	(0.015)	(0.013)	(0.017)	(0.014)
Party affiliation				
Democrats	0.491***	0.677***	0.342	0.583***
	(0.157)	(0.124)	(0.210)	(0.160)
Republicans	0.302*	0.380***	0.160	0.379**
	(0.165)	(0.128)	(0.214)	(0.161)
Other	0.292*	0.374***	0.091	0.321**
	(0.159)	(0.125)	(0.211)	(0.159)
2020 Vote				
Joe Biden	-0.396***	-0.293***	-0.247*	-0.194*
	(0.117)	(0.093)	(0.141)	(0.105)
Donald Trump	-0.165	-0.115	0.050	-0.040
	(0.134)	(0.105)	(0.159)	(0.113)
Other	-0.080	0.393**	-0.003	0.305*
	(0.209)	(0.157)	(0.246)	(0.183)
Didn't vote	-0.259**	-0.184*	-0.165	-0.168
	(0.120)	(0.096)	(0.147)	(0.107)
Constant	2.014***	1.557***	2.154***	1.782***
	(0.237)	(0.197)	(0.286)	(0.232)
Observations	1,680	2,520	1,201	1,807

Notes: Estimates come from linear regressions. Detailed descriptions of each variable can be found in appendix section VI.3. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Table 4.2e: Increasing treatment effect interactions with outcomes

	Amount allocated to Risk-Taker			
	Externality		Lottery	
	✓	x	✓	x
Control x Level x			-0.025	0.195***
			(0.023)	(0.029)
Ex Ante x Level x			0.050**	0.268***
			(0.023)	(0.028)
Ex Post x Level x	0.181***	0.127***	0.031	0.279***
	(0.022)	(0.022)	(0.020)	(0.023)
Observations	3,700	3,700	3,700	3,700

Notes: Estimates come from linear regressions. Coefficients report the effect of a particular treatment assignment interacted with the level of the safe option and interacted with a particular outcome. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

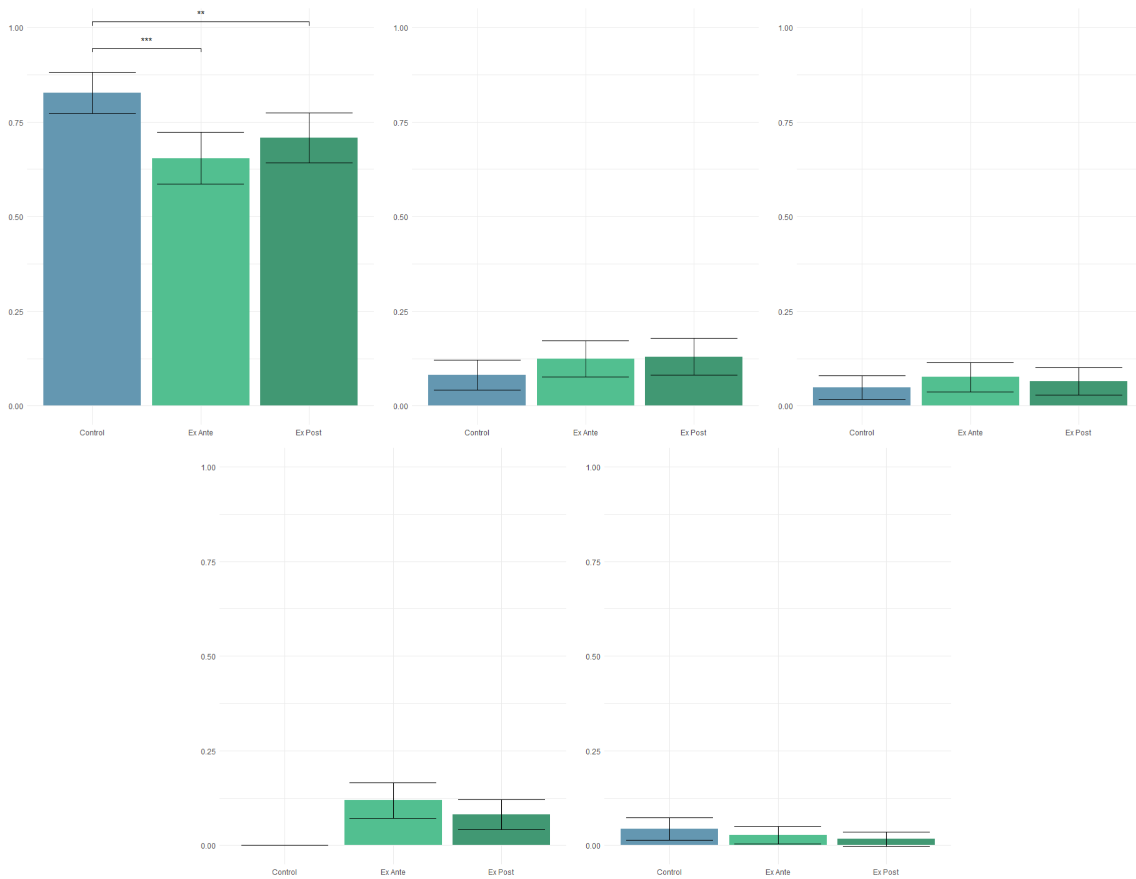
Turning to the results of the interaction effects with outcomes over the risk-takers personal lottery, a similar pattern across all treatment groups can be observed. While spectators increase the reward of the risk-taking decision maker as the value of the safe option rises if the personal lottery was unsuccessful, there is no significant increase if the personal lottery was successful. Importantly, this increase in the reward of the risk-taking decision maker is somewhat larger in the treatment groups than in the control group, if the personal lottery was unsuccessful. While, on average, spectators do not reward risk-takers more if the personal lottery was unsuccessful in the treatment compared to the control group, as reported in table 4.2, they seem to nonetheless respond more to the opportunity cost of choosing risk in the treatment group.

VI.2.3 Full Results of Belief elicitation

Figure 4.2a illustrates spectators' responses to the question "Why do you think some participants chose the risky option B?". Each panel reports the share of spectators by treatment condition who selected a particular option as the most important in answering that question.

Spectators were asked this question after each treatment or control block which allows for this comparison of answers across conditions. As already reported in the main text, the only significant treatment difference can be observed for the option "To maximize their own

Figure 4.2a: Full Spectator Beliefs about why decision makers took risk

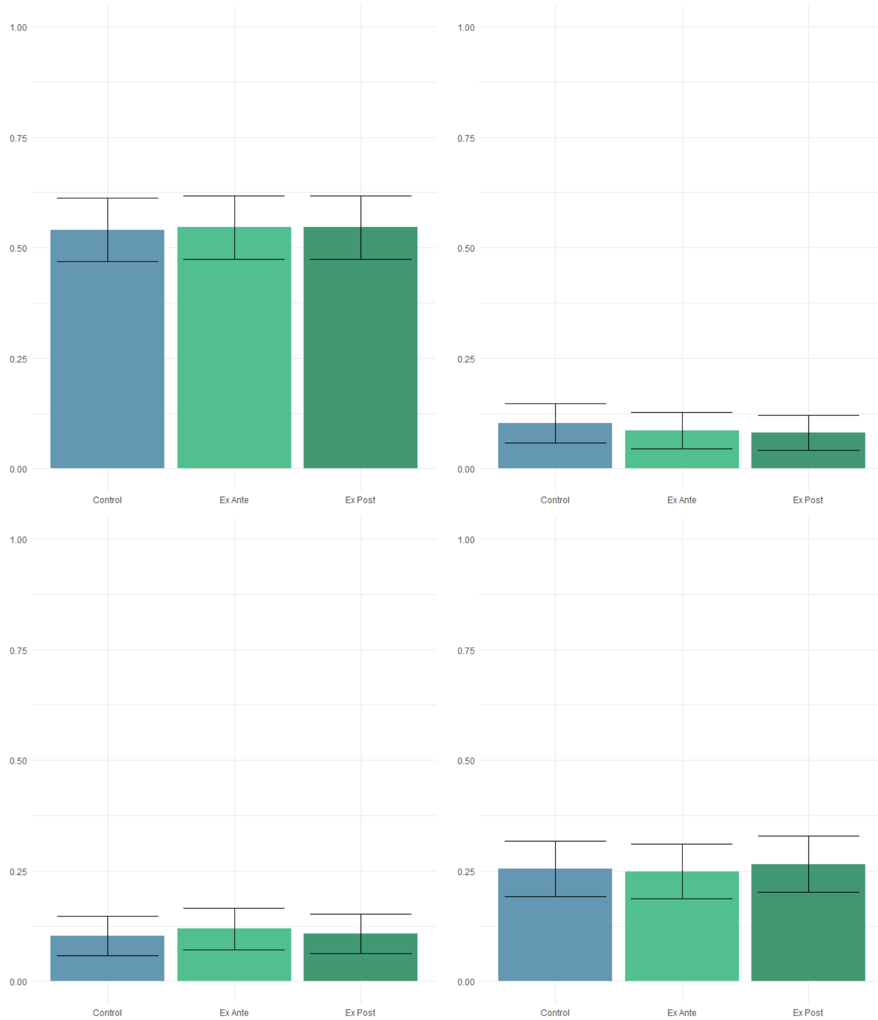


Notes: Spectators were asked why they think some participants chose the risky option B. From top left to bottom right, the panels report the share of spectators by treatment conditions who selected the option "To maximise their own payoff", "To maximize the pair's payoff", "To influence your allocation decision of the additional $4", "To generate the additional $2 for the other participants", "Other". The only significant difference between treatments can be observed for the first statement and the fourth, which was however only an option in the treatment conditions.

payoff". Spectators are less likely to select this option as the most important one in answering the question in either treatment condition compared to the control conditions. This difference can be explained by the option "To generate the additional $2 for the other participants", which is only an option in the treatment conditions but selected by a small share of spectators as the most important option in explaining decision makers' risk-taking.

After each treatment or control block of allocation decisions, spectators were also asked why they believe some decision makers took the safe option A. Figure 4.2b reports in each panel the share of spectators by treatment condition who selected a particular option as the most

Figure 4.2b: Full Spectator Beliefs about why decision makers took safe option

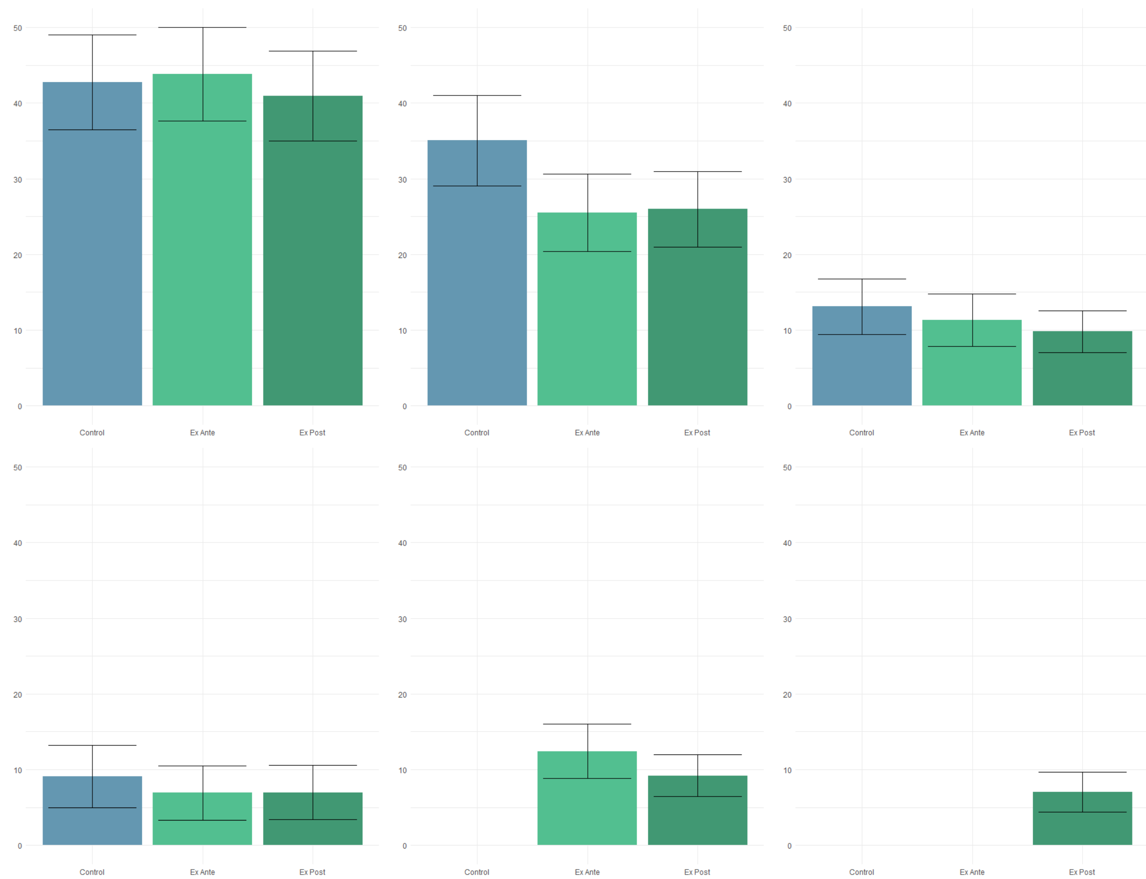


Notes: Spectators were asked why they think some participants chose the safe option A. From top left to bottom right, the panels report the share of spectators by treatment conditions who selected the option "To maximise their own payoff", "To maximize the pair's payoff", "To influence your allocation decision of the additional $4", "Other". There are no significant differences between treatments for any of the statements.

important in answering that question. Compared to figure 4.2a, it is striking how consistent spectator responses are across treatment conditions. There are no significant differences but the means are also almost identical across the conditions. This result suggests that there are no spillovers in beliefs about the decision maker who did not take the risk when externalities become part of the decision problem.

Spectators were also asked after each treatment or control condition block, which attributes

Figure 4.2c: Full Spectator Beliefs about allocation decisions



Notes: Spectators were asked which attributes they considered when making their allocation decisions. From top left to bottom right, the panels report the share of spectators by treatment conditions who selected the option "Inequality in earnings between participants", "The choices participants made", "The outcome of the personal lottery", "Other", "The potential benefit for the participant pool of choosing option B", "Whether the additional $2 generated for the participant pool". The fifth option was added for both treatment conditions and the sixth only for the ex post treatment condition. There are no significant differences between treatments for any of the statements.

they considered when making their allocation decisions. Figure 4.2c reports the share of spectators who selected a particular attribute as most important by treatment condition.

The most selected option across all treatment conditions is "Inequality in earnings between participants". Between 40-45% of spectators select this option as the most important in each condition. The second most chosen option is "The choices participants made". Here, although not statistically significant, there is a substantial drop of almost 10% in the share of spectators who select this option as most important between the control and treatment conditions. This difference can be explained by the small share of spectators who opt for the options ”The potential benefit for the participant pool of choosing option B” or ”Whether the additional \$2 generated for the participant pool” in the treatment conditions.

These beliefs are interestingly in line with the results reported in table 4.2c as spectators who are choice egalitarians in revealed and stated preferences seemingly trade-off their fairness criterion with a consideration for rewarding prosocial risk-taking in the treatment conditions.

VI.3 Description of Variables

Share b. Continuous variable indicating the absolute share of the \$4 bonus spectators allocated to the decision maker who chose the risky option B.

Age. Continuous variable indicating the self-reported age of subject i.

Gender. Dummy variable coded as 1 if subject i indicated to be male, 0 if subject i indicated to be female. Subjects who indicated ”other” or ”prefer not to answer” were coded as missing values.

Education. Categorical variable capturing subject i's highest level of education.

0: Primary education or less

1: Some high school

2: High school degree/GED

3: Some college

4: 2-year college degree

5: 4-year college degree

6: Master's degree

7: Professional degree (JD, MD, MBA)

8: Doctoral degree

Income. Categorical variable capturing the income bracket of subject i. Values are stated in US Dollar (\$).

1: Under \$10,000

2: \$10,000 to \$20,000

3: \$20,001 to \$30,000

4: \$30,001 to \$40,000

5: \$40,001 to \$50,000

6: \$50,001 to \$60,000

7: \$60,001 to \$80,000

8: \$80,001 to \$100,000

9: \$100,001 to \$150,000

10: \$150,001 to \$200,000

11: \$200,001 to \$350,000

12: \$350,001 to \$500,000

13: Above \$500,000

Employment. Dummy variable coded as 1 if subject i indicated that they are in either full or part time employment or a business owner and 0 if subject i indicated to be unemployed, retired or a student.

Economics. Dummy variable coded as 1 if subject i indicated that they have taken a module in economics or a related subject at University. A value of 0 indicates that subject i either has not taken a module in economics or has never attended higher education.

Left-Right Placement. Continuous variable capturing subject i's self-reported placement on a political left-right scale with 0 indicating "Left" and 10 indicating "Right".

Party affiliation. Categorical variable capturing subject i's self-reported political party affiliation. Those who responded with "Don't know" were coded as missing variables.

1: Democratic Party

2: Republican Party

3: Other

2020 Vote. Categorical variable capturing subject i's self-reported vote in the 2020 Presidential election. Those who responded with "Don't remember" or "Prefer not to say" were coded as missing variables.

1: Joe Biden

2: Donald Trump

3: Other candidate

4: Didn't vote

Risk seeking. Variable capturing subject i's willingness to take risks on a scale from 0 to 10, where 0 means "completely unwilling to take risks" and a 10 means "very willing to take risks".

Ambiguity aversion. Variable capturing subject i's degree of ambiguity aversion on a scale from 4 to 28. This variable combines subjects' responses to the four questions of item D12 which can be found in appendix VI.4 and is taken from the module on measuring ambiguity aversion in (Cavatorta and Schröder 2019). The possible responses to each question range from 1 indicating "strongly agree" to 7 indicating "strongly disagree". A higher value on this scale therefore indicates a lower degree of ambiguity aversion.

VI.4 Experimental Instrument

VI.4.1 Stage I: Decision Makers

Decisions 1-7 are presented in randomized order during the experiment and vary in the amount of the certain option A (from \$0.5 in decision 1 to \$3.5 in decision 7). Below, only decision 3 (with a certain option of \$1.50) is shown as an example. Text which is only shown in the treatment condition is highlighted.

Introduction:

Thank you for participating in this study. In the following, you will be asked to make a number of decisions that will influence the bonus payment you can receive for this study. Specifically, you will be asked seven times to decide between two options. Option A is always a certain payment while option B is a lottery with a potentially higher payoff. If you choose option B and irrespective of the outcome of the lottery, there is also a 50% chance that an additional \$2 will generate. If this is the case, we will use these \$2 to reward two randomly chosen participants from other studies. Below is an example decision you might face:

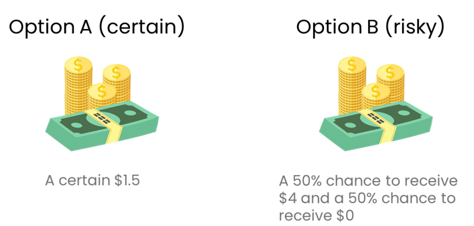


Figure 4.4a: Control Screen

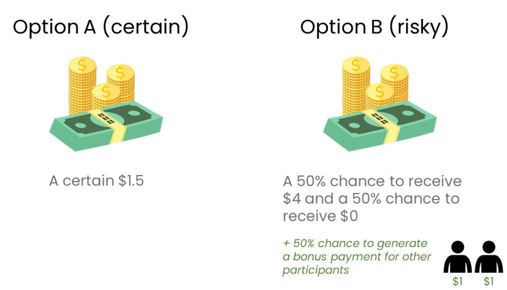


Figure 4.4a: Treatment Screen

While option B will remain the same throughout the seven decisions, the value of option A, which is a certain payment, will vary.

After you made your decisions, you will be paired with another participant and **one of the decisions** you made will be randomly selected to determine your bonus payment. All of the seven decisions you make have an equal chance of being selected for payment.

You will then be asked a set of questions about yourself and about the choices you just made.

Before you receive your bonus payment, a third participant will have the option to allocate an **additional $4** between you and the other participant as they wish. Depending on the choices you make and the decision of the third participant, you can therefore receive **a total of $8 in bonus payments** in this study. To receive a bonus payment, you need to complete the full study. This should take about 10 minutes.

Understanding:

Before you make your first decision, please answer the following questions. Your final payment will not depend on your answers to these questions. However, please answer to the best of your ability as your answers will impact the quality of our research.[8]

U1: How many decisions are you asked to make?

- 5
- 7
- 10
- Don't know

U2: Which one of the two options remains the same throughout the decisions?

- The certain option A
- The risky option B
- Don't know

U3: What will determine your final bonus payment? Please select all that apply.

- The sum of all the choices you make
- One randomly selected choice

[8]After subjects have submitted their answers to U1-U4, the correct answers will be displayed before they can proceed to the next page.

- The allocation decision of \$4 by a third participant between you and the other participant in your pair
- The sum of all the choices made by the other participant in your pair
- Don't know

U4: If you choose the risky option B, how high is the chance that \$2 will generate that we will use to pay additional participants?

- 0%
- 25%
- 50%
- 75%
- 100%
- Don't know

Example Decision:

Please carefully consider the below two options. The option you choose has a chance of 1 in 7 to determine your bonus payment for this study. Once you leave this screen, you cannot change the decision you made. Please remember that if you choose option B and irrespective of the outcome of the lottery, there is also a 50% chance that an additional $2 will generate that will be used to reward two randomly chosen participants from other studies. Which option would you like to choose?

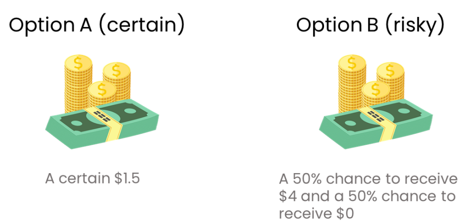


Figure 4.4b: Control Screen

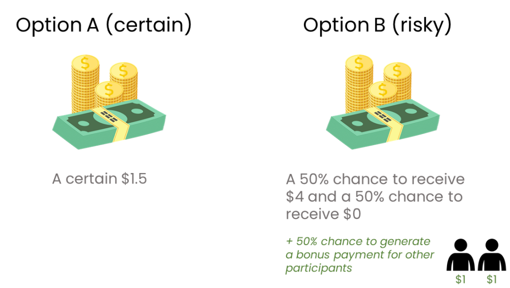


Figure 4.4b: Treatment Screen

Confidence: On a scale from 0 to 10, please indicate how confident you are in the decision you just made.

Reasoning Questions:

J1: Please explain your reasoning for the decisions you just made:

J2: Did you choose option B in any of the decisions:

- Yes
- No

J3: *[if previous question Yes]* Why did you choose option B in those cases? Please select all that apply.

- To maximise my own payoff
- To generate the additional $2 for the other participants
- To influence the allocation decision of the $4 by the third participant
- Other

J4: Which of the reasons you selected for choosing option B was most important for your decision making?

- To maximise my own payoff
- To generate the additional $2 for the other participants
- To influence the allocation decision of the $4 by the third participant
- Other
- All were of equal importance

Belief Questions:

B1: Please assume for now that the below decision will get selected for payment and the other participant in your pair did not choose the same option as you. Remind yourself of the choice you made. How much of the additional \$4 do you expect the third participant would allocate to you in this case if they were fully informed about your and the other participant's choices?

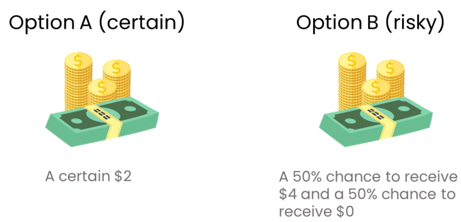


Figure 4.4c: Control Screen

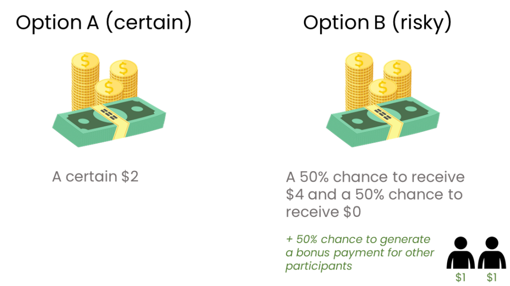


Figure 4.4b: Treatment Screen

Confidence: On a scale from 0 to 10, how confident are you in the estimate you just provided?

B2: On a scale from 0 to 10, how much autonomy do you feel you have over your final earnings?

VI.4.2 Stage II: Spectators

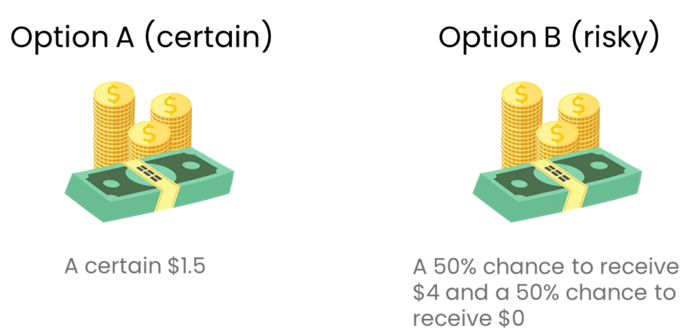
Decisions 1-20 are presented in randomized order during the experiment. Below, one decision from each treatment and control condition is shown as an example.

Introduction:

Thank you for participating in this study. In the following, you will be asked to decide on an allocation of money between two participants. You will be asked to make a total of 20 allocation decisions. While 19 of these decisions are hypothetical, 1 will determine the actual payment for a participant pair. Please consider each decision carefully as we will not inform you about which of the 19 decisions is the one that determines actual payoffs for the two participants. After making the 20 allocation decisions you will be asked a set of questions about yourself and about the choices you just made. You will also have the opportunity to earn an additional bonus payment.

Control Condition:

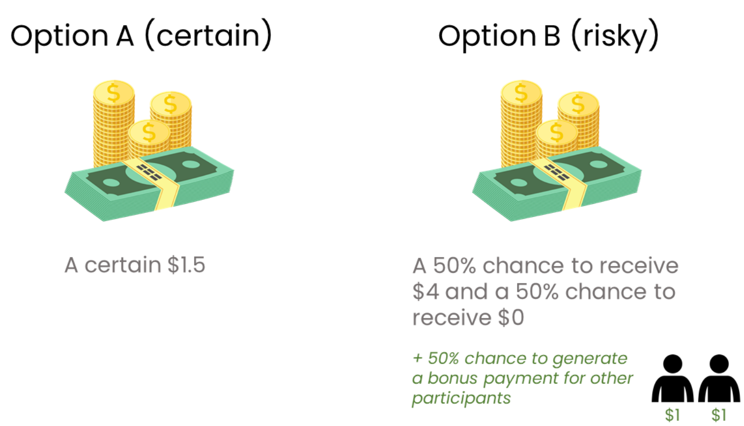
You will now be asked to make your first/next five allocation decisions. All of the participants within the pairs already had the chance to receive an income based on a choice they were asked to make. Specifically, they were asked to decide between two options. Option A is always a certain payment while option B is a lottery with a potentially higher payoff. Below is an example of a decision they might have faced:



While option B always remains the same across all potential scenarios you will face, the value of option A, which is a certain payment, may vary. You will be informed about **the decision the participants faced, the choice they made, and their resulting current earnings**. You then have the option to allocate an additional $4 between the pair.

Ex ante Treatment Condition:

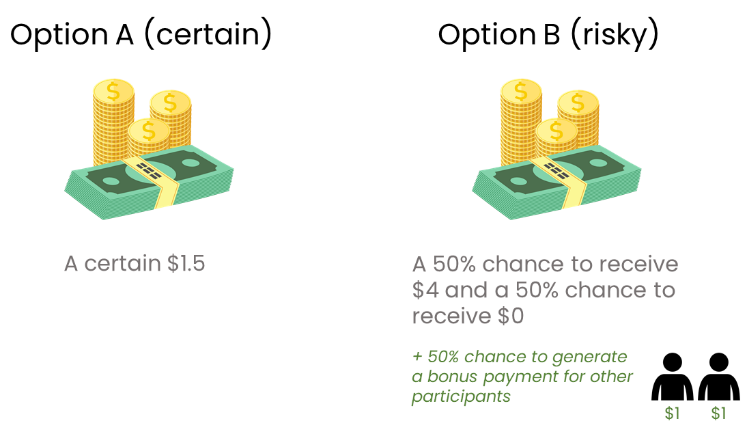
You will now be asked to make your first/next five allocation decisions. All of the participants within the pairs already had the chance to receive an income based on a choice they were asked to make. Specifically, they were asked to decide between two options. Option A is always a certain payment while option B is a lottery with a potentially higher payoff. If a participant chose option B and irrespective of the outcome of the lottery, there was also a 50% chance that an additional $2 generated. In these cases, the $2 have been used to reward two randomly chosen participants from other studies. Participants were aware of this possibility when making their decisions. Below is an example of a decision they might have faced:



While option B always remains the same across all potential scenarios you will face, the value of option A, which is a certain payment, may vary. You will be informed about **the decision the participants faced, the choice they made, and their resulting current earnings** but **not** about whether the **additional $2** generated. You then have the option to allocate an additional $4 between the pair.

Ex post Treatment Condition:

You will now be asked to make your first/next ten allocation decisions. All of the participants within the pairs already had the chance to receive an income based on a choice they were asked to make. Specifically, they were asked to decide between two options. Option A is always a certain payment while option B is a lottery with a potentially higher payoff. If a participant chose option B and irrespective of the outcome of the lottery, there was also a 50% chance that an additional \$2 generated. In these cases, the \$2 have been used to reward two randomly chosen participants from other studies. Participants were aware of this possibility when making their decisions. Below is an example of a decision they might have faced:



While option B always remains the same across all potential scenarios you will face, the value of option A, which is a certain payment, may vary. You will be informed about **the decision the decision the participants faced, the choice they made, their resulting current earnings, and whether the additional \$2 generated**. You then have the option to allocate an additional \$4 between the pair.

Understanding:

Before you make your first decision, please answer the following questions. Your final payment will not depend on your answers to these questions. However, please answer to the best of your ability as your answers will impact the quality of our research.[9]

U1: How many decisions are you asked to make in total?

- 10
- 20
- 30
- Don't know

U2: Which one of the two options remains the same throughout the decisions?

- The certain option A
- The risky option B
- Don't know

U3: How many of your decisions will result in an actual payment for the two participants?

- 0
- 1
- 5
- 10
- Don't know

U4: If a participant chose the risky option B, how high is the chance that \$2 will generate that we will use to pay additional participants?[10]

- 0%

[9]After subjects have submitted their answers to U1-U4, the correct answers will be displayed before they can proceed to the next page.

[10]This questions is displayed together with U1-U3 if spectators are randomly allocated to see one of the treatment conditions first. If spectators see the control conditions first, this question is displayed on its own after the introduction of the second condition.

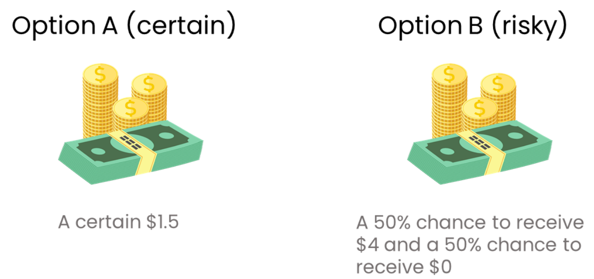
- 25%
- 50%
- 75%
- 100%
- Don't know

Reminder:

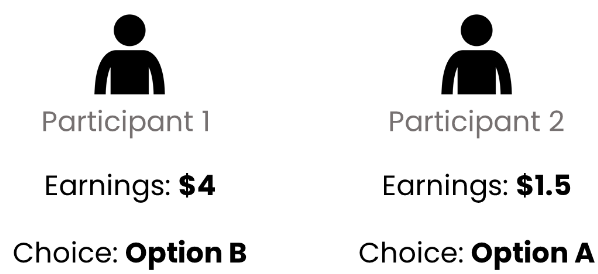
You will now make five/ten allocation decisions. Please remember that one of your decisions will determine the real payment of two individuals who participated in this study.

Control Decision Example:

Please carefully consider the below scenario. Participant 1 and 2 both faced the following decision:



The outcomes for both participants and the choices they made are given below:



You are now asked to allocate an additional amount of money between the pair. You can allocate this money as you wish. Please note, you have to allocate the total amount of $4.

Please specify the amount you would like to allocate to each participant. You may use up to two decimal points when specifying each amount. Please ensure the two values add up to $4 before proceeding.

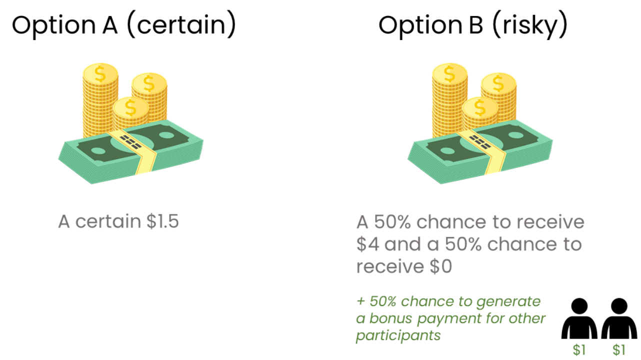
Participant 1 (in $):

Participant 2 (in $):

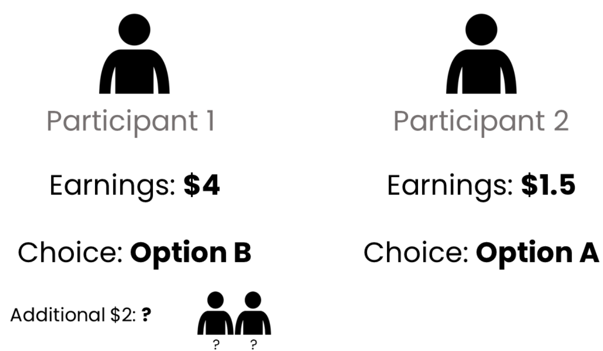
Confidence: On a scale from 0 to 10, please indicate how confident you are in the decision you just made.

Ex Ante Treatment Decision Example:

Please carefully consider the below scenario. Participant 1 and 2 both faced the following decision:



The outcomes for both participants and the choices they made are given below. Whether the additional $2 generated or not will be revealed after your allocation decision:



You are now asked to allocate an additional amount of money between the pair. You can allocate this money as you wish. Please note, you have to allocate the total amount of $4.

Please specify the amount you would like to allocate to each participant. You may use up to two decimal points when specifying each amount. Please ensure the two values add up to $4 before proceeding.

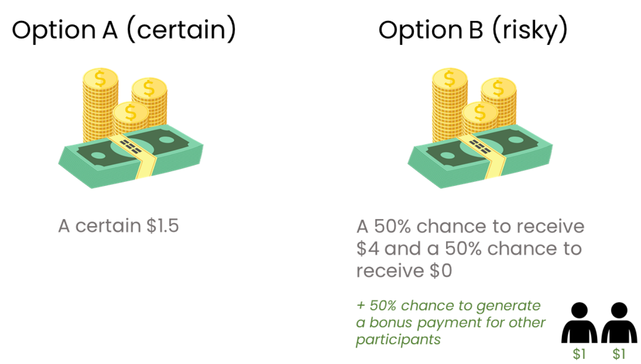
Participant 1 (in $):

Participant 2 (in $):

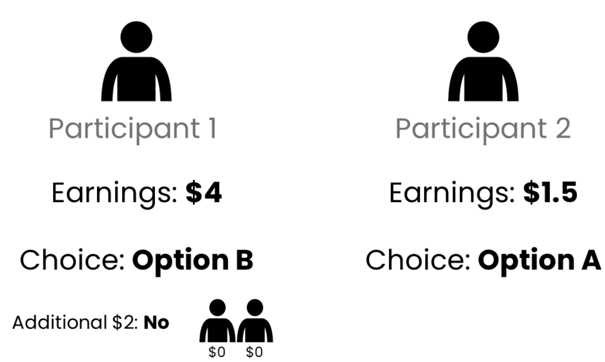
Confidence: On a scale from 0 to 10, please indicate how confident you are in the decision you just made.

Ex Post Treatment Decision Example:

Please carefully consider the below scenario. Participant 1 and 2 both faced the following decision:



The outcomes for both participants and the choices they made are given below. Because participant 2 chose the risky option B, an additional $2 generated and two randomly chosen participants from other studies **have been given** a $1 bonus payment ($2 in total) / Despite participant 2 having chosen option B, the additional $2 did not generate and the two randomly chosen participants from other studies **have not been given** a $1 bonus payment ($2 in total):



You are now asked to allocate an additional amount of money between the pair. You can allocate this money as you wish. Please note, you have to allocate the total amount of $4.

Please specify the amount you would like to allocate to each participant. You may use up to two decimal points when specifying each amount. Please ensure the two values add up to $4 before proceeding.

Participant 1 (in $):

Participant 2 (in $):

Confidence: On a scale from 0 to 10, please indicate how confident you are in the decision you just made.

Beliefs and Preferences[11]

B1: How did you decide on the allocation of income within the participant pairs?

B2a: Which of the following attributes did you consider when making your allocation decisions? Please select all that apply.

- Inequality in earnings between participants
- The choices participants made
- The outcome of the personal lottery in option B
- Whether the additional $2 generated for the participant pool[12]
- The potential benefit for the participant pool of choosing option B
- Other

B2b: How important were each of the attributes you just selected for your allocation decisions? Please allocate a total of 100 points across the attributes you selected. Please ensure that the more important an attribute was to your decision making, the more points you allocate to it.

B3a: Why do you think some participants chose the risky option B? Please select all that apply.

- To maximise their own payoff

[11] B1-B4 are asked after each condition. Highlighted options are only displayed in the two treatment conditions. B5-B9 are asked after all three conditions were completed.

[12] This option is only included in the ex ante treatment condition.

- To maximise the pair's payoff
- To generate the additional $2 for the other participants
- To influence your allocation decision of the additional $4
- Other

B3b: Which of the reasons you just selected do you think was the main reason why some participants chose the risky option B?

B3conf: On a scale from 0 to 10, please indicate how confident you are that this was the main reason why some participants chose the risky option B.

B4: Why do you think some participants chose the certain option A? Please select all that apply.

- To maximise their own payoff
- To maximise the pair's payoff
- To influence your allocation decision of the additional $4
- Other

B4b: Which of the reasons you just selected do you think was the main reason why some participants chose the certain option A?

B4conf: On a scale from 0 to 10, please indicate how confident you are that this was the main reason why some participants chose the certain option A.

B5: What percentage of decision makers do you believe chose the risky option B in the below scenarios? You will receive a bonus payment of 10ct for each estimate that is within +/-5 percentage points of the correct answer. If all of your estimates are correct, you will therefore be able to earn an additional bonus payment of $1.[13]

I think the percentage of decision makers who chose the risky option B in the scenario where

[13]Each spectator is asked to answer B5 for five randomly selected values of A in the control and treatment condition.

the certain option A was [**x**] and it was **not** possible to generate the additional $2 is:

10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

I think the percentage of decision makers who chose the risky option B in the scenario where the certain option A was [**x**] and it **was** possible to generate the additional $2 is:

10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

B6a: Which of the following do you believe apply? Please choose all options that you agree with. The $2 for the two other randomly chosen participants:

- Will decrease income inequality between all participants on Prolific Academic
- Will increase income inequality between all participants on Prolific Academic
- Will benefit the two participants who receive it
- Will not matter to the two participants who receive it
- Is unfair because they did nothing to receive it
- Will increase the total amount of money all participants on Prolific Academic have combined

B6b: Which of the statements you just selected do you agree with most? The $2 for the two other randomly chosen participants:

B7: On a scale from 0 (lower) to 10 (higher), what do you believe is the effect of the $2 the two randomly chosen participants might receive, if a decision maker chose option B, on income inequality between all participants on Prolific Academic?

B8: Do you believe risk-taking should be rewarded?

- Yes
- No
- I don't know
- I don't have an opinion on this

B9: Do you believe risk-taking for the benefit of society should be rewarded?

- Yes
- No
- I don't know
- I don't have an opinion on this

VI.4.3 Demographics

In this final part of the study, we will ask you a few of questions about yourself. Please read the questions carefully and answer honesty. This part should take only 2-3 minutes.

D1: How old are you?

D2: What is your gender?

- Female
- Male
- Other
- Prefer not to answer

D3: To which of these groups do you consider you belong? You can choose more than one group.

- American Indian or Alaska Native
- Asian
- Black or African-American
- Native Hawaiian or other Pacific Islander
- Spanish, Hispanic or Latino
- White
- Other group
- Prefer not to answer

D4: Which category best describes your highest level of education?

- Primary education or less
- Some high school

- High school degree/GED
- Some college
- 2-year college degree
- 4-year college degree
- Master's degree
- Doctoral degree
- Professional degree (JD, MD, MBA)
- Prefer not to answer

D5: What is your total (annual) household income before tax?

- Under $10,000
- $10,000 - $20,000
- $20,001 - $30,000
- $30,001 - $40,000
- $40,001 - $50,000
- $50,001 - $60,000
- $60,001 - $80,000
- $80,001 - $100,000
- $100,001 - $150,000
- $150,001 - $200,000
- $200,001 - $350,000
- $350,001 - $500,000

- Above \$500,000
- Don't know
- Prefer not to answer

D6: What is your current employment status?

- Full-time employee
- Part-time employee
- Self-employed or small business owner
- Medium or large business owner
- Unemployed and looking for work
- Student
- Not currently working and not looking for work (e.g. full-time parent)
- Retiree
- Prefer not to answer

D7: Have you ever taken a module on economics or a related subject area at university?

- Yes
- No
- I have never attended higher education

D8: In politics people sometimes talk of left and right. Where would you place yourself on the following scale?

D9: Which party do you feel closest to?

- Democratic Party
- Republican Party

- Other
- Don't know

D10: Who did you vote for in the recent 2020 Presidential Election?

- Joe Biden
- Donald Trump
- Other candidate
- Didn't vote
- Don't remember
- Prefer not to say

D11: Please tell us, in general, how willing or unwilling you are to take risks. Please use a scale from 0 to 10, where 0 means "completely unwilling to take risks" and a 10 means you are "very willing to take risks". You can also use any numbers between 0 and 10 to indicate where you fall on the scale.

D12: Please respond to the following statements by indicating the extent to which you agree or disagree with them on a scale from 1 (I strongly agree) to 7 (I strongly disagree).

- There is a right way and a wrong way to do almost everything
- Practically every problem has a solution
- I feel relieved when an ambiguous situation suddenly becomes clear
- I find it hard to make a choice when the outcome is uncertain

D13: Do you have any feedback or impressions regarding this study?

VI.5 Pre-Analysis Plan

VI.5.1 Motivation

Many decisions people make have the potential to create positive externalities. This is especially the case for decisions that are made under uncertainty: Starting a business, developing new technologies, or investing into new ventures, are all decisions that involve uncertainty about personal gains and losses, but have potential positive externalities for wider society. For example, a successful entrepreneur may create jobs, knowledge spillovers, or improve welfare through her products and services. The likelihood of a person choosing such entrepreneurial activities is however affected by whether and how societies reward this decision.[14] One channel that may affect such rewards are individuals' distributive preferences. If individuals hold a preference to reward risk takers who create positive externalities for society, then this creates demand for policies in line with such preferences. How potential externalities affect distributive preferences is however an open question. In this study, I will provide a theoretical framework and a first experimental test of how potential externalities affect distributive preferences when income is earned under uncertainty.

VI.5.2 Experimental Design

My basic experimental design follows in particular Cappelen et al. (2013a) by asking an impartial spectator to decide on a fair allocation of a monetary bonus between two decision makers. Prior to spectators making their distributive choice, decision makers have to choose between a lottery and a safe income. In the treatment condition, choosing the lottery does not just yield a potentially high reward for the decision maker if successful, but might also result in positive externalities for the participant pool.

Stage 1: Participants recruited via Prolific Academic participate in either the control or treatment condition that determines individual payoffs.

Control: Subjects are asked to decide between the following two options seven times, one choice being randomly selected for payment, each time with a slightly different value

[14] For example, providing entrepreneurs with tax breaks has been shown to incentivise similar entrepreneurial activities (e.g. Djankov et al. 2010; Da Rin et al. 2011; Venâncio et al. 2020).

for option 1:

Option 1: \$0.5/\$1/\$1.5/\$2/\$2.5/\$3/\$3.5

Option 2: A 50% chance to receive \$4 and a 50% chance to receive \$0.

Treatment: Subjects are asked to decide between the following two options seven times, one choice being randomly selected for payment, each time with a slightly different value for option 1:

Option 1: \$0.5/\$1/\$1.5/\$2/\$2.5/\$3/\$3.5

Option 2: A 50% chance to receive \$4 and a 50% chance to receive \$0. Irrespective of the outcome of the lottery, there is a 50% chance that an externality of \$2 will generate which will be used to reward two randomly chosen participants from other studies.

Stage 2: Participants are paired based on the procedure outlined in IV.2 and one of the seven decisions subjects made is randomly chosen to determine payoffs. Each pair consists of a subject that chose the lottery and a subject that chose the safe option for the randomly selected decision. This is however unknown to subjects prior to making their choice, to ensure that subjects cannot make strategic decisions about the likely composition of the pair. Importantly, only subjects within the same treatment condition are matched.

Stage 3: Impartial Spectators, who have not participated in the first two stages of the experiment, are asked to allocate a windfall bonus of \$4 between pairs of decision makers in one control and two treatment conditions. Each spectator makes allocation decisions in all three conditions, but the order in which spectators see these conditions is randomized:

Control condition: Spectators allocate the \$4 between pairs of decision makers from the control group in stage 1. They receive full information of the choices made and the resulting earnings of each decision maker.

Ex ante treatment condition: Spectators allocate the \$4 between pairs of decision makers from the treatment group in stage 1. They receive full information of the choices made and on the resulting earnings of the decision makers but no information on whether the externality realised or not.

Ex post treatment condition: Spectators allocate the \$4 between pairs of decision makers from the treatment group in stage 1. They receive full information of the choices made, the resulting earnings of the decision makers, and on whether the externality realised or not.

Figure 4.5a shows example scenarios spectators might face during the experiment. Importantly, spectators only compare the decisions of two subjects for the same choice set, i.e., the choices made when faced with the same safe and risky option. They also do not have the option to communicate with the participants.

Overall, spectators make 20 such allocation decisions - 5 in the control condition, 5 in the ex ante treatment condition, and 10 in the ex post treatment condition - with 19 being hypothetical and one resulting in actual payoffs for a participant pair. Spectators however do not know which of the allocation decisions will result in actual payoffs when making their decisions.[15]

VI.5.3 Research Question and Hypotheses

The primary research question of this study is *'How do positive externalities affect distributive choices for income earned under uncertainty?'*. I make several theoretical predictions:

Hypothesis 1a: The share of the bonus allocated to the subject choosing the lottery is higher in the ex ante treatment than in the control condition.[16]

Hypothesis 1b: The treatment effect in H1a increases as the value of the safe option A increases.[17]

Hypothesis 2a: The ex post treatment effects in H1a and H1b are larger when the externality realised than when it did not.

---

[15]This method is commonly used in experimental designs to increase the number of observed choices without affecting the behaviour of the subjects (Charness et al. 2016).

[16]That is, if the difference between the expected size of the externality and the expected personal benefit from choosing the risky option is positive or zero. Therefore, H1 will primarily be tested for decisions where option $A > \$1$.

[17]This effect could be nonlinear as it may be particularly strong in those cases where the EMV of the lottery is smaller than the EMV of the safe option (when the safe option is $> \$2$). That is, because risk takers incur an actual cost to their own expected payoff in those instances which spectators may wish to compensate.

Figure 4.5a: Example Spectator Scenarios in Control & Treatments

Notes: These figures illustrate example screens spectators face during the experiment. The top left panel shows a control condition scenario, the top right panel shows an ex ante treatment condition scenario, and the two bottom panels show ex post treatment scenarios with different outcomes for the lottery over the externality. For comparability, the personal lottery decision makers faced and its outcome is identical and positive in all four cases.

Hypothesis 2b: The ex post treatment effects in H1a and H1b are larger when the personal lottery was unsuccessful than when it was successful.

While not central to the main research question of this study, the first stage of the experimental design also allows me to test whether the choices made by decision makers are affected by the potential positive externalities of the risky option. H3 follows:

Hypothesis 3: Decision makers are more risk seeking in the treatment than in the control group.

H3 also provides a further justification for providing spectators with hypothetical as well as real decision scenarios, given that the distribution of choices is expected to be different in the treatment as opposed to the control group.

VI.5.4 Sampling

Based on a power analysis, a minimum total number of 840 spectator decisions is required for each control and treatment condition to achieve a power target of 0.9. Given that each spectator makes at least five individual allocation decisions for each condition, and to allow for potential attrition or other unforeseeable circumstances, a total of 180 spectators and 360 decision makers will be recruited. Of those 360 decision makers, 120 will be randomly allocated to the control group and 240 to the treatment group. This results in 3600 individual spectator decisions with 180 of those decisions determining the actual payoff of pairs of decision makers.

Spectators and decision makers will be recruited via Prolific Academic and the experiment will be coded in Qualtrics. Randomization will be done automatically via Qualtrics.

VI.5.4.1. Participation Criteria

Prolific Academic allows to restrict participation based on pre-defined critera. I will restrict participation to the following subjects:

1. Those currently resident in the United States.
2. Subjects who have not participated in the pilot study.

Additionally, I will exclude subjects from the main analysis if they have not completed all seven (or 20 for spectators) decisions.

VI.5.4.2. Matching of Decision Maker Pairs

As decision maker pairs have to consist of a person who chose the certain option A and a person who chose the risky option B for a given amount of option A, the matching of pairs will be conducted as follows:

1. Decision makers are grouped based on the first decision they saw during the experiment (the order of decisions is randomized for each subject).

2. Decision makers within each group receive a random number from a subset, depending on the choice they made for the given decision (either option A or option B).

3. Decision makers with the same numbers in the option A and option B subsets are paired.

4. Those decision makers who could not be paired are allocated to new groups based on the second decision they saw during the experiment.

5. Steps 2-4 are repeated until all decision makers are paired or all 7 decisions are exhausted for matching. In that case, the remaining decision makers are paired randomly with a random decision being selected, even if both decision makers chose the same option for this decision, and spectators are asked to decide on the allocation of the bonus between those pairs. These decisions may be used for additional analysis.

VI.5.5 Empirical Strategy

VI.5.5.1. Main Outcome Variable

The main outcome variable of interest to test hypotheses 1a-2b is the share of the bonus allocated to the subject in the decision pair who chose the risky option B. This is denoted as y_r and is equal to \$4-y_s, whereby y_s is equal to the share of the bonus allocated to the subject who chose the safe option A. Spectators can choose values with up to two decimal points to allow for sufficient variation in spectator choices.

VI.5.5.2. Hypothesis Testing

To estimate average treatment effects, I plan to use non-parametric Wilcoxon rank-sum tests. Additionally, I will run simple regression models with spectator-fixed effects to account for the fact that each spectator contributes 20 individual observations in my sample.
To test H1a, I then estimate the following simple model:

$$y_{r,d} = \alpha + \delta_1 TA_{r,d} + \epsilon_{r,d} \tag{4.5}$$

Here, α is equal to the share of the bonus y_r the risk-taking decision maker receives in the control condition where no externalities are present for a given decision d. δ_1 is then equal to the effect of the ex ante treatment condition ($TA_{r,d}$) on y_r. As I expect δ_1 to only be positive if the difference between the size of the expected externality and the benefit of choosing the lottery is positive, I restrict the above analysis to those decisions where option A > \$1.
To test H1b, I estimate the below model:

$$y_r = \delta_2 TA_r \times d_r + \epsilon_r \tag{4.6}$$

Here, δ_2 is equal to the combined effect of a \$1 increase in the safe option A and being in the ex ante treatment as opposed to the control condition on the share of the bonus y_r the risk-taking decision maker receives. H1b predicts that δ_2 will be positive.

To test H2a and H2b I will estimate the following models:

$$y_r = \delta_3 TP_r \times e_r + \epsilon_r \tag{4.7}$$

$$y_r = \delta_4 TP_r \times l_r + \epsilon_r \tag{4.8}$$

Here, e_r and l_r are dummy variables equal to 1 if the externality or personal lottery of the risk-taking subject realised, respectively. TP_r now refers to the ex post treatment condition. H2a predicts a positive δ_3 in equation 4.7 and H2b predicts a negative δ_4 in equation 4.8.

I will also test whether the effect predicted by H1b, and estimated in equation 4.6, increases when the externality realised as predicted by H2a; and whether it decreases when the personal lottery did not pay off, as predicted by H2b. To this end, I will estimate the following models:

$$y_r = \delta_5 TP_r \times d_r \times e_r + \epsilon_r \tag{4.9}$$

$$y_r = \delta_6 TP_r \times d_r \times l_r + \epsilon_r \tag{4.10}$$

Here, e_r and l_r are again dummy variables equal to 1 if the externality or personal lottery of the risk-taking subject realised, respectively. TP_r again refers to the ex post treatment condition. H2a predicts a positive δ_5 in equation 4.9 and H2b predicts a negative δ_6 in equation 4.10.

Finally, to test H3, I will calculate the proportion of decision makers who chose the risky option B for each level of option A and conduct a Pearson's chi-squared test of the frequency distributions for the treatment and control groups.

VI.5.5.3. Exploratory Analysis of potential mechanisms

If the previously outlined analysis supports H1a and y_r is larger in the (ex ante) treatment than control group, I plan to explore the underlying potential mechanisms for such an effect by testing whether spectator beliefs (B1-B9) differ significantly between the treatment and control conditions.

I further plan to estimate whether spectators follow particular fairness ideals following Cappelen et al. (2013a). Specifically, I will estimate the probability of spectators following an ex ante (choice egalitarian), or ex post (inequality averse) fairness ideal in their allocation decisions. I will then estimate whether the distribution of these fairness ideals differs between treatment and control conditions. I will also test whether the presence of externalities might create a trade-off between the fairness ideal of a spectator and a desire to reward the risk-taking decision maker. To this end, I will develop a theoretical model building on the model developed by Cappelen et al. (2007, 2013a, 2016) and structurally estimate its parameters.

VI.5.5.4. Heterogeneity Analysis

I plan to test for heterogeneous treatment effects along the following dimensions:

1. Level of confidence in distributive choice

3. Gender (D2)

4. Education level (D4)

5. Household income (D5)

6. Political left-right placement (D8)

7. Party affiliation (D9)

8. Risk preferences (D11)

9. Ambiguity aversion (D12)

VI.5.6 Ethics

This study has received ethical approval from King's College London. The reference number is MRSP-21/22-30100. No deception is being used in this experiment and no information is collected that would allow subjects to be personally identified.

VI.5.7 Power Analysis

Using the results of a pilot study conducted in December 2021, a simple power analysis can be conducted to estimate the required sample size for the main experiment. While none of the treatment effects reached conventional levels of significance in the pilot, the coefficients ranged from 0.298 to 0.598. I therefore assume a conservative treatment effect of 0.250, an alpha of 0.05 and use the standard deviation of the main outcome variable of 1.58 observed in the pilot sample. The figure below reports the required sample size for a given level of statistical power. To reach statistical power of 0.9 a sample size of at least 1679 is required. Importantly, all of the hypotheses tests will be conducted between the control condition and only one of the two treatment condition. Therefore, a total of 840 observations per control and treatment condition are required.

Figure 4.5b: Power Analysis: Hypothetical Treatment Effect = 0.25, SD = 1.58