



# Beliefs about inequality and the nature of support for redistribution<sup>☆</sup>

Aljosha Henkel<sup>a</sup> , Ernst Fehr<sup>b</sup> ,<sup>\*</sup> Julien Senn<sup>b</sup> , Thomas Epper<sup>c,d</sup>

<sup>a</sup> KOF Swiss Economic Institute, ETH Zurich, Leonhardstrasse 21, 8092 Zurich, Switzerland

<sup>b</sup> Department of Economics, University of Zurich, Blümlisalpstrasse 10, 8006 Zurich, Switzerland

<sup>c</sup> CNRS, IESEG School of Management, Univ. Lille, UMR 9221 - LEM - Lille Economie Management, F-59000 Lille, France

<sup>d</sup> iRisk Research Center on Risk and Uncertainty, France

## ARTICLE INFO

### JEL classification:

D31

D72

H23

H24

### Keywords:

Social preferences

Beliefs about inequality

Preferences for redistribution

Information

Inequality aversion

## ABSTRACT

Do beliefs about inequality depend on distributive preferences? What is the joint role of preferences and beliefs about inequality for support for redistribution? We study these questions in a staggered experiment with a broadly representative sample of the Swiss population conducted in the context of a vote on a highly redistributive policy proposal. Our sample comprises a majority of inequality averse subjects, a sizeable group of altruistic subjects, and a minority of predominantly selfish subjects. Irrespective of preference types, individuals overestimate the extent of income inequality. An information intervention successfully corrects these large misperceptions for all types, but essentially does *not* affect aggregate support for redistribution. These results hide, however, important heterogeneity because the effects of beliefs about inequality for demand for redistribution are preference-dependent: only inequality averse individuals, but not the selfish and altruistic ones, significantly reduce their support for redistribution. These findings cast a new light on the seemingly puzzling result that, in the aggregate, large changes in beliefs about inequality often do not translate into changes in demand for redistribution.

## 1. Introduction

Over the last decades, many countries have experienced a substantial increase in income and wealth inequality. This increased concentration of income and wealth has been particularly pronounced at the very top. In the US, for example, the share of income captured by the top 1% nearly doubled over the last four decades, from 10.4 percent in 1980 to 19.1 percent in 2020 (Alvaredo et al., 2013; World Inequality Database, 2023). This evolution is not limited to the US: over the last 40 years, the share of income captured by the top 1% increased by more than 20 percent in Germany and France, respectively, and by more than 25 percent in Switzerland (Foellmi and Martínez, 2017; Swiss Federal Department of Finance, 2022).

These increasing inequalities have put redistribution again on top of the political agenda and have given rise to civil movements such as the Wallstreet movement (“We are the 99%”). Likewise, in Switzerland several strongly redistributive initiatives have been put to vote in recent years. The latest example of such a vote took place in September 2021, when Switzerland voted on the so-called “99% initiative”—a policy

proposal aimed at increasing taxes on the capital gains of those at the very top of the income distribution.

But what motivates citizens to support such strongly redistributive proposals? One strand of research has highlighted the role of (biased) beliefs about inequality for support for redistribution (Cruces et al., 2013; Kuziemko et al., 2015; Karadja et al., 2017). Another strand of the literature has underscored the importance of (other-regarding) preferences (Fehr et al., Forthcoming; Fisman et al., 2017; Kerschbamer and Müller, 2020). These two strands of the literature have largely evolved separately. In this paper, we investigate the *joint* role of beliefs about inequality and other-regarding preferences for demand for redistribution. We are particularly interested in understanding whether preferences alter the effects of beliefs about income inequality for support for redistribution, i.e. whether the effects of beliefs about inequality for demand for redistribution are preference-dependent.

Recent field and laboratory evidence indicates that there are important heterogeneities in terms of other-regarding preferences, i.e., not all individuals put the same weight on the distribution of payoffs (see, e.g., Fisman et al., 2017; Kerschbamer and Müller, 2020; Fehr et al.,

<sup>☆</sup> The “” symbol indicates that the authors’ names are in certified random order. We thank Ricardo Perez-Truglia, co-editor, as well as three anonymous referees for their very constructive comments and suggestions during the review process. Thomas Epper acknowledges the support from the Métropole Européenne de Lille (MEL). Ernst Fehr acknowledges the support from the Research Priority Program of the University of Zurich on Equality of Opportunity.

<sup>\*</sup> Corresponding author.

E-mail addresses: [henkel@kof.ethz.ch](mailto:henkel@kof.ethz.ch) (A. Henkel), [ernst.fehr@econ.uzh.ch](mailto:ernst.fehr@econ.uzh.ch) (E. Fehr), [julien.senn@econ.uzh.ch](mailto:julien.senn@econ.uzh.ch) (J. Senn), [thomas.epper@cnrs.fr](mailto:thomas.epper@cnrs.fr) (T. Epper).

**Forthcoming**). Yet, little is known about the extent and the ways in which these preferences interact with beliefs about inequality. This raises several interesting new questions. First, do beliefs about inequality depend on preferences? While previous studies have established that individuals tend to have biased beliefs about inequality (see e.g. Cruces et al., 2013), there may be large heterogeneities in the population. For example, other-regarding individuals might have different beliefs about inequality than selfish individuals. In particular, other-regarding individuals may believe that there is more inequality than there actually is, and selfish individuals may believe that there is less inequality than there actually is. Much like motivated beliefs, such preference-dependent beliefs might be self-sustaining: it might be easier to remain selfish by convincing oneself that there is little inequality, and it might be easier to remain other-regarding by convincing oneself that there is a lot of inequality. Second, can we legitimately expect that beliefs about inequality will shape demand for redistribution to a similar extent across individuals with different preferences? For example, should we expect that beliefs about inequality will affect the demand for redistribution of other-regarding individuals and of selfish individuals in the same way? Would providing individuals with different preference types with information about inequality affect their demand for redistribution differently? For example, would the response to information about inequality of a selfish subject be smaller than, say, the response of an inequality averse individual?

We study these questions using a pre-registered online experiment with a sample that is broadly representative of the Swiss population (with respect to age, gender and geographical area) in the context of the vote on the 99% initiative, a highly redistributive proposal that aimed at increasing taxes for the top 1% by instituting a capital gains' tax. Our experiment consists of two waves, conducted six months apart. In the first wave, we elicit participants' other-regarding preferences using an incentivized money allocation task in which participants have to make distributional choices between themselves and another anonymous participant. In the second wave, we measure participants' beliefs about the income share received by the top 1% of income earners in Switzerland.<sup>1</sup> To assess the causal effects of beliefs about inequality for demand for redistribution, we randomly assign half of the participants to a treatment condition that provides them with credible *factual* information about the share of income received by the top 1% of income-earners in Switzerland. We then elicit all respondents' support for the 99% initiative by allowing them to make a real monetary donation to organizations that either actively campaign in favor of or against this referendum.

This staggered design, which purposefully decouples the elicitation of distributive preferences from the information intervention and the measurement of support for redistribution, allows us to study the causal effects of beliefs about inequality for support for the 99% initiative, and to investigate the extent to which these effects are preference-dependent. In addition, it also allows us to explore whether individuals with different preference types hold fundamentally different beliefs about the extent of inequality, and whether they update these beliefs differently when presented with credible and objective information.

Following the approach discussed in Fehr et al. (Forthcoming), we characterize preference heterogeneity in our sample by applying the Dirichlet Process means (DP-means) algorithm, a Bayesian nonparametric clustering algorithm that allows us to infer the prevailing social preference types in the population using the subjects' overall behavior in the money allocation task, and that endogenously assigns each individual to a preference type. This approach has several advantages. In particular, it enables the identification of preference types without committing to a pre-specified number of different preference types,

and it does not require an ex-ante specification or parameterization of types. It also does not presume a specific error structure. In other words, the algorithm enables the identification of behavioral clusters without assumptions on the number of existing preference clusters and the behavioral properties (e.g., the utility functions) of the different clusters, thereby enabling a flexible and data-driven exploration of heterogeneity.

Consistent with the findings of Fehr et al. (Forthcoming), we also document three distinct types with a clear behavioral interpretation: A large group of predominantly *inequality averse* individuals, a smaller group of *altruistic* individuals, and a minority of predominantly *selfish* individuals. Our sample is therefore very diverse in terms of how respondents weigh other people's payoffs: some individuals are highly sensitive to the distributional consequences of their choices, while others are predominantly self-interested.

We expect important heterogeneities in subjects' response to our information intervention depending on subject's preference type. Indeed, previous work has provided both theoretical and empirical evidence that other-regarding preferences are an important predictor of political support for redistribution (Fehr et al., Forthcoming; Kerschbamer and Müller, 2020). In particular, Fehr et al. (Forthcoming) show that while inequality averse individuals tend to support policies that primarily aim at reducing the incomes of the rich, altruistic (and selfish) individuals are considerably less supportive of such policies. For this reason, we expect support for the 99% initiative to be particularly strong among the inequality averse individuals. This is due to the fact that the 99% initiative was largely framed as a policy that is predominantly concerned with taxing the very rich. We also expect information about inequality to have a particularly large effect on these individuals. In addition, we speculate that this effect might be larger at higher incomes, as Fehr et al. (Forthcoming) have provided evidence that social preferences might play a particularly important role for more affluent individuals.<sup>2</sup> Specifically, we speculate that the effect of beliefs about inequality will be particularly pronounced for inequality averse individuals with an above-median income, relative to selfish individuals with an above-median income.<sup>3</sup>

We document several novel findings on how beliefs about inequality and other-regarding preferences jointly affect support for redistribution. First, we investigate how individuals *perceive* inequality, focusing on the inequality between the top 1% of the income earners and the rest of the population. We find that our subjects substantially overestimate the extent of income inequality, regardless of their preference type. While the top 1% of income earners actually receive 12% of the total annual income, the average respondent thinks that the top 1% receives about five times more (average belief: 54.2%). These misperceptions are widespread, with 94% of our sample overestimating inequality, and only 4% underestimating it.


Second, we explore whether the extent of these perceptions differs by preference type. We find that these misperceptions are largely orthogonal to respondents' preference type. That is, inequality averse individuals do not have more biased beliefs about inequality than the selfish or the altruistic individuals.

Third, we show that social preferences are an important predictor of support for the 99% initiative — consistent with past evidence that social preferences elicited experimentally can provide valuable insights into individuals' support for real world policies (see, e.g., Fehr et al., Forthcoming; Fisman et al., 2017; Kerschbamer and Müller, 2020). In our context, we find that — compared to selfish subjects — the largest support for the 99% initiative comes from the inequality averse respondents. Altruists also show significantly more support for the 99%


<sup>1</sup> In the following, we use “beliefs about (income) inequality” and “perceived (income) inequality” as synonyms for “beliefs about the income share received by the top 1% of the income earners”.

<sup>2</sup> We refer to subjects with an income above the median as affluent individuals.

<sup>3</sup> We pre-registered this conjecture as our main hypothesis.

initiative than selfish subjects, but the magnitude of the effect is somewhat lower than for the inequality averse (although not significantly different). Furthermore, consistent with previous findings (Fehr  *et al.*, Forthcoming), we find that the association between other-regarding preferences and support for the 99% initiative is rather small and insignificant among individuals with an income below the median, but large in magnitude and strongly significant among those with an income above the median.

Fourth, we show that providing respondents with accurate information about the extent of income inequality dramatically reduces these misperceptions, and that all social preferences types update their beliefs to a similar extent. While this informational shock successfully corrects subjects' large misperceptions, it essentially does *not* affect average support for redistribution, consistent with a recent meta-analysis on the causal effects of beliefs about inequality for demand for redistribution (Ciani *et al.*, 2021). This average result hides, however, important heterogeneity. Indeed, the downwards shocks in beliefs about inequality generates a large and significant *decrease* in donations in favor of the 99% initiative among inequality averse respondents, and it essentially does *not* affect the donations of the remaining subjects. This effect appears to be mainly driven by higher incomes individuals, consistent with what we hypothesized. We discuss multiple robustness checks that rule out alternative interpretations to these findings. In particular, we argue that the type-specific nature of our hypotheses, and our results, rule out that our findings are driven by demand effects — a concern against some information provision experiments (Haaland *et al.*, 2023). If our information intervention had generated large demand effects one would have expected that all preference types show a reduction in the support for redistribution, regardless of income levels. In contrast, however, we find that the effect is mainly driven by the affluent inequality averse individuals — as hypothesized.

Our paper connects to several strands of the literature. First, our paper relates to the large literature on the political economy of demand for redistribution. This literature has identified a list of determinants of support for redistribution, such as beliefs about income mobility (Piketty, 1995; Benabou and Ok, 2001; Benabou and Tirole, 2006; Alesina *et al.*, 2018), beliefs about the causes of success (Fong, 2001; Alesina and Angeletos, 2005; Almås *et al.*, 2020), and beliefs about income and wealth inequality (Cruces *et al.*, 2013; Karadja *et al.*, 2017; Fehr *et al.*, 2022), among others. More recently, several papers have highlighted the role of other-regarding preferences, such as the equality-efficiency tradeoffs (Fisman *et al.*, 2017) or inequality aversion and altruism (Fehr  *et al.*, Forthcoming), for the demand for redistribution. We contribute to this literature by investigating the joint role of preferences and beliefs about inequality for the demand for redistribution. We are also the first, to our knowledge, to measure support for redistribution using real monetary donations in the context of a real, upcoming political campaign.

Our paper also connects to the literature that uses information interventions to study subjective beliefs about the economy.<sup>4</sup> For example, recent papers have used information interventions to study topics as diverse as the role of beliefs about inequality or about the distribution of income for demand for redistribution (Cruces *et al.*, 2013; Karadja *et al.*, 2017; Fehr *et al.*, 2022; Kuziemko *et al.*, 2015), the link between labor market concerns and support for immigration (Haaland and Roth, 2020), the effects of beliefs about public debt for the demand for government spending and taxation (Roth *et al.*, 2021), the relationship between perceptions of existing spending levels on support for increased government spending on education (Lergetporer *et al.*, 2018), and the role of beliefs about the size of the gender wage gap for demand for policies intended to reduce it (Settele, 2022). We contribute to this literature by providing novel evidence on the misperceptions of income

inequality, and by showing that these misperceptions are widely held among individuals characterized by very different preference types.

While most of these survey experiments mainly focused on analyzing the average effects of a particular information intervention for demand for redistribution, several of them also investigated how political orientation modulates this effect. Overall, the evidence on whether political orientation affects subjects' response to information is mixed. Some papers find that it is mainly left-wing respondents who adjust their demand for redistribution after a shock in their beliefs about inequality (Alesina *et al.*, 2018; Fehr *et al.*, 2022; Settele, 2022) while others find that it is mainly right-wing respondents who adjust (Karadja *et al.*, 2017; Fenton, 2020). There are also some papers that find no evidence of heterogeneous effects (Lergetporer *et al.*, 2020; Haaland and Roth, 2023) whatsoever. In this paper, we take a different route and show that correcting misperceptions about inequality predominantly affects support for redistribution of inequality averse individuals, especially the most affluent ones. These findings cast a new light on the seemingly puzzling result that, in the aggregate, large changes in beliefs about inequality often do not translate into changes in demand for redistribution, as documented in a recent review of the literature (Ciani *et al.*, 2021).

## 2. Experimental design

### 2.1. Overview

Our main aim is to study how beliefs about inequality and distributive preferences jointly predict demand for redistribution. In particular, we are interested in understanding whether individuals with different preference types hold different beliefs about the extent of inequality and whether the causal effect of beliefs on support for redistribution differs across preference types. This endeavour requires the following elements: (i) a clean and independent measure of social preferences, (ii) a baseline measure of beliefs about inequality, (iii) an exogenous shock to beliefs about inequality, and (iv) an incentivized measure of demand for redistribution. We approach this task by conducting a staggered experiment with two waves.<sup>5</sup> The key features of our experiment are summarized in Fig. 1.

The main goal of the first wave is to measure the distributive preferences of Swiss voters. We also use this wave to collect information on respondents' socio-demographics and on a set of beliefs that have been shown to matter for support for redistribution, such as beliefs about the determinants of individual success, beliefs about intergenerational income mobility, prior (i.e., pre-intervention) beliefs about poverty, and beliefs about the distributions of income and wealth. In addition, we measure other economic preferences using the preference survey module by Falk *et al.* (2022). For details on the measurement of these additional variables, see Appendix B.1.

We conduct the second wave six months later. In this second wave, we first measure respondents' prior beliefs about top income inequality in Switzerland (i.e., their beliefs about the share of total income received by the top 1% of income earners). We then exogenously shock these beliefs for half of the respondents by providing them with credible and objective information about the extent of top income inequality in Switzerland. Subsequently, we measure demand for redistribution by allowing all subjects to make a real monetary donation to civic

<sup>4</sup> For a recent review of the literature of information interventions, see Haaland *et al.* (2023).

<sup>5</sup> The advantage of separating the elicitation of distributional preferences (first wave) from the belief elicitation (second wave) is that it mitigates the risk of spillovers between the two tasks. This is important as such spillovers might introduce confounds. For example, one could imagine that subjects' stated beliefs about the degree of inequality could make inequality more salient and thereby affect their decisions in the money allocation task used to identify preferences.

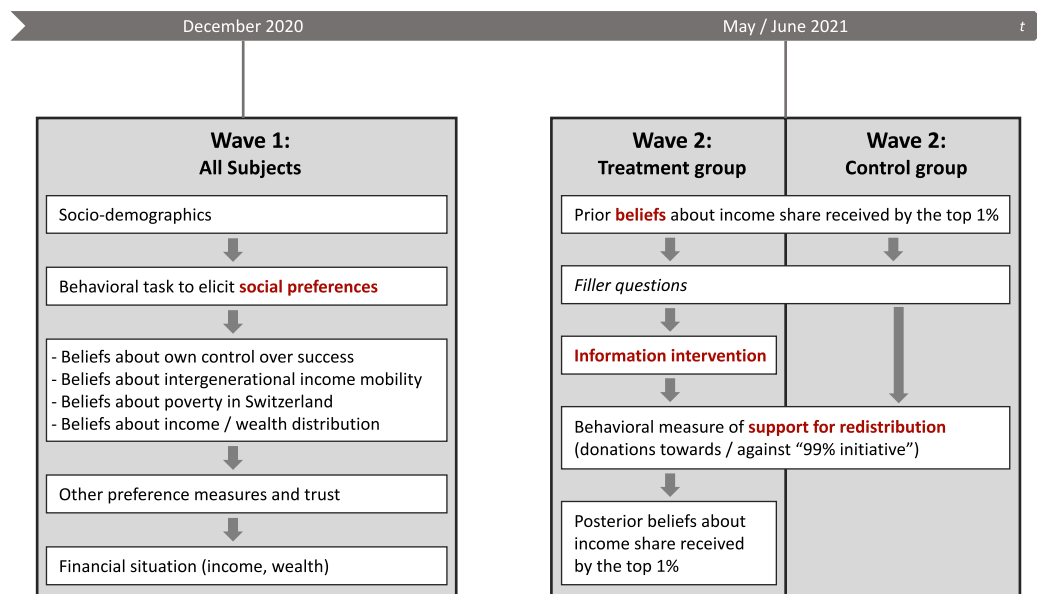


Fig. 1. Overview of our staggered experiment with two waves.

groups that either support or oppose the 99% initiative. Last, we re-elicited subjects' beliefs about income inequality in order to assess whether they updated their beliefs.

We provide details on the two waves and on our subject pool in the next subsections.

## 2.2. Wave 1: Eliciting social preferences

The first wave of our study, which we conducted in December 2020, is aimed at uncovering the distribution of other-regarding preferences among Swiss voters. We approach this task by eliciting respondents' distributional preferences in a set of 12 incentivized money allocation tasks.<sup>6</sup> In each of these tasks, the participants have to decide how to allocate experimental currency units (ECUs) between themselves and an anonymous other participant of the study.

Fig. 2(a) depicts the various budget lines for which subjects had to make a decision. In some decision situations, the decision maker can give up some of her own payoff to *increase* the payoff of the other. In other decision situations, they can pay to *decrease* the payoff of the other. These different choice situations systematically vary the cost and the joint payoff consequences of redistribution; thereby allowing us to identify a wide range of other-regarding behaviors.

Fig. 2(b) illustrates how a typical choice situation was presented to participants. We represented the available choices numerically and graphically in order to make the trade-offs and the associated payoff implications transparent. There were always seven interpersonal allocations (labeled by 1 to 7) available per choice situation, and all of them were located on a budget line. Each available allocation consisted of a specific distribution of ECUs between the participant (bars labeled by "You receive") and the other person (bars labeled by "other person receives"). In this example, the slope of the budget line is -2, indicating that for every ECU the decision maker gives up, the other player receives 2 ECUs. Perfect equality in payoffs can be achieved by choosing allocation 4.

<sup>6</sup> The task used for the clustering and type identification is based on Fehr et al. (Forthcoming, 2023).

## 2.3. Wave 2: Beliefs and information provision experiment

### 2.3.1. Measuring prior beliefs about income inequality

Six months after the first wave, we contacted the *same* subjects to participate in the second part of the study. This wave starts with questions aimed at measuring respondents' prior beliefs about income inequality. More specifically, we elicit subjects' beliefs about the share (in percent) of the total national income that is received by the top 1% of income-earners in Switzerland. We elicit these beliefs using a two-step approach: We first ask subjects to think about the total income that is received by all the people in Switzerland, and to indicate how much they believe is received by the top 1% of people with the highest incomes by selecting one out of ten possible brackets (e.g., 0%–10%, 11%–20%, 21%–30%, ..., 91%–100%). We then ask respondents to refine their answer by providing a point estimate within the range they chose.<sup>7</sup>

We purposely chose to elicit respondents' beliefs about the *income* share received by the top 1% (as opposed to, for example, the wealth share of the top 1%, or another measure of inequality) because it more closely corresponds to the main purpose of the 99% initiative, which is to increase income taxes for the top 1% of the income earners (and is largely framed as such). We did not incentivize these beliefs because we were not interested in eliciting correct beliefs but in getting an estimate of their *subjective* perception of the extent of income inequality, i.e., what comes to their mind when they think about the extent of inequality.

### 2.3.2. Providing respondents with objective information on income inequality

We implement our information intervention after the elicitation of subjects' prior beliefs about the income inequality and several filler questions. We provide credible and objective information about the top income inequality in Switzerland to a random selection of subjects. More specifically, we randomly assign subjects to one of two groups: (i) a treatment group in which subjects are provided with factual information about the degree of income inequality, or (ii) a pure control group in which subjects receive no information.

<sup>7</sup> For example, if a subject first answered that the top 1% receive between 21% and 30% of the total income, then they had to provide a precise estimate within this interval. We provide the exact wording in Appendix B.2.

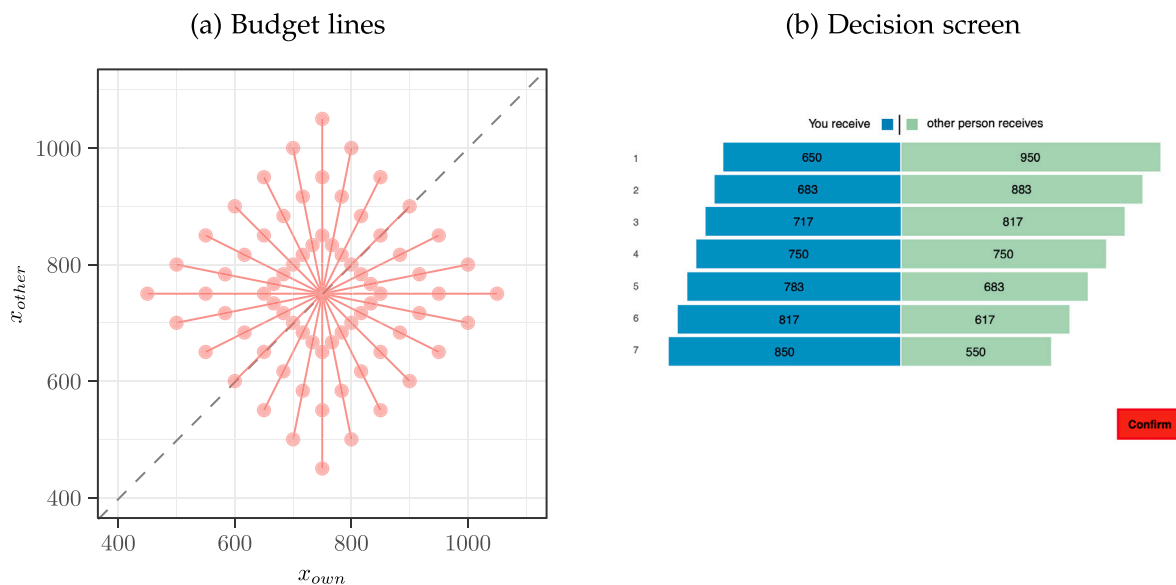


Fig. 2. Measuring other-regarding preferences with a money allocation task.

Note: Fig. 2(a) depicts the twelve decision situations (budget lines) for which subjects had to make a decision. For each budget line (i.e., for each decision situation), subjects were asked to choose an interpersonal allocation of ECUs between themselves ( $x_{own}$ ) and the other person ( $x_{other}$ ). Each “budget line” crossed the 45-degree line and comprised seven different (discrete) interpersonal allocations (represented by dots in Fig. 2(a)). Fig. 2(b) illustrates how a typical choice situation was presented to participants.

Subjects in the treatment group are told the share of the total income that is received by the top 1%. We convey this information, and contrast it with their prior beliefs, using the following sentence:

*“You told us that you believe that the 1% of people with the highest incomes in Switzerland receive [xx]% of the total annual income. According to the objective data collected by the Federal Department of Finance, the top 1% actually receive 12% of the total annual income.”*

We explicitly mention our data source (the Federal Department of Finance) in order to increase the chance that our participants believe the information we provide them with.<sup>8</sup> To further illustrate the discrepancy (if any) between respondents’ misperceptions and reality, we also provide them with a graphical representation that contrasts these two figures. For details on the information intervention, see Appendix B.3.

### 2.3.3. Measuring demand for redistribution in the context of the 99% initiative

After the information intervention, we measure all respondents’ support for the 99% initiative. We start by describing the content of this initiative using a wording that is close to the wording used in the official voting booklets sent to all the Swiss voters:

*“In September 2021, Switzerland will vote on the 99% initiative. The initiative aims at increasing taxation of the richest 1%. The resulting tax revenue shall be used to reduce the taxes for low and middle labor incomes or increasing social transfers. The initiative wants to reach this goal by taxing capital incomes (i.e., incomes that result from capital ownership like, for example, dividends or interest incomes from stocks and bonds) beyond a threshold at a rate that is 50% higher than labor incomes of the same amount.”*

<sup>8</sup> We measured participants’ trust in the Federal Department of Finance at the end of wave 2 in order to verify that they consider the information trustworthy. Respondents could indicate how trustworthy they find the Federal Department of Finance on a scale from 1 to 7, where 1 means “not trustworthy at all” and 7 means “very trustworthy”. Overall, participants find the Swiss Federal Department of Finance very trustworthy (Mean = 5.53; Standard Deviation = 1.20).

We then measure respondents’ support for, or opposition to, this initiative by allowing them to make a *real monetary donation* to civic organizations that either actively support or actively oppose the 99% initiative.<sup>9</sup> To that end, we endow them with 20 Swiss Francs (CHF). We first ask them whether they lean towards donating money to a civic group that supports the 99% initiative, or whether they lean towards donating to a civic group that opposes the 99% initiative. We then ask them to make a donation of up to CHF 20 to the organizations they lean towards. For example, if a subjects reveals that they lean towards donating to a civic group that opposes the 99% initiative, they have then to decide in a second stage how many — out of the CHF 20 — they want to donate to such organizations. These donations are incentive compatible because the donations are effectively made to such civic groups, and because the money the subjects did not donate can be kept by themselves.<sup>10</sup>

### 2.3.4. Exit questionnaire and posterior beliefs

At the end of the second wave, we measure treated subjects’ posterior beliefs about income inequality. We asked them to think about the total income that is received by all the people in Switzerland, and to indicate what they think is the share (in percent) of the total income that is received by the top 1% of people with the highest incomes. These posterior beliefs allow us to evaluate whether treated subjects update their beliefs in line with the information we provide them with.<sup>11</sup>

<sup>9</sup> We purposefully did not disclose the identity of these civic groups in order to avoid that our subjects condition their donations on their subjective beliefs about the different organizations. However, we provided them with examples of such civic groups.

<sup>10</sup> We decided to measure support for the 99% initiative using this two-step approach for two reasons. First, by giving subjects the possibility to make donations to organizations that *oppose* the 99% initiative, we give subjects more leeway to express their opposition to such a policy proposal, compared to just giving them the choice of only donating 0 francs to a group that supports it. Second, this two-step procedure increases the time spent thinking about the decision and thereby likely reduces the amount of random answers.

<sup>11</sup> Measuring both priors and posteriors allows us to precisely pin down the extent to which participants learn from the information we provide them, and to distinguish the effects of changes in beliefs from mere priming effects (Haaland et al., 2023). While one might be concerned that measuring

Finally, we conclude the survey with additional questions on subjects' socio-demographics and their personality traits.

#### 2.4. Data collection, sample, and experimental protocol

Both waves were conducted online with a broadly representative sample of the German- and the French-speaking population of Switzerland and were collected in collaboration with the LINK Institute.<sup>12</sup> In both waves, respondents were paid a show-up fee for their participation, provided that they completed the survey until the end. In the first wave, we also incentivized respondents' choices in the money allocation task by paying out each subject on the basis of their decision in one randomly chosen decision situation. In the second wave, respondents could earn whatever fraction of the CHF 20 they decided to keep instead of donating it to civic groups campaigning in favor or against the 99% initiative.<sup>13</sup>

Our final sample comprises the 1031 subjects who participated in both waves. Descriptive statistics on participants' main socio-demographic characteristics can be found in Table B.1 in Appendix B.4. Overall, our sample is broadly representative of the Swiss voting population in the German and the French language areas with respect to age, gender, geographical area, and income. The average respondent in our sample is 47.6 years old, the share of men is 52.4%, the share of French-speaking respondents is 24.6%, and the median income is CHF 6000. Table B.1 also shows that the control and the treatment groups are well balanced across the main observable characteristics.

An important concern in studies comprising multiple waves is selective attrition. In Table B.2 in Appendix B.5, we show that attrition between the two waves is orthogonal to the treatment assignment, to the social preference type, and to the main observable characteristics, i.e., we have very little reasons to worry about selective attrition in our sample.<sup>14</sup>

Another potential concern with online studies is that subjects do not pay attention to the questions they are asked and simply rush quickly through the survey. In order to measure respondents' attentiveness and to proxy data quality, we added two attention checks to the second wave. These attention checks are aimed at measuring whether participants read survey items carefully before answering them (Berinsky et al., 2014). Data quality is remarkably high in our sample: 79.2% of the subjects correctly answered both attention checks, and only 9.7% failed to pass both checks.<sup>15</sup>

both priors and posteriors about the same object might generate stronger demand effects, Roth and Wohlfart (2020) find no evidence that the elicitation of prior beliefs affects learning. In addition, subjects' desire for consistency might even mute the effect of information (Falk and Zimmermann, 2013; Haaland et al., 2023), and recent methodological contributions suggest that even strong experimentally-induced demand effects tend to result in only very modest behavioral responses (de Quidt et al., 2018). To further mitigate such concerns, we added a number of "filler" questions between the information intervention and the elicitation of posterior beliefs. For details on the exact wording of the prior and posterior belief elicitation, see Appendix B.2.

<sup>12</sup> For logistical reasons, we did not conduct the experiment in the Italian-speaking part of Switzerland, which comprises approx. 8% of the Swiss population.

<sup>13</sup> Median time to complete wave 1 was 39 min, for which respondents were paid an average of CHF 35.5 (including a show-up fee of CHF 15; the exchange rate between points in the money allocation task and Swiss Francs was 40 points per CHF 1). Median time to complete wave 2 was 24 min, for which respondents were paid an average of CHF 19.7 (including a show-up fee of CHF 10).

<sup>14</sup> In total, 1383 subjects participated in wave 1 in December 2020, and 1031 subjects participated in wave 2 in May/June 2021, i.e., the rate of attrition is 25.5%.

<sup>15</sup> Some online samples contain considerable proportions of respondents who do not correctly answer attention check questions even in shorter surveys, with fail rates between a third and a half of the sample (see, e.g., Berinsky et al., 2014). Thus, our pass rates can be considered very high.

We pre-registered our experimental design, the main hypotheses, the main outcome variables, and the sample sizes before conducting the second wave of the study.<sup>16</sup> Ethics approval was obtained from the Human Subjects Committee of the Department of Economics of the University of Zurich (OEC IRB #2021-032).

### 3. The empirical distribution of social preferences

Before turning to the beliefs about inequality and their effect for demand for redistribution, we investigate the empirical distribution of social preferences in our sample. We follow Fehr et al. (Forthcoming, 2023) and characterize preference heterogeneity in our sample using the Dirichlet Process (DP-) means (Kulis and Jordan, 2012), a Bayesian nonparametric clustering algorithm that allows to cluster individuals into groups on the basis of their behavioral similarities.

The DP-means algorithm groups individuals into clusters according to their *behavioral similarities*. In our context, clusters are based on the 12 distributional choices made by subjects in the money allocation task, and similarity is measured by "how close" an individual's allocation profile is to the average allocation of a cluster. Ultimately, individuals' are assigned to the cluster whose centroid — i.e., the vector of mean allocations across the 12 distributional choices — is closest to their own allocation profile in the 12-dimensional space of interest.<sup>17</sup>

An important aspect of the DP-means approach is that it enables the identification of preference types without committing to a prespecified number of different preference types. Moreover, this approach does neither require an ex-ante specification or parameterization of types, nor does it presume a specific error structure. This means that it remains ex-ante agnostic about key distributional assumptions, and it does not constrain heterogeneity to lie within a predetermined set of models or parameter space.<sup>18</sup> In addition, the DP-means algorithm allows for all possible type partitions of the data spanning from a representative agent up to as many types as there are individuals in the population. In this way, it endogenously determines (i) the actual number of types, (ii) the assignment of each individual to one of the types, and (iii) the behavioral (preference) properties of each type.<sup>19</sup>

This procedure reveals the existence of three fundamentally distinct preference types. We depict their distribution in the Table 1. The largest group (46.5% of the sample) comprises subjects who make predominantly payoff-equalizing choices. These subjects show both a willingness to pay to *increase* the payoff of others who are worse off, and a willingness to pay to *decrease* the payoff of others who are better off, consistent with models of inequality aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). We therefore assign this cluster the label "*inequality averse*". The second largest group (38.1% of the sample) comprises individuals who display a strikingly different form of other-regarding behavior: They are also willing to pay in order to increase the payoff of those worse off, but they are generally *not* willing to pay to reduce the payoff of those who are better off, i.e., they display a high degree of altruism towards the poor but are *not* willing to reduce the income of those who are better off. This form of other-regarding

<sup>16</sup> The study is pre-registered on the AEA RCT registry (<https://www.socialscienceregistry.org/trials/7716>).

<sup>17</sup> We briefly summarize the key features of this approach in Appendix C.1. For an extensive discussion of this procedure and some applications, see Fehr et al. (2023).

<sup>18</sup> In this regard, our approach differs from previous work (e.g. Bellemare et al., 2008; Fisman et al., 2015, 2017; Bruhin et al., 2018) that characterized preference heterogeneity on the basis of structural assumptions on preferences and error terms.

<sup>19</sup> The fact that the number of types adapts to the data has important benefits (see Kulis and Jordan, 2012). Most notably, as previous work has shown (see Comiter et al., 2016), this feature of the algorithm yields higher quality type-separation than methods that specify the number of types prior to clustering (such as *k*-means).

**Table 1**  
Distribution of behavioral types.

Cluster	Shares
Cluster 1 (Inequality Averse)	46.5%
Cluster 2 (Altruistic)	38.1%
Cluster 3 (Selfish)	15.4%
Total	100.0%

*Note:* The table displays the distribution of individuals to the three clusters (in percent) that emerge in our dataset. The behavioral interpretation of the clusters (indicated in parenthesis in the left column) is based on the interpretation of each cluster's typical behavior provided in Figure C.1.

behavior is consistent with an altruistic concern for the worse off (Charness and Rabin, 2002) and with altruistic other-regarding behavior that incorporates an equity-efficiency tradeoff (Fisman et al., 2007, 2015). We therefore label this behavioral cluster the “altruistic” cluster. The last group comprises the remaining 15.4% of the individuals that make predominantly own-payoff maximizing choices and can therefore be labeled as being “predominantly selfish”.

We discuss the characteristic behavior of these qualitatively different behavioral types in Appendix C.2. For an extensive discussion of the identification, the characterization and the validation of these behavioral types, see Fehr et al. (Forthcoming, 2023). We also explore the link between individual characteristics and social preferences in Appendix C.3.

#### 4. Hypotheses

Our experimental design allows us to shed light on the effects of beliefs about inequality and other-regarding preferences on demand for redistribution. It also allows us to assess whether the effects of beliefs about inequality are preference-dependent.

In previous work, Fehr et al. (Forthcoming) have provided theoretical arguments and empirical evidence that social preferences affect individuals' support for redistribution.<sup>20</sup> Building on these results, we hypothesize that social preferences elicited in our money allocation task can help understand support for the 99% initiative.

**Hypothesis 1.** *Individuals with social preferences, in particular the inequality averse, are more supportive of the 99% initiative, compared to selfish individuals.*

We predict that the effect may be particularly strong for the inequality averse because Fehr et al. (Forthcoming) also provide some empirical evidence that different social preference types might differ in their support for particular policy proposals. In particular, they show that while inequality averse individuals tend to support policies that primarily aim at reducing the incomes of the rich, altruistic individuals are considerably less supportive of such policies. For this reason, we expect support for the 99% initiative — which was largely framed as a policy that is predominantly concerned with taxing the very rich — to be particularly strong among the inequality averse individuals. Moreover, we speculate that this effect may be particularly large at higher incomes. This is because, at lower incomes, self-interested individuals tend to already have a high demand for redistribution, leaving little scope for social preferences to influence demand for redistribution (see also Fehr et al., Forthcoming).<sup>21</sup>

Our second hypothesis relates to the effect of the information intervention for support for the 99% initiative. Previous work has shown

that information interventions sometimes have muted average effects, but that these results may hide substantial heterogeneity. For example, some studies have suggested that subjects' response to information interventions such as, e.g., information about the respondent's position in the income distribution, may depend on their political orientation (see, e.g., Karadja et al., 2017; Fehr et al., 2022; Settele, 2022).<sup>22</sup> Relatedly, we expect that correcting people's misperceptions about the extent of inequality will primarily affect the demand for redistribution of the inequality averse respondents — who are also predicted to be the ones for which the 99% initiative is predicted to be particularly appealing (see Hypothesis 1) — and that this effect might be particularly large for the inequality averse respondents with an income above the median. In contrast, we expect that individuals who primarily care about their own payoff will remain largely insensitive to new information about the extent of inequality. More precisely, we preregistered the following hypothesis.

**Hypothesis 2.** *A (downward) shock in beliefs about income inequality will cause a larger reduction in the demand for redistribution for inequality averse individuals compared to selfish individuals — an effect that we expect to be particularly pronounced at higher incomes.*

We preregistered this conjecture as a one-sided hypothesis because preliminary data on beliefs about the distribution of income gathered in wave 1 indicated that the majority of individuals substantially overestimate the extent of income inequality. It was therefore clear at the moment of preregistration and when we designed our information intervention that the information would shocks beliefs about income inequality of most respondents *downwards* while there was no reason to expect that subjects would, on average, revise their beliefs upwards. Thus, because we are shocking beliefs about income inequality *downwards* for the vast majority of the individuals, there is no reason to expect that demand for redistribution will increase. If anything, a decrease in beliefs about inequality should either *decrease* demand for redistribution, or leave it unchanged. Because of the clear directional prediction of this conjecture, we evaluate it using one-sided tests. For all other statistical results, we use two-sided tests.

In addition, our experimental design also allows to study the extent and the ways in which beliefs and beliefs updating differ across preference types. A relatively large literature has documented that people have biased beliefs about inequality (e.g., Cruces et al., 2013; Kuziemko et al., 2015; Karadja et al., 2017). Yet, little is known about whether these biases exist for the whole population or whether they depend on individuals' preferences. In particular, we are not aware of any empirical study that systematically relates beliefs about inequality to other-regarding preferences. While beliefs might be identically biased across the whole population, there are also good reasons to think that social preferences and beliefs about inequality are correlated. In particular, it is quite plausible that other-regarding individuals believe that there is *more* inequality than there actually is, while selfish individuals believe that there is *less* inequality than there actually is. Much like motivated beliefs, these preference-dependent beliefs might be self-sustaining: it might be easier to remain selfish by convincing oneself that there is little inequality, and it might be easier to remain other-regarding by convincing oneself that there is a lot of inequality. Thus, we conjecture that beliefs may differ by preference type. In particular, we conjecture that other-regarding individuals believe that there is more inequality than selfish individuals (Hypothesis 3).

<sup>20</sup> In Appendix E, we reproduce the simple theoretical framework discussed in greater details in Fehr et al. (Forthcoming).

<sup>21</sup> We discuss the differential role of altruism and inequality aversion, and the role of income, for the political support for different types of policy proposals in greater details in Fehr et al. (Forthcoming).

<sup>22</sup> Note, however, that the evidence of heterogeneity along political orientation is not unequivocal. In Appendix A, we review the literature that has analyzed such heterogeneous responses to information interventions in the context of preferences for redistribution and show that while some papers have found evidence of heterogeneous response by political orientation, others did not find such heterogeneities.

**Hypothesis 3.** *Other-regarding individuals overestimate the extent of inequality more than selfish individuals.*

Our last hypothesis relates to belief updating across the different preference types. While prior beliefs might differ across preference types (Hypothesis 3), whether and how individuals update their beliefs depending on their type is unclear (Hypothesis 4). On the one hand, it is possible that the beliefs updating process depends on the preference type, and that individuals are unwilling to revise their beliefs upon seeing information that contradicts their priors. For example, other-regarding individuals might be less likely to revise their beliefs downwards if they learn that there is less inequality than they initially thought. In a similar vein, selfish individuals might be reluctant to update their beliefs upwards if they learn that there is more inequality than they initially thought. On the other hand, it is also possible that all individuals update their beliefs “rationally” upon being presented with credible and objective information about income inequality, which would suggest that the beliefs updating process is independent of preferences. However, although there is uncertainty regarding the preference-dependence and the extent of belief updating, it appears reasonable to conjecture that all preferences types respond at least partly to the provision of credible information about inequality.

**Hypothesis 4.** *The information intervention corrects misperceptions about income inequality for all preference types.*

## 5. Income inequality in Switzerland: facts and misperceptions

Before turning to the joint role of beliefs about inequality and other-regarding preferences for political support for redistribution (Hypothesis 1 and Hypothesis 2), we discuss the actual extent of income inequality in Switzerland. We also explore whether beliefs about inequality depend on preferences (Hypothesis 3) and the extent to which beliefs updating depends on preferences (Hypothesis 4).

### 5.1. The distribution of income inequality in Switzerland

Over the last century, the share of total income received by the top 1% of income earners has fluctuated around approximately 10% in Switzerland. While this share dropped well below 10% in the late sixties, it has increased by more than 27% between 1981 and 2010 (Foellmi and Martínez, 2017). By the end of 2018, the top 1% of income earners received 12.08% of the total income in Switzerland (Swiss Federal Department of Finance (2022)). While this evolution is broadly comparable to other European countries such as Germany or France, it is in stark contrast with the United States, which have experienced a much stronger increase in the income concentration in recent years (Alvaredo et al., 2013) with the top 1% receiving close to 20% of the total US-income in 2022 (World Inequality Database, 2023).

### 5.2. Respondents’ (mis)perceptions of income inequality

Fig. 3 depicts the distribution of subjects’ misperceptions about the income share received by the top 1%, defined as the difference between subjects’ prior belief of the income share received by the top 1% and the truth (12%).<sup>23</sup> On average, respondents largely *overestimate* the share of total income received by the top 1%: They believe that the top 1% receive 54.2% of the total income (SD: 25.6 pp), while the actual share of 12% is almost five times lower (*t*-test that the mean equals 12%,  $p < 0.001$ ). In terms of distribution, 93.9% of the sample overestimates inequality, while 3.8% underestimates it, and only a minority of 2.3% of the respondents has correct beliefs.<sup>24</sup>

These results might come as a surprise given that some studies have found that misperceptions about inequality are more balanced. For

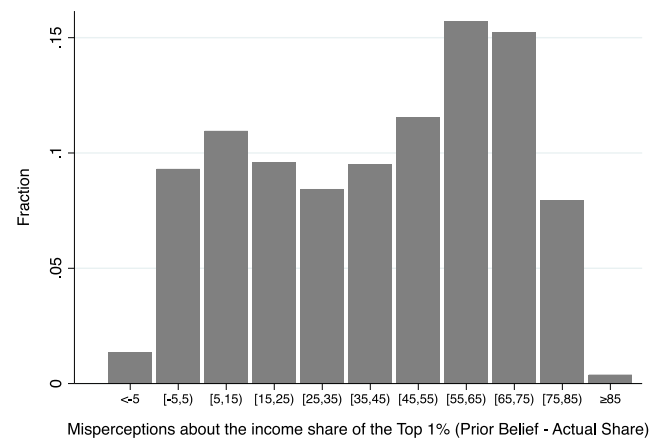


Fig. 3. Misperceptions about the income share of the top 1%.

Note: This figure shows the distribution of misperceptions about the income share of the top 1% of income earners in Switzerland. The x-axis reports misperceptions, i.e., the difference between the respondent's prior belief and the actual income share received by the top 1% (in 10 pp-width bins). The y-axis reports the percentage of subjects in each bin.

example, Cruces et al. (2013) show that misperceptions about relative income are relatively balanced in a sample of Argentinians, i.e., a substantial share of individuals also underestimates inequality. Interestingly, however, they also find that the majority of their sample (55%) underestimates how rich they are (i.e., they overestimated inequality). Overall, our study is not the only one in which subjects overestimate one dimension of income inequality. For example, Karadja et al. (2017) find that 86% of Swedes believe that they are poorer than they actually are (which can be interpreted as an overestimation of inequality). Likewise, Fehr et al. (2022) find that Germans systematically underestimate their true place in the worlds' income distribution.

Do these misperceptions depend on subjects' preferences (Hypothesis 3)? To shed light on this question, we depict the distribution of subjects' misperceptions about the income share of the top 1% as a function of respondents' preference type in Fig. 4. The figure indicates that the misperceptions are rather similar across the different preference types. On average, the inequality averse respondents overestimate the share received by the top 1% by 43.1 pp (SD: 26.5 pp), the altruistic respondents overestimate it by 41.8 pp (SD: 24.2 pp), and the selfish subjects by 40.7 pp (SD: 27.0 pp). A Kruskal–Wallis test cannot reject the null hypothesis that individuals with different preference types overestimate inequality to a similar extent ( $p = 0.478$ ). Altogether, these results suggest that our respondents generally overestimate inequality, and that these misperceptions are not preference-dependent.<sup>25</sup>

### 5.3. Correcting misperceptions about income inequality

Can the large and widespread misperceptions about the extent of income inequality documented above be corrected, and how do the different preference types update their beliefs (Hypothesis 4)? To assess the causal effects of beliefs about inequality, we provided all the subjects from the treatment group with credible and objective information about the share of total income received by the top 1% of income

<sup>24</sup> The results are qualitatively similar if we rely on the beliefs about the income earned by the top 1% measured in wave 1. While the elicitation technique differed between wave 1 and wave 2 (see Appendix B.1 and B.2 for details), the correlation between prior beliefs about the income of the top 1% elicited in wave 2 and the beliefs elicited in wave 1 is positive and highly significant ( $\rho = 0.34$ ,  $p < 0.001$ ). See Appendix D.2 for details.

<sup>25</sup> In Appendix D.1, we show that income does not predict misperceptions. We also show that it does not predict posterior beliefs in Appendix D.4.

<sup>23</sup> In Appendix D.3, we show the distribution of prior beliefs.

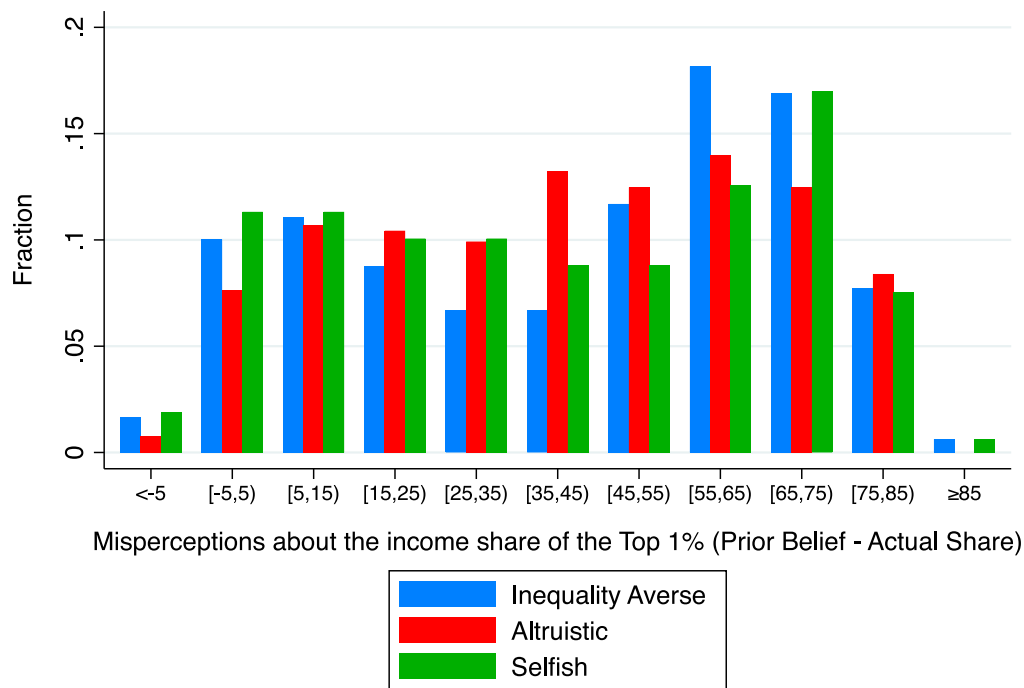


Fig. 4. Misperceptions about the income share of the top 1% by preference type.

Note: This figure shows the distribution of misperceptions about the income share of the top 1% of income earners in Switzerland (by preference type). The x-axis reports misperceptions, i.e., the difference between the respondent's prior belief and the actual income share received by the top 1% (in 10 pp-width bins). The y-axis reports the percentage of subjects in each bin.

earners (see Section 2.3.2). We then elicited their knowledge about the income distribution once more at the end of the second wave. This allows us to assess whether and how respondents updated their beliefs, and whether beliefs updating depends on preference types.

Overall, the vast majority (77.3%) of treated subjects holds correct posterior beliefs, and this holds true for all preference types (for details, see Appendix D.4). In Fig. 5, we depict the belief updating of the treated subjects, i.e., the difference between their posterior and their prior beliefs, as a function of their initial misperceptions and their preference type. The figure shows that the vast majority of subjects who initially overestimated the income share received by the top 1% shifts their beliefs downwards by the correct magnitude, while the small share of individuals who initially underestimated this share shift their beliefs upwards. Importantly, individuals from all three preference types update in the correct direction and by the right magnitude. Moreover, the Figure underscores that the three preferences types update in the right direction not only on average, but for each level of prior beliefs. This is an important result, as it shows that beliefs about inequality adjust accurately, independently of preference types and irrespective of how inaccurate prior beliefs are. While one could have expected, for example, that inequality averse subjects might “want” to believe that there is a lot of inequality and be reluctant to revise their beliefs downwards even in the face of truthful information, our results show that this is clearly not the case. Overall, these results suggest that it is unlikely that different preference types hold “motivated beliefs” that help them justify their preferences.

The fact that prior beliefs and beliefs updating are independent of preferences has an important implication for our results on the effects of beliefs about inequality on support for redistribution discussed in the next section. If we find that the information intervention affects support for redistribution for only some preference types but not for others, we can rule out that this preference-specific effect is explained by differences in prior beliefs or differences in belief updating across preference types.

## 6. The effects of beliefs about inequality and other- regarding preferences on support for the 99% initiative

In this section, we investigate the empirical role of social preferences and the causal effect of beliefs about income inequality for respondents' support for the 99% initiative. We also investigate whether the effects of changes in beliefs about inequality are preference-specific.

### 6.1. Descriptive analysis

For our analysis, we code all the donations made to organizations that oppose the 99% initiative with negative values<sup>26</sup> and we combine them with the donations made to organizations that support the 99% initiative. Our main dependent variable, which we refer to as “support for the 99% initiative”, thus ranges from -20 (if the respondent makes the largest possible donation against the 99% initiative) to +20 (if the respondent makes the largest possible donation in favor of the 99% initiative). We display the distribution of this variable in Figure D.6 in Appendix D.5.<sup>27</sup>

We hypothesized that individuals with social preferences, in particular the inequality averse, would be more supportive of the 99% initiative (Hypothesis 1), and that a shock in beliefs would predominantly reduce their demand for redistribution, compared to selfish individuals (Hypothesis 2). Moreover, we speculated that this effect might be particularly large at higher incomes. We first shed light on these questions at the descriptive level. We display the average donations for the 99% initiative as a function of respondents' preference type and their income in Fig. 6.

<sup>26</sup> For example, we recode a donation of CHF 15 against the 99% initiative with a value of -15.

<sup>27</sup> In Appendix D.5 (Figure D.7), we also plot the average donations by preference type in the control group. This allows to see how social preferences types relate to support for the 99% initiative in absence of the information intervention.

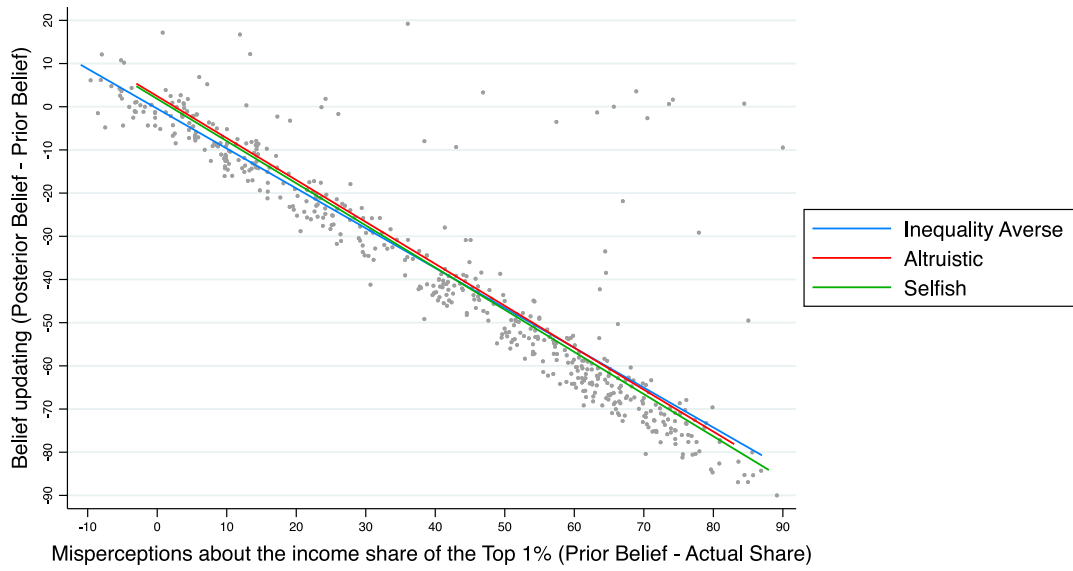


Fig. 5. Belief updating by misperceptions and preference type.

Note: This figure shows how respondents in the treatment group update their beliefs as a function of their misperceptions of income inequality and their preference type (three separate fitted lines). Belief updating (y-axis) is defined as the difference between treated subjects' posterior and their prior belief about the income share of the top 1%. The x-axis reports misperceptions, i.e., the difference between the respondent's prior belief and the actual income share received by the top 1%. Dots represent individual observations. Fitted lines are based on three separate OLS regressions where the dependent variable is belief updating and the independent variable is misperceptions.

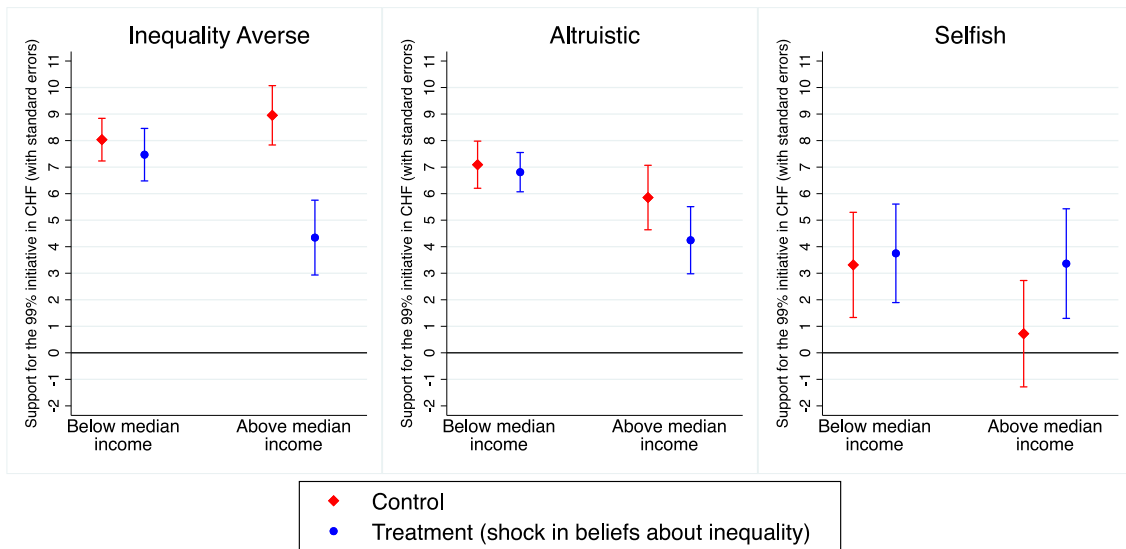


Fig. 6. The role of income and social preferences for the effect of the information shock on the donations in favor of the 99% initiative.

Note: The figure shows the average donation in favor of the 99% initiative (with standard errors). Donations to organizations that oppose the 99% initiative are coded as negative values. The control group comprises subjects who were not exposed to the information intervention. The treatment group comprises subjects who were informed about the true level of income inequality.

Several results are worth highlighting. First, other-regarding respondents tend to be a lot more supportive of the 99% initiative than selfish subjects, consistent with Hypothesis 1. While selfish subjects donate an average of CHF 2.54 in favor of the 99% initiative, those with other-regarding preferences donate more than twice as much (they donate an average of CHF 6.73, test of difference:  $p < 0.001$ ). Second, average donations decrease with income. For respondents with an income below the median, the average donation in favor of the 99% initiative is of CHF 6.89. This donation drops by more than 35%, to CHF 5.10, for respondents with an income above the median (test of difference,  $p = 0.011$ ). Finally, the effect of the information intervention appears to be the largest among inequality averse subjects (Hypothesis 2), and this effect seems mainly driven by the more affluent individuals. While inequality averse respondents with an income above the median donate

an average of CHF 8.95 in the control group, those in the treatment group donate an average of CHF 4.34 (test of difference,  $p = 0.012$ ), i.e., the affluent inequality averse subjects who revise their beliefs about inequality downwards display much lower support for the 99% initiative than those whose beliefs are not shocked downwards.

## 6.2. Regression analysis

To shed further light on the causal effect of beliefs about inequality and their interaction with other-regarding preferences, we estimate the following model

$$\text{Support}_i = \beta_0 + \beta_1 \text{Treatment}_i + \beta_2 \text{IA}_i + \beta_3 \text{Altruistic}_i + \beta_4 \text{Treatment}_i \times \text{IA}_i + \beta_5 \text{Treatment}_i \times \text{Altruistic}_i + \Gamma' X_i + \varepsilon_i \quad (1)$$

where  $\text{Support}_i$  is our measure of support for the 99% initiative based on subjects' donations.  $\text{Treatment}_i$  is an indicator variable that takes the value one if the respondent is in the information treatment.  $\text{IA}_i$  is a dummy that takes the value one if the respondent is inequality averse, and  $\text{Altruistic}_i$  is a dummy that takes the value one if the respondent is altruistic. The two interaction variables,  $\text{Treatment}_i \times \text{IA}_i$  and  $\text{Treatment}_i \times \text{Altruistic}_i$ , are aimed at capturing the possible interactions between the treatment and the two other-regarding types. The omitted category in these regressions are the predominantly selfish individuals assigned to the control group. For some of our regressions, we also include a rich set of individual-level controls,  $X_i$ , which comprise respondents' socio-demographics, proxies for their economic preferences and general trust (Falk et al., 2022), proxies for their financial situation, as well as their prior beliefs about the income share of the top 1%, the determinants of success, the income and wealth distributions in Switzerland, financial mobility, poverty in Switzerland, and distrust in politicians. Finally,  $\epsilon_i$  is an individual-specific error term.

We report the results of our estimates in Table 2.<sup>28</sup> Columns 1 and 2 show the average effect of the information intervention on the full sample. On average, a downward shock in beliefs about inequality causes a small but insignificant reduction in donations in favor of the 99% initiative ( $p = 0.106$ ). This result is consistent with a recent meta-analysis showing that while presenting subjects with information about inequality generally yields large changes in beliefs and successfully corrects misperceptions, it very often does not substantially affect demand for redistribution (Ciani et al., 2021). This result hides, however, substantial heterogeneity — as we will discuss below.

Columns 3 and 4 show the results of our main regression on the whole sample. This estimation reveals two important results. First, it shows that individuals with social preferences are significantly more supportive of the 99% initiative than selfish subjects, consistent with Hypothesis 1. This effect is particularly large for inequality averse respondents, who donate an average of CHF 6.87 more than selfish subjects (+0.62 of a standard deviation,  $p < 0.001$ ). Altruistic subjects also donate more than the selfish (CHF +5.00, +0.45 of a standard deviation,  $p = 0.001$ ), but the magnitude of this effect is lower than for the inequality averse ( $p = 0.054$ ). Column 4 shows that these effects survive the inclusion of a large set of control variables.

Second, it shows that the effects of the information shock are heterogeneous. We predicted that the information intervention would predominantly reduce the demand for redistribution of inequality averse individuals compared to selfish individuals (Hypothesis 2). Columns 3 and 4 provide evidence of such a heterogeneous response to the information intervention by preference type. For example, column 3 shows that the information shock yields a large reduction in donations for inequality averse subjects (CHF −4.30, approximately −0.4 of a standard deviation,  $p = 0.023$ ) compared to the treated selfish subjects, consistent with our predictions. However, the treatment does not significantly affect the donations of the altruistic subjects (compared to the treated selfish) nor the donations of the selfish.<sup>29</sup>

Furthermore, prior work suggests that the effect of social preferences might be particularly pronounced at higher incomes (Fehr et al., Forthcoming). To test this hypothesis we estimate Eq. (1) separately for respondents with an income below the median, and for respondents with an income above the median. We depict the results of these estimations in columns 5–8.<sup>30</sup>

<sup>28</sup> For transparency, we also depict the full regression results in Table D.1 in Appendix D.6.

<sup>29</sup> Note that the treatment effect on the inequality averse (CHF −4.30) is larger in magnitude but not significantly different from the treatment effect on the altruistic subjects (CHF −2.95,  $p = 0.347$ ). This result is not inconsistent with our Hypothesis 2, which conjectures that treated inequality averse subjects will reduce their support for the 99% initiative in comparison to treated selfish — in particular at higher incomes.

<sup>30</sup> 51 subjects did not disclose their income and are thus not included in columns 5–8.

Consistent with previous findings, we find that — in the control group — the association between other-regarding preferences and support for the 99% initiative is particularly pronounced among individuals with incomes above the median. While the association is insignificant among individuals with an income below the median (see, e.g., column 6), it is relatively large and strongly significant among those with an income above the median (columns 7–8).<sup>31</sup>

Turning to the effect of the information intervention, we find that our treatment does not affect the demand for redistribution of subjects with an income below the median. Treated selfish subjects increase their support for the 99% initiative by CHF 1.73, but the effect is not significant ( $p = 0.540$ ). Compared to the treated selfish, treated inequality averse subjects reduce their support for the 99% initiative by CHF 2.31 and treated altruists by CHF 2.29, but neither of these effects is statistically significant. The picture is very different among subjects with an income above the median. There, too, treated selfish subjects slightly increase their support for the 99% initiative (CHF +2.55), but the effect is not significant ( $p = 0.348$ ). However, we observe a sharp decrease in support for the 99% initiative for the inequality averse subjects with an income above the median (CHF −6.54, or −0.59 of a standard deviation), compared to the treated selfish subjects with an income above the median ( $p = 0.022$  for  $H_0 : \beta_4 \geq 0$ , see the p-values for the relevant test at the bottom of the Table), consistent with what we conjectured in Hypothesis 2.<sup>32</sup> Finally, note that altruistic subjects with an income above the median also reduce their support for the 99% initiative (CHF −4.19) compared to the treated selfish with an above the median income, but the effect is smaller in magnitude than for the inequality averse and insignificant ( $p = 0.200$ ).

Overall, these results highlight the heterogeneous effects of beliefs about inequality and other-regarding preferences for demand for redistribution. While all subjects tend to have biased beliefs about the extent of income inequality, correcting these misperceptions significantly reduces the demand for redistribution of the inequality averse compared to the treated selfish subjects, in particular among individuals with higher incomes, consistent with what we conjectured.<sup>33</sup>

### 6.3. Robustness and additional analyses

One potential concern with information experiments conducted online is that some individuals might not pay attention to the information that is presented to them, which could considerably alter the results of the study (Berinsky et al., 2014). To account for this potential problem, our study included attention checks. Less than 10% of the subjects failed to correctly answer our screener questions, which is very reassuring and relatively low compared to related studies.<sup>34</sup> Overall, our main results are broadly robust to excluding subjects who did not pass the attention checks (see Appendix D.7), although the statistical significance of some of the results weakens when we restrict the sample

<sup>31</sup> The inequality averse subjects with an above-median income donate significantly more than the selfish (Column 8: CHF +6.41 or 0.58 of a standard deviation,  $p = 0.003$ ). Likewise, the altruistic subjects with an above-median income donate significantly more than the selfish (Column 8: CHF +5.43 or 0.49 of a standard deviation,  $p = 0.014$ ). In contrast, the coefficient for altruistic subjects with an income below the median is insignificant (Column 6:  $p = 0.140$ ) and the one for inequality averse subjects with an income below the median is only very weakly significant (Column 6:  $p = 0.067$ ).

<sup>32</sup> While we find evidence for Hypothesis 2, note that a test of equality of coefficient fails to reject the null hypothesis that the interaction coefficient between treatment and inequality aversion ( $\text{Treat} \times \text{IA}$ ) equals the interaction coefficient between treatment and altruism ( $\text{Treat} \times \text{Altruistic}$ ) ( $p = 0.354$ ).

<sup>33</sup> In Appendix D.9, we show that using self-reported measures of support for redistribution as a dependent variable instead of our incentivized measure of support for the 99% initiative yields qualitatively very similar results.

<sup>34</sup> For example, between 30 and 50 percent of the participants fail to pass the attention checks in the studies reported in Berinsky et al. (2014).

**Table 2**  
Determinants of donations in favor of the 99% initiative.

	Full sample				Below median income		Above median income	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	−1.106 (0.684)	−0.999 (0.668)	1.983 (1.881)	2.042 (1.825)	0.436 (2.694)	1.727 (2.815)	2.643 (2.860)	2.553 (2.718)
IA			6.871*** (1.519)	5.341*** (1.514)	4.721** (2.123)	4.501* (2.454)	8.232*** (2.288)	6.407*** (2.130)
Altruistic			5.001*** (1.554)	4.350*** (1.536)	3.777* (2.156)	3.730 (2.523)	5.133** (2.337)	5.434** (2.194)
Treat x IA			−4.298** (2.145)	−3.909* (2.091)	−1.001 (2.981)	−2.315 (3.077)	−7.252** (3.381)	−6.542** (3.247)
Treat x Altruistic			−2.951 (2.129)	−3.258 (2.066)	−0.716 (2.932)	−2.289 (3.069)	−4.252 (3.356)	−4.195 (3.266)
Constant	6.644*** (0.471)	−0.708 (4.592)	1.568 (1.376)	−4.795 (4.875)	3.314* (1.964)	−1.754 (6.243)	0.721 (1.996)	−11.661 (9.310)
Beliefs	No	Yes	No	Yes	No	Yes	No	Yes
Socio-demographics	No	Yes	No	Yes	No	Yes	No	Yes
Education	No	Yes	No	Yes	No	Yes	No	Yes
Occupation	No	Yes	No	Yes	No	Yes	No	Yes
Preference measures	No	Yes	No	Yes	No	Yes	No	Yes
Wealth bracket dummies & financial assets	No	Yes	No	Yes	No	Yes	No	Yes
p-value(Ho: Treat x IA ≥ 0)			0.023	0.031	0.369	0.226	0.016	0.022
p-value(Ho: Treat x Altruistic = 0)			0.166	0.115	0.807	0.456	0.206	0.200
p-value(Ho: Treat x IA = Treat x Altruistic)			0.347	0.645	0.868	0.988	0.234	0.354
p-value(Ho: Treatment = 0)	0.106	0.135	0.292	0.263	0.872	0.540	0.356	0.348
p-value(Ho: Treatment + Treat x IA = 0)			0.025	0.067	0.658	0.641	0.011	0.023
p-value(Ho: Treatment + Treat x Altruistic = 0)			0.332	0.214	0.809	0.638	0.360	0.359
R <sup>2</sup>	0.003	0.132	0.029	0.146	0.020	0.136	0.038	0.232
Observations	1031	1030	1031	1030	558	557	422	422

*Note:* OLS regression. The dependent variable is the donation amount towards an organization in favor of the 99% initiative for the full sample (columns 1-4), for subjects with an income below the median (columns 5-6), and for subjects with an income above the median (columns 7-8). Subjects who did not disclose their income are not included in columns 5-8. Beliefs include subjects' prior beliefs about the income share of the top 1%, their prior beliefs about income and wealth distributions in Switzerland, as well as their beliefs about the determinants of success, financial mobility, poverty in Switzerland, and distrust in politicians. Socio-demographics include age, age squared, a dummy variable indicating whether the respondent is male, a dummy variable indicating whether the respondent's native language is French, and a dummy indicating whether the respondent is married. Education includes dummies indicating a respondent's highest educational achievement (compulsory school, vocational training, high school, university, or other). Occupation includes dummies indicating a respondent's occupation status (currently has a full-time job, a part-time job, is a student, is a pensioner, is unemployed, or other), and a dummy variable indicating whether the respondent has experienced unemployment in the past. Preference measures from the global preference survey (Falk et al., 2022) include measures of risk aversion, (positive and negative) reciprocity, impatience, and general trust in strangers. Wealth bracket dummies and financial assets include dummy variables for each wealth level, and a dummy variable indicating whether the respondent has invested part of her wealth in financial assets. For a detailed explanation of the covariates, see Appendix B.1. Robust standard errors are reported in parentheses.  $p$ -value(Ho: Treat  $\times$  IA  $\geq 0$ ) indicates the  $p$ -value associated with a one-sided test evaluating the null hypothesis that the information intervention has a non-negative effect on the donations of the inequality averse subjects.

\* Levels of significance: \* $p < .1$ .

\*\* Levels of significance: \*\* $p < .05$ .

\*\*\* Levels of significance: \*\*\* $p < .01$ .

to subjects who passed both attention checks. However, note that these restrictions considerably reduce our sample size and our ability to detect treatment and interaction effects.

In addition, we also show in Appendix D.8 that our results are also broadly robust to focusing only on subjects who overestimated the income share of the top 1%, or to restricting the sample on subjects with “large” misperceptions, i.e., misperceptions greater than 10 percentage points — following the approach in Karadja et al. (2017) and Cruces et al. (2013).

The analysis reported throughout the paper focuses on subjects' donations to organizations that are in favor or against the 99% initiative. While this has the advantage of being a behavioral measure of support for a real world policy proposal, this measure has the drawback that it might conflate self-interest with policy preferences. In particular, it might be that selfish subjects donate little in support (or against) the 99% initiative because they are maximizing their own payoff, which might also explain why they do not react to the treatment. In Appendix D.9, we show that this is *not* the case. Indeed, our main results are broadly robust to using self-reported measures of support for redistribution as a main dependent variable.

Another potential concern with information provision studies are experimenter demand effects (see, e.g., Haaland et al., 2023). While


demand effects can be a serious concern in some settings, we believe that they are unlikely to be a threat to our results due to the preference-specificity of the predictions we are testing. Indeed, we hypothesized (Hypothesis 2) that a shock in beliefs about inequality would predominantly decrease the demand for redistribution of inequality averse subjects, in particular those with an income above the median. If our results were predominantly driven by experimenter demand effects, we should observe that *all* subjects adjust their demand for redistribution as a response to our information treatment, irrespective of their preference type and their income. This is however *not* what we find. Instead, the information intervention affects the demand for redistribution of inequality averse subjects, in particular the affluent ones, consistent with what we conjectured. In addition, recent methodological contributions have shown that strong demand effects (generated on purpose by the experimenter) result in only very modest behavioral responses in similar survey experiments (de Quidt et al., 2018).

## 7. Concluding remarks

Over the last decade, several studies have highlighted the role of beliefs about inequality as well as the role of other-regarding preferences for support for redistribution. These two strands of the literature

have largely evolved separately. In this paper, we studied the nature of support for redistribution by exploring the *joint* role of social preferences and beliefs about inequality. We also explored whether beliefs about inequality and beliefs updating depend on preferences. We investigated these questions by conducting an online experiment with a representative sample of the Swiss population in the context of the 99% initiative, a highly redistributive policy proposal aimed at increasing taxes on the richest individuals in Switzerland.

We showed that the vast majority of individuals overestimate the extent of income inequality in Switzerland, and that these misperceptions are independent of preference types. We also showed that all subjects update their beliefs correctly upon receiving information about the true share of total income received by the top 1% of income earners, irrespective of their preferences.

In addition, we showed that social preferences are an important predictor of support for the 99% initiative—consistent with past evidence that social preferences elicited experimentally can be helpful to understand people's support for real world policies (see, e.g., Fehr  *et al.*, Forthcoming; Fisman *et al.*, 2017; Kerschbamer and Müller, 2020). In our context, we find that support for the 99% initiative is particularly large among inequality averse respondents, especially at higher incomes.

Turning to the effects of beliefs about inequality for support for redistribution, we showed that while our information intervention successfully corrected subjects' misperceptions, it essentially did *not* affect average support for redistribution, consistent with a recent meta-analysis on the causal effects of beliefs about inequality for demand for redistribution (Ciani *et al.*, 2021). This average result hides, however, important heterogeneity. Indeed, the downwards shocks in beliefs about inequality generated a large and significant *decrease* in donations in favor of the 99% initiative among inequality averse respondents, but it did *not* affect the donations of the remaining subjects. This effect appears to be mainly driven by higher income individuals, consistent with what we hypothesized.

Altogether, these results highlight the joint importance of beliefs about inequality and other-regarding preferences for demand for redistribution. They also underscore the possibly large heterogeneous effects that information interventions might generate and that remain hidden in aggregate analyses. While previous work has explored the heterogeneous effects of information (about inequality) by political ideology, we show the existence of significant heterogeneities by preference type. These heterogeneous effects could explain the somewhat puzzling finding that information interventions often successfully generate large corrections in beliefs about inequality *without* substantially affecting demand for redistribution, as documented in a recent meta-analysis (Ciani *et al.*, 2021).

### Declaration of competing interest

(1) None of the authors have any relevant, material or financial interests that relate to the research described in this paper.

(2) None of the authors have held any position in organizations that relate to this research.

(3) No party outside of the authors has had the right to review the manuscript prior to submission.





### Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.pubeco.2025.105350>.

### Data availability

Data will be made available on request.

### References

- Alesina, Alberto, Angeletos, George-Marios, 2005. Fairness and redistribution. *Am. Econ. Rev.* 95 (4), 960–980.
- Alesina, Alberto, Stantcheva, Stefanie, Teso, Edoardo, 2018. Intergenerational mobility and preferences for redistribution. *Am. Econ. Rev.* 108 (2), 521–554.
- Almås, Ingvald, Cappelen, Alexander W., Tungodden, Bertil, 2020. Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians? *J. Political Econ.* 128 (5), 1753–1788.
- Alvaredo, Facundo, Atkinson, Anthony B., Piketty, Thomas, Saez, Emmanuel, 2013. The top 1 percent in international and historical perspective. *J. Econ. Perspect.* 27 (3), 3–20.
- Bellemare, Charles, Kröger, Sabine, Van Soest, Arthur, 2008. Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica* 76 (4), 815–839.
- Benabou, Roland, Ok, Efe A., 2001. Social mobility and the demand for redistribution: the POUM hypothesis. *Q. J. Econ.* 116 (2), 447–487.
- Benabou, Roland, Tirole, Jean, 2006. Belief in a just world and redistributive politics. *Q. J. Econ.* 121 (2), 699–746.
- Berinsky, Adam J., Margolis, Michele F., Sances, Michael W., 2014. Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *Am. J. Political Sci.* 58 (3), 739–753.
- Bolton, Gary E., Ockenfels, Axel, 2000. ERC: A theory of equity, reciprocity, and competition. *Am. Econ. Rev.* 90 (1), 166–193.
- Bruhin, Adrian, Fehr, Ernst, Schunk, Daniel, 2018. The many faces of human sociality: Uncovering the distribution and stability of social preferences. *J. Eur. Econ. Assoc.* 17 (4), 1025–1069.
- Charness, Gary, Rabin, Matthew, 2002. Understanding social preferences with simple tests. *Q. J. Econ.* 117 (3), 817–869.
- Ciani, Emanuele, Fréget, Louis, Manfredi, Thomas, 2021. Learning About Inequality and Demand for Redistribution: A Meta-Analysis of In-Survey Informational Experiments. OECD.
- Comiter, Marcus, Cha, Miriam, Kung, HT, Teerapittayanon, Surat, 2016. Lambda means clustering: automatic parameter search and distributed computing implementation. In: 2016 23rd International Conference on Pattern Recognition. ICPR, IEEE, pp. 2331–2337.
- Cruces, Guillermo, Perez-Truglia, Ricardo, Tetaz, Martin, 2013. Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *J. Public Econ.* 98, 100–112.
- de Quidt, Jonathan, Haushofer, Johannes, Roth, Christopher, 2018. Measuring and bounding experimenter demand. *Am. Econ. Rev.* 108 (11), 3266–3302.
- Falk, Armin, Becker, Anke, Dohmen, Thomas J., Huffman, David, Sunde, Uwe, 2022. The preference survey module: A validated instrument for measuring time, risk, and social preferences. *Manag. Sci.* 69 (4), 1935–1950.
- Falk, Armin, Zimmermann, Florian, 2013. A taste for consistency and survey response behavior. *CESifo Econ. Stud.* 59 (1), 181–193.
- Fehr, Ernst  Epper, Thomas  Senn, Julien, 2023. The fundamental properties, stability and predictive power of distributional preferences. University of Zurich, Department of Economics, Working Paper, 439.
- Fehr, Ernst  Epper, Thomas  Senn, Julien, 2024. Social preferences and redistributive politics. *Rev. Econ. Stat.* (Forthcoming).
- Fehr, Dietmar, Möllerstrom, Johanna, Perez-Truglia, Ricardo, 2022. Your place in the world: Relative income and global inequality. *Am. Econ. J.: Econ. Policy* 14 (4), 232–268.
- Fehr, Ernst, Schmidt, Klaus M., 1999. A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114 (3), 817–868.
- Fenton, George, 2020. How elastic are preferences for redistribution? New results on partisan polarization. SSRN Electron. J. <http://dx.doi.org/10.2139/ssrn.3538441>.
- Fisman, Raymond, Jakiela, Pamela, Kariv, Shachar, 2017. Distributional preferences and political behavior. *J. Public Econ.* 155, 1–10.
- Fisman, Raymond, Jakiela, Pamela, Kariv, Shachar, Markovits, Daniel, 2015. The distributional preferences of an elite. *Science* 349 (6254), aab0096.
- Fisman, Raymond, Kariv, Shachar, Markovits, Daniel, 2007. Individual preferences for giving. *Am. Econ. Rev.* 97 (5), 1858–1876.
- Foellmi, Reto, Martínez, Isabel Z., 2017. Volatile top income shares in Switzerland? Reassessing the evolution between 1981 and 2010. *Rev. Econ. Stat.* 99 (5), 793–809.
- Fong, Christina, 2001. Social preferences, self-interest, and the demand for redistribution. *J. Public Econ.* 82 (2), 225–246.
- Haaland, Ingar, Roth, Christopher, 2020. Labor market concerns and support for immigration. *J. Public Econ.* 191, 104256.
- Haaland, Ingar, Roth, Christopher, 2023. Beliefs about racial discrimination and support for pro-black policies. *Rev. Econ. Stat.* 105 (1), 40–53. [http://dx.doi.org/10.1162/rest\\_a.01036](http://dx.doi.org/10.1162/rest_a.01036).
- Haaland, Ingar, Roth, Christopher, Wohlfart, Johannes, 2023. Designing information provision experiments. *J. Econ. Lit.* 61 (1), 3–40.
- Karadja, Mounir, Möllerstrom, Johanna, Seim, David, 2017. Richer (and holier) than thou? The effect of relative income improvements on demand for redistribution. *Rev. Econ. Stat.* 99 (2), 201–212.

- Kerschbamer, Rudolf, Müller, D., 2020. Social preferences and political attitudes: An online experiment on a large heterogeneous sample. *J. Public Econ.* 182.
- Kulis, Brian, Jordan, Michael I., 2012. Revisiting k-means: New algorithms via Bayesian nonparametrics. In: *Proceedings of the 29th International Conference of Machine Learning*.
- Kuziemko, Ilyana, Norton, Michael I., Saez, Emmanuel, Stantcheva, Stefanie, 2015. How elastic are preferences for redistribution? Evidence from randomized survey experiments. *Am. Econ. Rev.* 105 (4), 1478–1508.
- Lergetporer, Philipp, Schwerdt, Guido, Werner, Katharina, West, Martin R., Woessmann, Ludger, 2018. How information affects support for education spending: Evidence from survey experiments in Germany and the United States. *J. Public Econ.* 167, 138–157.
- Lergetporer, Philipp, Werner, Katharina, Woessmann, Ludger, 2020. Educational inequality and public policy preferences: Evidence from representative survey experiments. *J. Public Econ.* 188, 104226. <http://dx.doi.org/10.1016/j.jpubeco.2020.104226>.
- Piketty, Thomas, 1995. Social mobility and redistributive politics. *Q. J. Econ.* 110 (3), 551–584.
- Roth, Christopher, Settele, Sonja, Wohlfart, Johannes, 2021. Beliefs about public debt and the demand for government spending. *J. Econometrics* 231 (1), 165–187.
- Roth, Christopher, Wohlfart, Johannes, 2020. How do expectations about the macroeconomy affect personal expectations and behavior? *Rev. Econ. Stat.* 102 (4), 731–748.
- Settele, Sonja, 2022. How do beliefs about the gender wage gap affect the demand for public policy? *Am. Econ. J.: Econ. Policy* 14 (2), 475–508.
- Swiss Federal Department of Finance, 2022. Verteilung der Pflichtigen nach Stufen des steuerbaren Einkommens. <https://www.estv.admin.ch/dam/estv/de/dokumente/estv/steuerstatistiken/direkte-bundessteuer/statistik-dbst-perzentilen-2018-de.xlsx.download.xlsx/statistik-dbst-perzentilen-2018-de.xlsx>.
- World Inequality Database, 2023. World inequality database. <https://wid.world/>. (Accessed 28 November 2023).