# Which Reference Groups Matter and How? A Relative Income Information Experiment with Administrative Data

*By* Xiaogeng Xu, Satu Metsälampi. Michael Kirchler, Kaisa Kotakorpi, Peter Hans Matthews and Topi Miettinen*

*Received wisdom holds that income rank matters for life satisfaction, but causal evidence on the nature and impact of income comparisons is limited. We randomize individuals from a representative sample of mid-career Finns to receive personal rank information from one of several reference groups. We find strong evidence of the effect of rank information on income satisfaction, but weaker effects on life satisfaction, and some evidence of real effects in experimental and administrative data. Effects are strong in narrow reference groups and weak and insignificant in the national one. Finally, we discuss the implications for income transparency policies.*

*JEL: D63, D8, D91, I31*

*Keywords: Relative position, individual welfare, fairness, comparison group, information provision*

In this paper, we focus on the fundamental question of how knowledge about income rank — relative to compatriots, neighbors, colleagues, age cohort, or educational peers — affects various measures of welfare. It is not a difficult one to motivate. As Luttmer (2005) and Clark and D'Ambrosio (2015) remind us, the notion that relative position has an important influence on human behavior is as old as social science itself (Smith, 1759; Mill, 1859; Veblen, 1899; Duesenberry, 1949; Festinger, 1954). Recent work on inequity aversion, fairness, and concern for income rank (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Cappelen et al., 2007; Bellemare, Kröger and Van Soest, 2008; Kuziemko et al., 2014; Martinangeli and Windsteiger, 2021) found much inspiration in this literature, and has shaped how we understand departures from narrow self-interest in choice data. Some studies suggest that relative income or rank may matter even more for happiness and satisfaction than absolute income (Clark, Frijters and Shields, 2008; Boyce, Brown and Moore, 2010; Fehr and Charness, 2023), and there is reason to believe that, due to social media, the importance of social comparison for happiness is on the rise (Haidt, 2024).[1]

In controlled laboratory experiments, the relevant reference group is assigned by design, leaving

* Xu: Hanken School of Economics and Helsinki GSE, Arkadiankatu 22, Fi-00101 Helsinki, xiaogeng.xu@hanken.fi; Metsälampi: University of Turku, satu.metsalampi@utu.fi; Kirchler: University of Innsbruck, michael.kirchler@uibk.ac.at; Kotakorpi: Tampere University and Finnish Centre of Excellence in Tax Systems Research, kaisa.kotakorpi@tuni.fi; Matthews: Middlebury College, Aalto School of Business, and Helsinki GSE, pmatthew@middlebury.edu; Miettinen: Hanken School of Economics and Helsinki GSE, Arkadiankatu 22, Fi-00101 Helsinki, topi.miettinen@hanken.fi. Previously circulated under the title "The Welfare Consequences of Learning Where One Stands: Evidence From A Large Field Experiment". We acknowledge the financial support from The Academy of Finland, project (grant no. 332550), Finnish Centre of Excellence in Tax Systems Research / Academy of Finland (grant no. 346250), the Austrian Science Fund (FWF, project SFB F6310), Finnish Cultural Foundation (grant no. 00210723) and the Yrjö Jahnsson Foundation (grant no. 20177011) and Fulbright Finland Foundation. We are grateful for the feedback from B. Bartling, D. Fehr, E. Fehr, I. Haaland, J. Haushofer, C. Kreiner, J. Mollerstrom, A. Oswald, R. Perez-Truglia, T. Ravaska, R. Schwaiger, U. Simonsohn, M. Sutter, S. Stantcheva, E. Ø. Sørensen, J. Tukiainen, B. Tungodden, U. Weitzel, E. Wengström, L. Windsteiger, C. Young, and audiences at Helsinki GSE, Kiel Institute, TSE, BEEL 10th Anniversary Workshop Birmingham, VATT, ESA 2022 Boston & Bologna, $38^{th}$ Summer Seminar of Finnish Economists, FAIR Midway conference Tromsø 2022, SABE 2022 Lake Tahoe, IIPF Linz 2022, $43^{rd}$ Finnish Economic Association Meeting 2023, NHH-FAIR, ECBE 2023, ESA World Meeting 2023, WEI workshop, EEA 2023, NCBEE 2023 and the LSE Well-Being Seminar 2024. The last author is grateful for the hospitality of University of Michigan in early stages of the project. The study was approved by Hanken's Research Ethics Committee on December 12, 2019. The RCT ID is AEARCTR-0011720.

[1]See Vogel et al. (2014, 2015); Reer, Tang and Quandt (2019); Twenge et al. (2022) on how negative effects of social media play out through increased social comparison.

researchers to wonder which group(s) are the basis for comparison in natural settings, and which have the strongest effects on behavior, outcomes and well-being. These questions are central in a number of contexts. In organizations, comparisons to bosses or co-workers will have different effects on incentives, performance, and job turnover (Englmaier and Wambach, 2010; Bartling, 2011; Kőszegi, 2014; Card et al., 2012; Dube, Giuliano and Leonard, 2019; Cullen and Perez-Truglia, 2022). In the public sphere, narrow and broad comparison groups will influence policy preferences and the consequent behavioral responses (Fehr, Mollerstrom and Perez-Truglia, 2022). If what matters to people is "inequality as *experienced* difference" (Bowles and Carlin, 2020), then inequality within occupation, neighborhood, or educational peers is both salient and felt whereas within some national distribution it may be an abstraction. In a similar vein, Genicot and Ray (2020) argue that aspirations relate to "lives that are on display". Finally, if relative income concerns matter for individual welfare, they also matter for aggregate welfare, and it is critical to know which reference groups matter most to people.

We invited a representative sample of 20,000 mid-career Finns aged 35 to 45 to participate in a pre-registered information provision experiment and about 6,600 invitees responded.[2] We did so in cooperation with Statistics Finland and were able to link the experimental data with administrative data. We first elicited incentivized beliefs about income rank in various reference groups, and then assigned participants at random to one of five information treatments, using a variant of the information protocol summarized in Haaland, Roth and Wohlfart (2023). Each of the treatments provided rank information in a particular reference group: based on the treatment, individuals learned about their rank in the distribution of disposable income for their occupation, age cohort, educational level, municipality, or the national distribution. In addition, there was a no-information control treatment. We report estimates of the causal effects of rank information on various measures of subjective well-being and real outcomes measured in our survey and administrative data.

As a major finding, we report significant causal effects of information about rank on income-related well-being measures: satisfaction with disposable income, wage satisfaction, and perceived fairness of own income. We also find and rationalize weaker (but not very precisely estimated) effects on job and life satisfaction. Furthermore, the effects on satisfaction with disposable income are strong for circumscribed groups—educational level, occupation, age cohort, and municipality—but small and insignificant for the national reference group. Educational level emerges as a particularly important reference group in which information about income rank has an effect on all income-related well-being measures. We also show that our key results are very robust to different specifications, by providing a multiverse analysis (Simonsohn, Simmons and Nelson, 2020), a method that is well suited for experiments with large representative samples linked to administrative data that cannot be shared on open science platforms.

We also find a significant real effect on earned income in the occupational reference group: Learning that income rank is lower than expected among people in the same occupation significantly increases earned income in official income registers in 2021. Estimated effects in other narrow reference groups (but not national) are consistent.[3] We also observe a treatment effect, on the intensive margin, on (real) charitable donations within our survey: learning that income rank is lower than expected lowers donations.

Further, results from an additional treatment that allows endogenous choice of reference group information, support the interpretation that more circumscribed reference groups matter most to people. That is, individuals are most likely to choose to find out their rank in the occupational reference group, while respondents seldom choose to learn about their position in the national income distribution.

Our main contribution is the comparison of the effects of relative income information across refer-

[2] The pre-analysis plan and the oTree codes of the experimental software can be found at the link.

[3] See Jäger et al. (2024) for a similar effect among occupational peers on intentions to search for jobs and negotiate wages.

ence groups. In their survey, Clark and D'Ambrosio (2015) credit Hyman (1942) with the identification of the relevant reference group as a critical feature of the psychology of status. The national reference group has been prominent in the earlier literature (Easterlin, 1974; Alesina, Di Tella and MacCulloch, 2004; Di Tella, Haisken-De New and MacCulloch, 2010; Wilkinson and Pickett, 2010). There are other studies that have examined the effects of relative income or rank for one (assumed) reference group on happiness or satisfaction, including the nation as a whole, workplace, neighborhood, education groups or age cohorts (Brown et al., 2008; Clark, Westergård-Nielsen and Kristensen, 2009; Clark and Senik, 2010; Ferrer-i-Carbonell, 2005; Godechot and Senik, 2015; McBride, 2001; Perez-Truglia, 2020). Our findings on stronger effects in narrow reference groups suggest that some of the literature, often focusing on national comparisons, may underestimate the true magnitude of social status effects.

Indeed, there are few, if any, comprehensive comparisons of causal effects of relative income information across reference groups. In descriptive research, Clark and Senik (2010) report that work colleagues are the most important reference group, at least for income comparisons. Reck, Slemrod and Vattø (2022) investigate the network of searches through the Norwegian income registers and finds that narrow employment and household networks featuring homophily arise as particularly relevant. Neither provides estimates of causal effects of the learned information, however. One of the earliest exceptions is Yamada and Sato (2013), who report on a series of hypothetical discrete choice experiments designed to "avoid the problems associated with researcher-imposed reference" incomes. Our own experiment does this in the field, with real incomes. Cullen and Perez-Truglia (2022) and Fehr, Mollerstrom and Perez-Truglia (2022) use methods similar to ours and randomly assign participants to receive relative income information in two alternative reference groups: boss vs. peers, and national vs. global, respectively. The study closest to ours is perhaps Hvidberg, Kreiner and Stantcheva (2023) who use an information provision experiment and Danish administrative data to examine the nature and effects of misperception about rank in various reference groups. However, they examine the effect on fairness perceptions of revealing relative income information in several reference groups at the same time, and thus they cannot estimate the causal effects of rank information in different reference groups. Further, all their reference groups are narrower than our age reference group due to conditioning on cohort throughout. Our experiment, on the other hand, is designed to allow a comparison of the causal effects of rank information in the broad national reference group, and various narrower ones. We achieve this by randomizing participants into treatments where they receive information about rank in one (and only one) reference group.

In addition to our key aim of analyzing the effects of relative income information in different reference groups, our paper makes three additional contributions. First, just as it is common to limit attention to a single frame of reference, it is also prevalent to focus on a single measure of happiness or welfare. We elicit both narrower income satisfaction measures relating to disposable income (Veblen, 1899; Frank, 1989; Corneo and Jeanne, 1997), wage income, and the perceived fairness of own income (Fehr and Schmidt, 2006; Almås, Cappelen and Tungodden, 2020), and broader satisfaction measures with life and job satisfaction, in an attempt to identify effects on different dimensions of well-being. Moreover, we study longer-run effects on wage income and turnover in register data (Card et al., 2012; Jäger et al., 2024), as well as on incentivized outcomes such as charitable donations within our survey.[4]

Second, our protocol offers another reminder that *experienced* or *believed* position is different than *actual* rank, and extends previous results on imperfect knowledge concerning rank (Karadja, Mollerstrom and Seim, 2017; Fehr, Mollerstrom and Perez-Truglia, 2022; Hvidberg, Kreiner and

[4]Unlike Card et al. (2012) or Perez-Truglia (2020), however, our setting abstracts from the social image effects of income rank, in which interested individuals can discover each other's ranks. We also note that anyone in Finland can enter a tax office to request access to information about the taxable income of another taxpayer, but that information about income ranks in various distributions is not easily accessible.

Stantcheva, 2023; van Rooij et al., 2024). We uncover some intriguing patterns in these misperceptions, such as the prevalent underestimation of income rank, and the high correlation of biases in rank beliefs between some but not all reference groups.

Third, we consider an important policy application of these results, namely, the consequences of income transparency policies. Our results highlight the importance of the structure of baseline misperceptions for the welfare effects of transparency policies. Even though relative income concerns are typically viewed as giving rise to a negative externality, we find large positive aggregate effects of providing rank information in many of the distributions that we consider. This can be attributed to the positive effects of personal rank information on pessimists, who outnumber optimists almost nine to one. It should be noted, however, that our analysis relates to the immediate effects of rank information on subjective well-being, and abstracts from potential broader effects e.g., via labor markets (Cullen and Pakzad-Hurson, 2023; Cullen, 2024).

In sum, our contribution is to provide novel evidence on two central but understudied questions in the study of well-being: how does knowledge about position in different reference groups matter to people, and for what measures of well-being is relative position important? Our results should matter not just to researchers interested in richer explanations of human behavior but also to policy-makers for the design of effective policies.

The remainder of the paper continues as follows. Section I outlines our pre-registered design and its implementation, including data on balance, selection, and attrition. Section II summarizes and dissects our main results, starting with characterization of misperceptions both within and across distributions in Subsection II.A, then studying the causal effects in Subsections II.B, II.C, II.D, and their robustness in Subsection II.E. Section III offers further thoughts on implications for policy. Section IV concludes.

## I. Research Design

We conducted a pre-registered information provision experiment in cooperation with Statistics Finland (SF) in the summer of 2021 after conducting a pilot study with 2500 invitations in late 2020.[5] We designed a personalized online survey (in Finnish and Swedish) containing incentivized belief elicitation and outcome measures, an information provision treatment, and standard survey questions (see English translations of the survey and oTree code for the experimental software at the pre-registration). We invited a representative sample of 35 to 45 year old Finns who had not permanently left the labor force at the time of the survey to participate.[6] The survey data is linked to SF's administrative records.

### A. *Survey design*

The survey consists of five sections: background, incentivized income rank belief elicitation, income rank information treatment, outcomes, and summary. The survey is the same for all respondents except for the information provision treatment, which varies according to the treatment assignment.

[5]The pre-analysis plan (PAP) of the pilot study can be found at https://doi.org/10.17605/OSF.IO/YJ5U4. The PAP for the pilot was posted at the end of the pilot data collection and before the data analysis. The PAP for the main collection can be found at https://doi.org/10.17605/OSF.IO/DJQ3G and was posted before the data collection started. To comply with a later external requirement, we posted another registration at https://www.socialscienceregistry.org/trials/11720. This second registration is entirely based on the OSF-registration of the main collection.

[6]More precisely, we restrict the sample to people born between 1975 and 1985, who had a Finnish social security number in 2010 (to approximate having lived in Finland for at least 10 years) excluding residents of Åland Islands, whose mother tongue is either Finnish or Swedish, who have non-missing income and occupational information (for 2018) and whose family status is not "child" in SF register data. SF oversampled participants with basic and upper secondary education to account for the expected unit non-response implied by the response rates in the pilot survey. We chose to focus on this target population as people in their mid-career have had a chance to establish themselves in the labor force and still have active years ahead so that information about relative income has a chance of affecting their career and other choices.

**Background and belief elicitation.** Respondents log in with a personal username and password provided in the survey invitation.[7] In the first part of the survey, participants answer questions about their birth year, gender, marital status, highest completed education, occupation, and municipality of residence.[8]

In the income rank belief elicitation section participants are asked to report their beliefs concerning their disposable income rank in 2018 among individuals in each of five reference groups (e.g., the national reference group includes all adults in Finland).[9] The beliefs are elicited for each reference group in random order. To incentivize the assessment, we rewarded participants whose rank assessment for a reference group chosen at random was "correct." Following the method of belief elicitation of Schlag and Tremewan (2021), an answer was considered correct if it fell within the same 5-point interval (e.g., 0-5%, 6-10%, ..., 96-100%) as the actual rank.[10] The participants learn whether they receive the bonus only at the end of the survey where they are told the correct answer in the randomly drawn reference group (see Figure A2 in Appendix). We clarify the definition of disposable income and remind the participants of the definition along the survey wherever there is a question regarding income.[11]

**Treatments.** The third section of the survey is the information provision treatment. The participants are provided information concerning their disposable income rank among individuals in the reference group corresponding to the treatment they are randomly assigned to.[12] This allows us to identify the causal effect of that piece of information alone.

Our seven treatments are summarized in Table 1. The participants in the CONTROL treatment receive no information about their income rank. The participants in the five treatments with exogenously assigned information (AGE, MUNICIPALITY, EDUCATION, OCCUPATION, NATIONAL) receive information about their income rank in the corresponding reference group. For instance, the participants in treatment EDUCATION are informed of their income rank among all Finns with the same level of education as the participant (see Figure A1 for an example of the treatment information). The broadest reference group is the NATIONAL reference group. In the endogenous information treatment CHOICE participants choose one of the five reference groups and later receive the chosen information. The rank information is provided alongside the perceived rank which the participants reported in the previous section.[13] We also ask the participants to answer a control question to ensure that they understand the information correctly before proceeding.

The conceptual motivation behind analyzing these different reference groups stems from the fact that other-regarding agents' behavior and well-being is partly driven by relative income concerns. Without better understanding which comparisons are particularly important, we lack central knowledge for the design of incentive schemes at organizational and interventions at societal level.

[7]Logging in with the username was necessary for the tailored information in the survey and helped prevent duplicate participation, see details in Appendix B.

[8]The last three questions concern their situation in 2018, the latest year data was available in the SF registers at the time of study. The rank and reference group information used in the information provision treatments concern the year 2018, and the goal of these questions is to help the respondent to recall their situation in the relevant time period and allow us to determine whether self-reported reference groups match those in the administrative data used in constructing the treatment rank information.

[9]The respondents indicate the percentage of the population of each reference group who they believe had lower disposable income than their own (see the survey screens at the link).

[10]This method is simple to understand and robust to bias generated by risk-aversion, unlike some scoring rules widely used in the literature.

[11]Following the definition of disposable income by Statistics Finland, we state in the survey: "By income, we refer to the total after tax annual income, which contains income from labor and capital, as well as all transfers and subsidies like unemployment benefits and pensions (i.e., disposable income)." This differs from the definitions used in some existing studies: Karadja, Mollerstrom and Seim (2017) use both labor and capital income before taxes, including pensions but exclusive of transfers such as unemployment insurance, Fehr, Mollerstrom and Perez-Truglia (2022) use household level income, and Hvidberg, Kreiner and Stantcheva (2023) focus on wage income alone. We use the *in-pocket-money* definition because it approximates well the standard of living and potential consumption, and is closely related to individual well-being.

[12]Notice that this group is not necessarily the same based on which their belief elicitation was rewarded.

[13]We also clarify that the reference group for which the income rank information is provided is not necessarily the randomly chosen reference group that determines the bonus payoff.

Table 1—. Treatments and reference groups

| Treatment | Description |
|---|---|
| CONTROL | No information about income rank |
| AGE | Exogenous information: income rank among people born in the same year |
| MUNICIPALITY | Exogenous information: income rank among adults living in the same municipality |
| EDUCATION | Exogenous information: income rank among people with same level of education (Level of education defined as basic, upper secondary, bachelor, master or higher. Classification is based on ISCED 2011.) |
| OCCUPATION | Exogenous information: income rank among people with same occupation (Classification is based on classification of occupations 2010 on 2-digit level, which follows the structure of ISCO-08, e.g. "teaching professionals", "sales workers.") |
| NATIONAL | Exogenous information: income rank among adult Finns |
| CHOICE | Endogenous information: income rank among the chosen reference group |

*Notes:* This table presents the treatments and reference groups used in the treatments. The reference group *National* is defined as all adult (aged 18 or older) Finns who had non-zero incomes in 2018. The other reference groups are subsets of *National*, such that reference group *Age* refers to individuals who had non-zero incomes in 2018 and were born in the same year as the participant, *Municipality* refers to all adult Finns with non zero incomes who lived in the same municipality in 2018, etc. See the occupational groups at SF's website.

Indeed, job turnover and performance, educational decisions, and human capital investments, consumer choices, and housing or financial market decisions may well be influenced by relative income concerns or by image related to relative income (Bursztyn and Jensen, 2017; Genicot and Ray, 2020). If inequality aversion, fairness concerns and conspicuous motives are driven by experienced and salient differences (Bowles and Carlin, 2020), then narrow reference groups, such as municipality, age, occupation, or education are more important than broader ones, such as the national one. Considerations related to merit, and instrumental motives (Hirschman and Rothschild, 1973; Clark and Senik, 2010; Almås, Cappelen and Tungodden, 2020) might turn attention towards occupational and educational reference groups. As societal policy is designed predominantly at the level of the nation state and the equality narratives in public media often emphasize the national comparisons, the national reference group could receive special attention in relative income comparisons. The national reference group also constitutes an important benchmark against which other reference groups can be compared to.

**Outcomes.** The outcome section of the survey consists of a series of standard survey questions concerning individual well-being and views toward societal and political issues, and a set of real stakes questions. There are six blocks of questions: (i) individual well-being (fairness perceptions of own income, satisfaction with disposable income, life satisfaction, job/wage satisfaction, job meaningfulness, and job search intentions); (ii) ideal income distribution, trust in institutions, and attitudes toward policies (tax, labor market, welfare, migration, and trade); (iii) just world beliefs; (iv) self-assessment and social preferences; (v) political orientations; and (vi) real stakes questions. The order of the first three blocks and the order of the questions within each block are both randomized (with some exceptions, see Appendix B for a full description of questions and question ordering principles).[14]

In this paper we focus on the impact of rank information on individual well-being (block (i)).[15]

[14]Because we can only ask the participants who are active in the labor market about their job and wage satisfaction, the questions related to one's job are asked after the participant has answered a question on their current employment status.

[15]This paper discusses the outcomes and treatments described as "Project 1" in the pre-analysis plan at https://doi.org/10.17605/OSF.IO/DJQ3G. Other outcomes and the endogenous information treatment CHOICE will comprise

In this block, participants were asked to use sliders to report various measures of individual welfare, including satisfaction with disposable income, a featured outcome, from "disappointed" to "pleased," as well as fairness of disposable income ("unfairly low" to "fair" to "unfairly high"), life satisfaction ("extremely dissatisfied" to "extremely satisfied"), likelihood of job search in the next six months ("very unlikely" to "very likely"), wage and job satisfaction ("not at all satisfied" to "very satisfied") and meaningfulness of one's job ("not at all meaningful" to "very meaningful"). To prevent a priming effect, there is no default position on any slider in the survey (see the survey screens at the link).[16] With the exception of wage and job satisfaction, the order was randomized: we asked the former before the latter in order to encourage assessments of job satisfaction *net* of wage satisfaction.

These outcomes are motivated by the existing literature on relative income concerns. It is disposable income that matters for conspicuous consumption motives (Veblen, 1899; Frank, 1985; Corneo and Jeanne, 1997; Kuhn et al., 2011); thus, satisfaction with disposable income is of key interest. The vast literature on inequity aversion and fairness (Fehr and Schmidt, 2006; Almås, Cappelen and Tungodden, 2020) calls for gauging fairness perception of own income relative to others. The wage satisfaction is particularly central for merit-based assessments of relative income which are known to be emphasized by the majority of the population irrespective of the extent of redistributive measures taken in a country (Almås, Cappelen and Tungodden, 2020). We also wanted to understand the difference between wage and job satisfaction, thus justifying the elicitation of the latter. Life satisfaction is a commonly used measure in the literature (Easterlin, 1974; Clark and D'Ambrosio, 2015) and is available in central international surveys. Thus, it is of interest to understand the effect of relative income information on general life satisfaction alongside more narrow income satisfaction measures. When we present our results, we divide the outcome measures into income-related measures, and broader well-being measures not specifically related to income.

### B. *Survey implementation*

The online survey was developed with oTree (Chen, Schonger and Wickens, 2016) and hosted on the server of Hanken School of Economics, Helsinki, Finland. The survey was personalized and contained embedded information concerning the respondents. The information, provided by SF, included the respondent's occupational group, disposable income rank in the five reference groups, as well as the treatment to which the respondent was randomly assigned.[17]

Both the pilot and main study were pre-registered. Unless otherwise specified, the results reported below are based on specifications in the pre-analysis plan, and the exceptions are reviewer requests. Data collection took place in the Summer of 2021 (see Figure A3 for the timeline of the study). The invitation letters were sent to 20,000 individuals by SF via mail, accompanied with email and text message reminders.[18] Participants received 15 euro for completing the survey and an additional 5

subsequent projects. Based on editor and reviewer requests, we have made the following deviations from the pre-registered list of outcomes: we dropped the job search intentions variable and included instead (within survey, incentivized) charitable donations, voluntary tax donations and lottery ticket purchases, and earned income (income register data in the year of the intervention (2021), and the year after.

[16] When answering with any slider, the participants need to tap on the slider and a thumb shows up at the tapped location. We use the visual analogue scale (VAS) in all survey (outcome) items discussed in this paper except for job search intentions which requires a categorical answer. The analogue ratings not only give greater resolution of scale and can be considered continuous, but helps mitigate the concern for discrete likert scale. Bond and Lang (2019) show that if we let people answer about their happiness by choosing from a few discrete categories without knowing the underlying distribution, we cannot easily compare the average level of happiness between groups. While Kaiser and Oswald (2022) alleviate the concern by showing the linear relationship between happiness and behavior, the concern may apply to the measurement of some outcomes other than happiness, such as fairness.

[17] The randomization was done by SF. The invitees were assigned into treatments according to 36 strata based on the following characteristics: gender (male, female), income (three classes by percentiles with cutoffs at 33.3% and 66.6%), statistical grouping of municipalities (urban municipalities, semi-urban municipalities, rural municipalities) and educational degree (basic education, other).

[18] The letter included a general description of the study and how the survey data is used, link to a data protection description,

euro depending on their response in the incentivized belief elicitation task. Participants could also use (some or all of) their payment as a donation to a charity (Save the Children), as a voluntary tax, or choose to receive a corresponding amount of lottery tickets in the real stakes questions (block vi). One of these three outcomes/purposes was randomly drawn for each participant at the end of the survey and the amount spent was subtracted from the amount sent as a gift card. The lottery tickets were sent via mail by SF, and the donations to charity and voluntary taxes were handled by Hanken School of Economics. The receipts of total sums of donations were posted online and messaged to the participants after data collection ended. The participation payments were sent as gift cards via text message or mail by SF.

### C. *Sample and data*

Table B1 presents information provided by SF on the survey sample and respondents. Out of the 20,000 invited individuals (column 1), 6,642 (33%) started completing the survey (column 3). Respondents tend to be somewhat more highly educated, have higher incomes and are more likely to reside in the Metropolitan area than non-respondents. Summary statistics in Table B2 show that these differences are more muted when comparing the respondents (column 1, and column 2 for individuals who completed the survey) to the target population (column 3).

All observations from treatments AGE, MUNICIPALITY, EDUCATION, OCCUPATION, NATIONAL and CONTROL, (also incomplete answers), are included in our main sample of analysis.[19] The number of invitees and the response and completion rates by treatment are reported in Table B3. Table B4 shows the relation between the overall and post-treatment attrition and background characteristics of the respondents. Women, and also respondents assigned to the treatment MUNICIPALITY tend to drop out more often, but there are no significant differences in attrition between treatments at any time after the participants have received the rank information treatments. Finally, Table B5 presents the summary statistics of socio-economic variables in the five treatments against the CONTROL treatment.

The analyses in this paper use survey data and register data. The match rate between the responses in the survey data and register data is high.[20] The outcome, treatment and control variables used in the main specification are described in Table A1.

### D. *Econometric specification*

We adopt the following econometric framework to evaluate the effects of relative income information:[21]

$$Y_i^k = \beta_0 + \beta_1(ER_i^j - R_i^j) + \beta_2 T_i^j + \beta_3 T_i^j \cdot (ER_i^j - R_i^j) + \gamma \mathbf{X_i} + u_i. \tag{1}$$

Here $Y_i$ is the value of outcome $k$ for individual $i$, $R_i^j$ is $i$'s actual rank in distribution $j$, $ER_i^j$ is the same individual's belief about her rank in $j$, so that $ER_i^j - R_i^j$ is her misperception about rank, $T_i^j$ is a treatment indicator that is equal to 1 if $i$ is shown her actual rank in distribution $j$, and 0 otherwise, and $\mathbf{X_i}$ is a vector of control variables. Our outcome variables relate to one's satisfaction with income and to general satisfaction (not necessarily related to income). Equation (1) is our pre-

link to the survey (URL and QR code), and a personal username and password.

[19]Alternative data restrictions, including restricting the sample to only respondents who completed the survey, are applied in the robustness checks (see Section II.E).

[20]The match rate is 0.99 for birth year and gender, 0.90 for municipality of residence and educational level, and 0.74 for occupational group.

[21]The design is similar to a few recent information provision experiments summarized in Haaland, Roth and Wohlfart (2023).

registered main specification, but we also carry out a multiverse analysis as exploratory analysis—i.e., the specification curve analysis of Simonsohn, Simmons and Nelson (2020)—in Section II.E, where we confirm the robustness of our findings to different specifications.

The model is estimated over five subsamples that each include the control group and one of our treatment groups. In each case, the main coefficient of interest is $\beta_3$, which estimates the causal effect of information about rank in a particular reference group on the relevant outcome. Coefficient $\beta_1$ on the other hand measures the relationship between misperception about rank and the outcome in the CONTROL condition, where no information is provided. We expect the coefficient $\beta_3$ to be of a sign opposite to $\beta_1$: This would indicate that our treatment truly provides meaningful information that serves to at least partially reverse the implications of initial misperceptions on well-being. On the other hand, $\beta_2$ provides an estimate of the treatment effect for those whose initial misperception about rank is zero, and hence we expect $\beta_2$ to be zero: as we describe in our PAP, information about rank should not matter for those individuals who have correct information in the first place.

A crucial feature of our design is that it allows us to cleanly identify the causal effects of rank information in each particular reference group. Two remarks on the interpretation of our estimates are in order. First, to be specific, the coefficient $\beta_3$ measures the intention-to-treat effect of information provision, rather than the effect of belief updating directly. To the extent that belief updating is less than complete, the causal effects of rank beliefs on well-being would be larger than the ITT effects reported here. In this sense, our estimates provide a lower bound of the causal effects of rank beliefs. However, our main focus is not on the effects of rank beliefs but on the effects of rank information overall, and conditional on misperception.

Second, while $\beta_3$ provides the reduced form estimate of the effects of rank information in a particular reference group, the mechanisms behind this reduced form effect may be manifold. As discussed above, individuals likely update their beliefs about their position in the particular reference group, but could also update beliefs about positions in other reference groups, to the extent that positions, and misbeliefs, are correlated across distributions. This phenomenon, which has been referred to as belief spillover or cross-learning in the literature, is a natural by-product of information provision experiments (Haaland, Roth and Wohlfart, 2023). Such cross-learning is embodied in our reduced form estimates, and our main interest lies in comparing such composite effects of information about rank rather than the effects of income rank per se or the identification of spillovers. We discuss the implications for the interpretation of our results further at the end of Section II.

## II. Results

### A. *Misperceptions about income rank*

Figure 1 contrasts the distributions of actual and perceived income ranks in reference groups *education* (panel a), *occupation* (panel b), *municipality* (panel c), *age* (panel d) and *national* (panel e). Most respondents report a position that is lower than their actual rank; 90% underestimate their position in the national and municipal income distributions, 80% in Education, 75% in Age and 70% in Occupation.[22] In a similar vein, misperceptions about rank in the national distribution

[22]This is consistent with what we hypothesize in the pre-analysis plan concerning the nature of misperceptions. Systematic underestimation of own position in the national income distribution is in line with Karadja, Mollerstrom and Seim (2017) and Bublitz (2020) who found systematic underestimation of national income rank in a study of six countries. (It is worth noting that our results are similar despite the quite different income profiles of our respective samples.) By contrast, Cruces, Perez-Truglia and Tetaz (2013), Fernández-Albertos and Kuo (2018), and Engelhardt and Wagener (2018), Fehr, Mollerstrom and Perez-Truglia (2022) and Hvidberg, Kreiner and Stantcheva (2023) find more balanced misperceptions among Argentinian, Spanish, German and Danish individuals, respectively, but notice that our sample consists of 35-45-year-old active in labor force. In a survey experiment across Australia, India, Mexico, Morocco, Netherlands, Nigeria, South Africa, Spain, United Kingdom and United States, Hoy and Mager (2021) find that respondents struggle to place themselves in the correct quintile in the national distribution of household incomes and that respondents in high income countries have more accurate beliefs.

are largest: Respondents underestimate their position by 22 percentage points on average (see Table C1). The smallest average misperception concerns position in the age group's distribution.

Misperceptions are significantly correlated and correlation is strongest between National and Municipality income distributions ($\rho = 0.79$, $p < 0.001$, panel c in Figure C2).[23] Visual comparison of misperceptions for the broad National reference group vs. the narrower reference groups is suggestive of this, too (see Figure C1).

Table C2 presents the results of a regression of absolute misperceptions on individual characteristics. The determinants of misperceptions are similar in all reference groups, consistent with strong correlations between misperceptions (panel c in Figure C2). In line with the results reported in Hvidberg, Kreiner and Stantcheva (2023), women tend to hold more inaccurate views than men. Having children is associated with less accurate perceptions, whereas living with a spouse is associated with more accurate perceptions. Highly educated individuals hold more accurate beliefs about own income position.

### *B. Effects of relative income information on well-being*

**Graphical analysis**. We start by grouping our outcome variables into income-related and broader measures of well-being, and illustrating the effects of rank information in the different reference groups in Figures 2 and 3.

The diagrams depict the difference in average well-being between those who were treated with rank information in a particular reference group and those who did not receive such information. The data are divided into three bins corresponding to initial misperception: respondents with positive misperception (i.e. negative surprise for the treated); (approximately) correct belief; and negative misperception (i.e. positive surprise).

The first and most important conclusion from Figure 2 is the existence of a significant "misperception gradient" for almost all reference groups for income-related well-being measures: the respondents who experienced a positive (negative) surprise about their rank were more (less) pleased with their disposable income, perceived more (less) fairness about their disposable income, and were more (less) satisfied with their wage. In short, there is causal evidence that rank information matters for well-being, a central result.

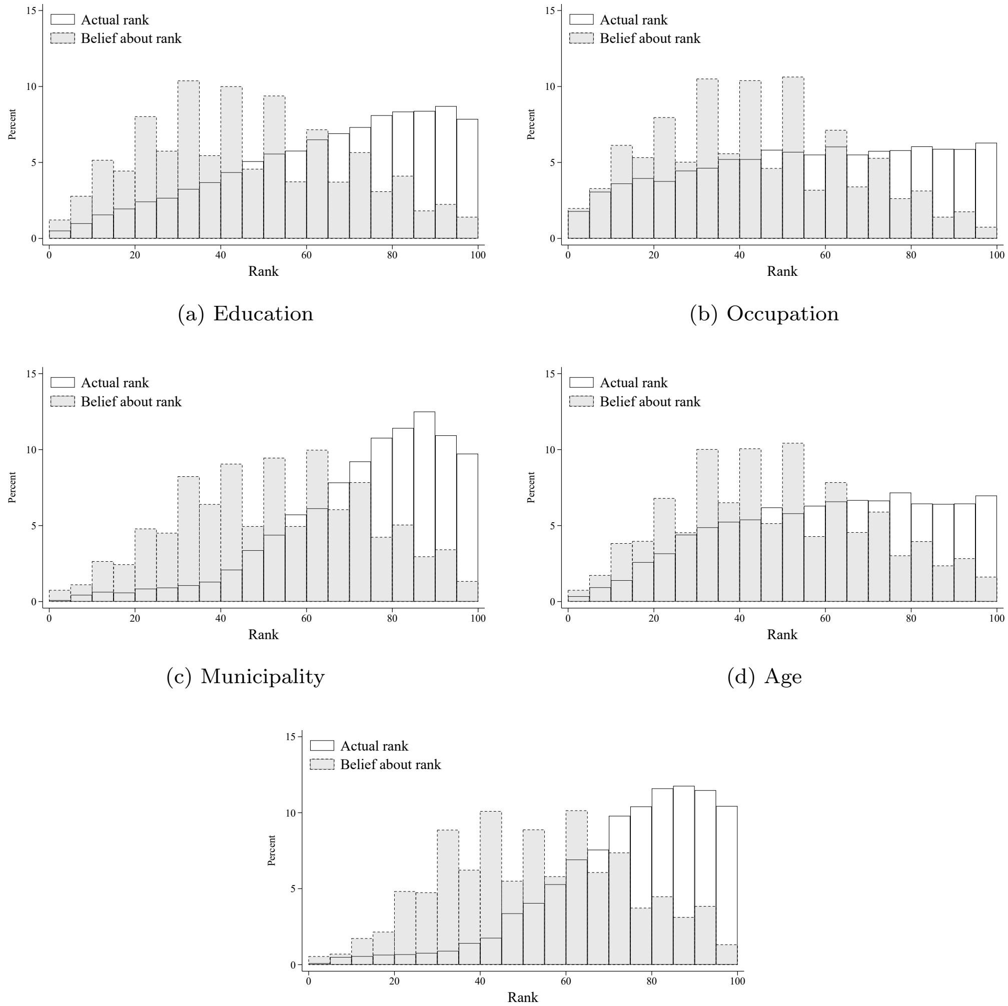
There are two further takeaways from Figure 2. First, the figure allows us to examine the hypothesis that information provision should not matter to those whose initial beliefs are correct, an important check on the internal validity of our design. Indeed, the figure indicates that across all reference groups, among the respondents whose initial beliefs were correct, the differences between those who learn their actual position and those who do not, is close to zero.

Second, Figure 2 allows a preliminary examination of whether negative and positive surprises have different effects on well-being. With few exceptions, the figure provides scant evidence of such asymmetric effects, a result at odds with previous studies (Di Tella, Haisken-De New and MacCulloch, 2010). Admittedly, the effects of negative surprises are imprecisely estimated in the figure, due to the fact few people overestimate their position, which makes asymmetries difficult to detect. We provide a more comprehensive econometric analysis of the above results below.

In contrast to Figure 2, the coefficient plots in Figure 3, for non-income related measures, exhibit no obvious "misperception gradient" for any of the reference groups. This indicates that the information about relative position has little effect on the dimensions of well-being that are not directly related to income. We interpret this as evidence that the enjoyment a worker derives from her job, as opposed to the compensation for that job, ought not depend on relative income rank.

**Econometric analysis.** Table 2 reports estimated coefficients from Equation (1) for income

[23]The correlations of misperceptions partly reflect correlations among actual ranks, which are significant and substantial (panel b in Figure C2 for correlations between beliefs).




(a) Education

(b) Occupation

(c) Municipality

(d) Age

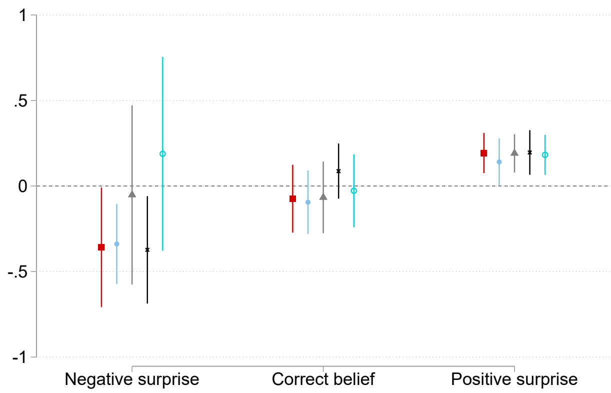
(e) National

Figure 1. . Beliefs about disposable income rank and actual rank in reference groups

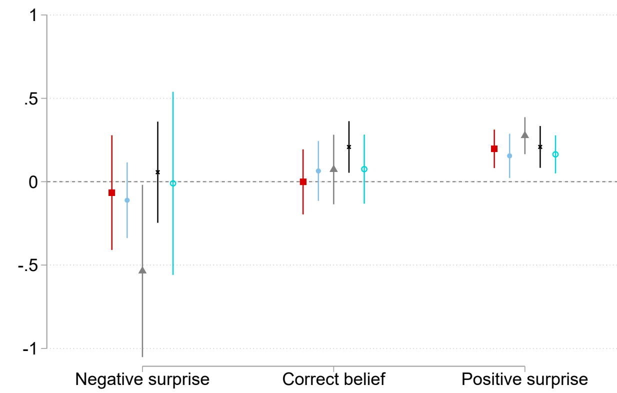
*Notes:*: Distributions of beliefs about disposable income ranks (blue) and actual ranks (gray) in reference groups a) Education, b) Occupation, c) Municipality, d) Age and e) National. Perceived rank is elicited in the belief elicitation section of the survey (see survey screens at the link). Actual rank is based on register data (variable "kturaha") provided by Statistics Finland. The actual rank in reference group *National* refers to the individual's rank among all adult (aged 18 or older) Finns who had non-zero income in 2018. The other reference groups are subsets of *National*, such that e.g. the rank in *Education* refers to the individual's rank among all adult Finns who had non-zero income in 2018, and who had the same highest level of education in 2018. The figures use data from the full survey sample.

related well-being. This formalizes the results in panel (a) of Figure 2, and provides further confirmation of our primary finding.[24] As shown in the first column of Table 2, for example, receiving information about rank in the age cohort affects income satisfaction in the predicted direction.

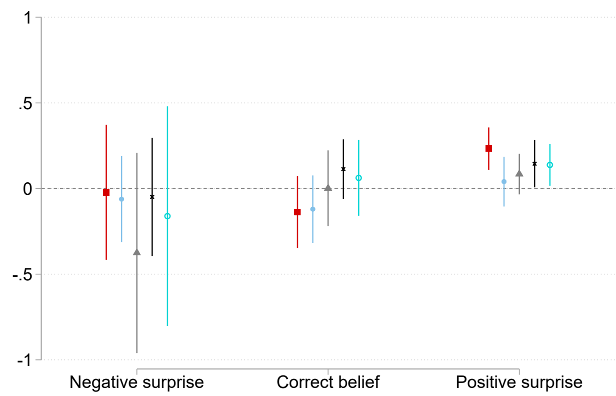
[24]The graphs are based on a specification without control variables (only misperception (bin), treatment, and their interaction are included). The specification reported in Table 2 includes income rank as a control. In Appendix C, we report results of a specification without income rank but including various demographic controls. Section II.E conducts a comprehensive specification curve robustness analysis.




(a) Satisfaction with own disposable income




(b) Fairness of own disposable income




(c) Wage satisfaction

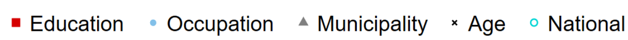



Figure 2. . Income related effects

*Notes:* Each bar plots the difference (95% confidence interval) of satisfaction with disposable income (panel a), perceived fairness of income (panel b) and wage satisfaction (panel c) between the respondents who see their position in the corresponding reference group and the respondents in the control group who do not see their position. The dependent variable is standardized by subtracting the control group mean from each observation and then dividing by the control group standard deviation. "Negative surprise" refers to those who overestimate their income rank by more than 10 percentage points, "positive surprise" to those who underestimate their rank by more than 10 percentage points and "correct belief" to those whose assessment of their position is less than or equal to 10 percentage points away from the true position in absolute terms.

The estimate of $\beta_3$ implies that when a respondent believes her income rank among people of her age is 10 percentage points lower than her actual position, informing her about the actual position would increase her income satisfaction by about 0.11 standard deviations, $\hat{\beta}_2 + (-0.1)\hat{\beta}_3$.[25]

The estimated coefficients of Treatment×Misperception ($\beta_3$) in the top panel in Table 2 are all negative and significant except in the NATIONAL treatment. Further, these coefficients have the opposite sign - and are in some cases close in magnitude to - the estimated Misperception ($\beta_1$) coefficients, which suggests that information provision is indeed an "antedote," partial or sometimes almost full, to the original misbelief. In different terms, this implies that when the coefficients do

[25]Misperception is defined as the difference between belief and actual position. A misperception of 0.01 means that the believed rank is 1 percentage point higher than the actual position.

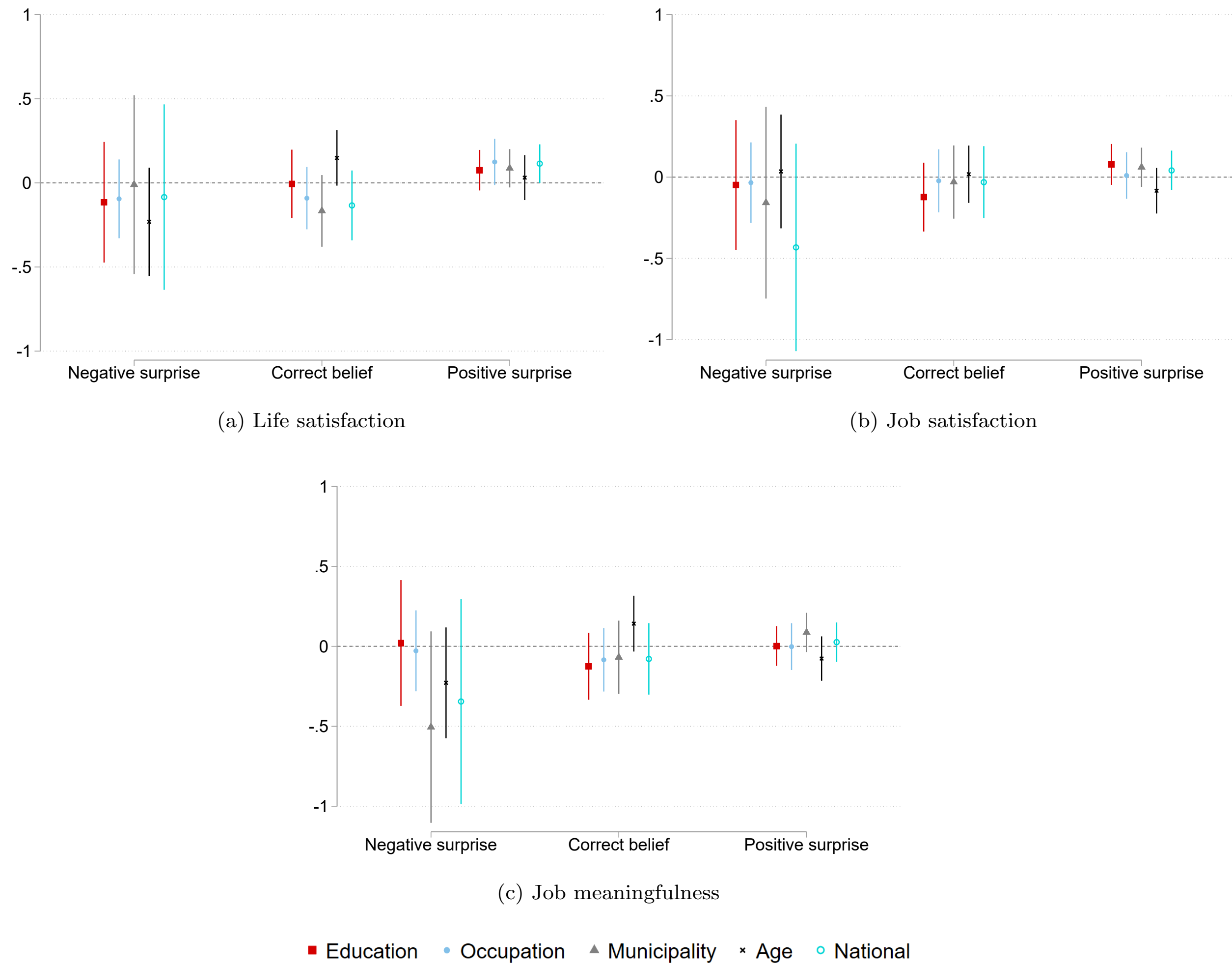



Figure 3. . Non-income related effects

*Notes:* Each bar plots the difference (95% confidence interval) of life satisfaction (panel a), job satisfaction (panel b), and perceived meaningfulness of job (panel c) between the respondents who see their position in the corresponding reference group and the respondents in the control group who do not see their position. The dependent variable is standardized by subtracting the control group mean from each observation and then dividing by the control group standard deviation. "Negative surprise" refers to those who overestimate their income rank by more than 10 percentage points, "positive surprise" to those who underestimate their rank by more than 10 percentage points and "correct belief" to those whose assessment of their position is less than or equal to 10 percentage points away from the true position in absolute terms.

offset, misbelief no longer predicts satisfaction. Our key results for the other income-related well-being measures, reported in the lower panels of Table 2, are qualitatively similar to those for income satisfaction.

Table 2 also provides further support for the finding that there is no pure treatment effect: The estimated effect of information for those with correct beliefs (estimate of $\beta_2$) is in most cases zero.[26]

To analyze potential asymmetries in the effects of information, we have also estimated a spline specification that allows for separate $\beta_3$ coefficients for negative and positive surprises. The results

[26]There are three instances where we observe a pure treatment effect: the effect of income rank in the national distribution on satisfaction with income (significant at 5% level), the perception of fairness of own disposable income (significant at 1% level) in the age distribution, and the effect of national income rank on wage satisfaction (significant at 1% level, see Table 2).

are reported in Table C10 in the Appendix. With few exceptions, we cannot reject symmetry. [27]

We next turn to the broader measures of well-being, the most common of which is general life satisfaction. The estimates reported in Table 3 indicate that the estimated Treatment$\times$Misperception ($\beta_3$) coefficients for life satisfaction are smaller than those for income satisfaction, and mostly statistically insignificant. [28] It should be noted though that the effects are quite imprecisely estimated: While the point estimates of $\beta_3$ for life satisfaction are in all reference groups smaller than the corresponding estimates for income satisfaction, the confidence intervals are partially overlapping; in those instances, we cannot rule out effects of a similar magnitude.

Results for job satisfaction and job meaningfulness are qualitatively similar to those for life satisfaction (see Figure 3 and Table 3).

To provide additional perspective on these results, we also followed reviewers' suggestions and carried out three further analyses that corroborate our main results. First, to increase power, we have estimated a joint regression where the treatment arms are pooled. A challenge with this specification is that it is not clear how to define the benchmark misperception for those in the control group, and what the correct interpretation of the treatment effect is; that is, this analysis does not easily lend itself to the intuitive interpretation we discussed in connection with Table 2, where initial misperception in a given reference group is countered by the treatment with information on the magnitude of that misperception. To provide one though admittedly imperfect solution, we redefine both rank and misperception to be equal to their corresponding averages across all five reference groups for those in the control arm, but use the rank and misbelief for the relevant treatment distribution otherwise. Table C3 in the appendix reports the results for a specification that is inspired by Equation (1) in the main text: well-being measures are regressed on a single treatment indicator, modified rank and misbelief, and the interaction of treatment and misbelief. The results are similar to the main specification in that the effects of rank information on income satisfaction and fairness are sizable and significant. In this specification, we find a significant effect also on life satisfaction, even though the point estimate remains smaller in magnitude than the estimates for the income-related measures.

Second, we have re-estimated our main specification for two consolidated outcomes, namely, the first principal components of the income-related and non-income-related measures of subjective well-being. The results, reported in Tables C4 and C5 in the appendix, are reassuring. Information about rank is a significant causal determinant of the consolidated income-based measure of well-being for all narrow reference groups with strongest and most significant effect in education. However, the effect is weak and insignificant in the national reference group. Moreover, none of the information treatments is significant for our non-income-based consolidated measure.

Third, because both the order of the first three question blocks and the order of questions within blocks were randomized, we can look for the presence of order effects in our data. This is important for two related reasons. First, given our topic and the one-time continuous engagement embodied in our design, there is reason to be concerned about priming and experimenter demand effects. Second, in the absence of a follow-up survey of reported well-being, the exogenous variation in question order affords a modest test of persistence. To this end, Tables C6 and C7 in the appendix report main treatment effects - the coefficient $\beta_3$ in our main specification - for all feasible "positions" of the well-being block for all well-being measures and all reference populations. There is no discernible pattern and, in particular, no evidence that the treatment effect diminishes when the relevant outcome questions are asked later.

[27] For a full set of results, see Table C10, and Tables C11, C12, C13, C14 and C15 for the estimation results of other outcomes.

[28] Despite standardizing outcomes to a standard deviation of 1, comparing effect sizes across different outcomes should be taken with a grain of salt due to differences in scales.

Table 2—. OLS results for the effect of income rank information on income related well-being

| | (1) Age | (2) Municipality | (3) Education | (4) Occupation | (5) National |
|---|---|---|---|---|---|
| **Satisfaction with own disposable income** | | | | | |
| Treatment | 0.025 | -0.010 | -0.120 | -0.081 | 0.176* |
| | [-0.084,0.134] | [-0.152,0.133] | [-0.244,0.005] | [-0.184,0.023] | [0.018,0.333] |
| | (0.655) | (0.896) | (0.059) | (0.127) | (0.029) |
| Misperception | 1.267*** | 1.575*** | 1.780*** | 1.365*** | 1.388*** |
| | [0.878,1.656] | [1.180,1.970] | [1.446,2.115] | [1.046,1.684] | [0.986,1.790] |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Misperception × Treatment | -0.829** | -0.740** | -0.987*** | -0.657** | 0.046 |
| | [-1.330,-0.328] | [-1.262,-0.218] | [-1.435,-0.539] | [-1.051,-0.264] | [-0.528,0.621] |
| | (0.001) | (0.005) | (0.000) | (0.001) | (0.874) |
| Bonferroni corrected p-value | (0.012) | (0.055) | (0.000) | (0.011) | (1.000) |
| Observations | 1521 | 1501 | 1519 | 1505 | 1498 |
| **Perceived fairness of own disposable income** | | | | | |
| Treatment | 0.166** | 0.039 | -0.032 | 0.049 | 0.111 |
| | [0.055,0.276] | [-0.114,0.193] | [-0.159,0.095] | [-0.054,0.152] | [-0.043,0.264] |
| | (0.003) | (0.615) | (0.618) | (0.353) | (0.157) |
| Misperception | 0.734*** | 0.861*** | 1.180*** | 0.943*** | 0.720*** |
| | [0.334,1.134] | [0.448,1.273] | [0.831,1.528] | [0.624,1.262] | [0.304,1.137] |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) |
| Misperception × Treatment | -0.321 | -0.811** | -0.781*** | -0.314 | -0.187 |
| | [-0.817,0.175] | [-1.358,-0.265] | [-1.229,-0.333] | [-0.693,0.065] | [-0.747,0.372] |
| | (0.204) | (0.004) | (0.001) | (0.105) | (0.511) |
| Bonferroni corrected p-value | (1.000) | (0.037) | (0.006) | (1.000) | (1.000) |
| Observations | 1521 | 1501 | 1519 | 1505 | 1498 |
| **Wage satisfaction** | | | | | |
| Treatment | 0.102 | 0.059 | -0.057 | -0.044 | 0.244** |
| | [-0.016,0.221] | [-0.093,0.211] | [-0.196,0.082] | [-0.159,0.070] | [0.083,0.405] |
| | (0.091) | (0.444) | (0.421) | (0.447) | (0.003) |
| Misperception | 1.203*** | 1.247*** | 1.586*** | 1.165*** | 1.137*** |
| | [0.806,1.601] | [0.835,1.658] | [1.230,1.942] | [0.794,1.535] | [0.705,1.569] |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Misperception × Treatment | -0.291 | -0.084 | -0.835*** | -0.290 | 0.454 |
| | [-0.823,0.241] | [-0.633,0.466] | [-1.314,-0.356] | [-0.724,0.144] | [-0.141,1.049] |
| | (0.284) | (0.765) | (0.001) | (0.190) | (0.134) |
| Bonferroni corrected p-value | (1.000) | (1.000) | (0.006) | (1.000) | (1.000) |
| Observations | 1413 | 1398 | 1400 | 1400 | 1393 |

OLS regressions using robust standard errors with 95% confidence intervals in brackets and unadjusted p-values in parentheses estimating the effects of income rank information. The dependent variables are standardized by subtracting the control group mean from each observation and then dividing by the control group standard deviation. The dependent variable perceived fairness, measured with a slider (0: Unfairly low, 50: Fair, 100: Unfairly high) is recoded as 50 - | slider value - 50 | to reflect range from Unfair to Fair. Misperception is defined as belief minus rank, and the difference in percentage points is divided by 100, so a misperception of 0.01 means that the believed rank is 1 percentage point higher than the actual position. Rank in the reference group corresponding to the treatment was used as a control variable but omitted from the table. Reported Bonferroni corrected p-values adjusted to the number of pairwise tests (10) between treatments concern Misperception × Treatment.

Table 3—. OLS results for the effect of income rank information on non-income related well-being

| | (1) Age | (2) Municipality | (3) Education | (4) Occupation | (5) National |
|---|---|---|---|---|---|
| **Life satisfaction** | | | | | |
| Treatment | 0.009 | 0.010 | -0.033 | -0.030 | 0.066 |
| | [-0.113,0.130] | [-0.140,0.160] | [-0.163,0.097] | [-0.143,0.082] | [-0.086,0.218] |
| | (0.887) | (0.895) | (0.617) | (0.594) | (0.392) |
| Misperception | 0.843*** | 0.855*** | 1.099*** | 0.661*** | 0.817*** |
| | [0.448,1.237] | [0.459,1.250] | [0.743,1.454] | [0.334,0.988] | [0.420,1.213] |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Misperception × Treatment | -0.400 | -0.144 | -0.360 | -0.468* | -0.016 |
| | [-0.921,0.122] | [-0.681,0.393] | [-0.807,0.087] | [-0.848,-0.089] | [-0.552,0.519] |
| | (0.133) | (0.598) | (0.114) | (0.016) | (0.953) |
| Bonferroni corrected p-value | (1.000) | (1.000) | (1.000) | (0.156) | (1.000) |
| Observations | 1521 | 1501 | 1519 | 1505 | 1498 |
| **Job satisfaction** | | | | | |
| Treatment | -0.011 | 0.125 | -0.035 | 0.004 | 0.034 |
| | [-0.135,0.112] | [-0.043,0.293] | [-0.179,0.108] | [-0.117,0.124] | [-0.132,0.200] |
| | (0.856) | (0.145) | (0.627) | (0.952) | (0.689) |
| Misperception | 0.475* | 0.129 | 0.558** | 0.185 | 0.203 |
| | [0.055,0.896] | [-0.320,0.577] | [0.173,0.944] | [-0.189,0.559] | [-0.244,0.651] |
| | (0.027) | (0.573) | (0.005) | (0.332) | (0.373) |
| Misperception × Treatment | 0.174 | 0.369 | -0.269 | 0.054 | 0.054 |
| | [-0.391,0.740] | [-0.226,0.964] | [-0.769,0.231] | [-0.373,0.480] | [-0.543,0.652] |
| | (0.546) | (0.224) | (0.291) | (0.804) | (0.858) |
| Bonferroni corrected p-value | (1.000) | (1.000) | (1.000) | (1.000) | (1.000) |
| Observations | 1413 | 1398 | 1400 | 1400 | 1393 |
| **Job meaningfulness** | | | | | |
| Treatment | 0.007 | 0.067 | -0.048 | -0.048 | 0.015 |
| | [-0.117,0.130] | [-0.100,0.235] | [-0.189,0.094] | [-0.171,0.075] | [-0.152,0.183] |
| | (0.915) | (0.430) | (0.508) | (0.443) | (0.856) |
| Misperception | 0.193 | 0.136 | 0.306 | -0.184 | 0.139 |
| | [-0.209,0.595] | [-0.299,0.572] | [-0.062,0.675] | [-0.532,0.163] | [-0.292,0.570] |
| | (0.346) | (0.539) | (0.103) | (0.299) | (0.527) |
| Misperception × Treatment | 0.127 | 0.130 | -0.110 | -0.149 | 0.072 |
| | [-0.428,0.682] | [-0.467,0.726] | [-0.579,0.358] | [-0.564,0.266] | [-0.520,0.664] |
| | (0.654) | (0.670) | (0.644) | (0.481) | (0.812) |
| Bonferroni corrected p-value | (1.000) | (1.000) | (1.000) | (1.000) | (1.000) |
| Observations | 1413 | 1398 | 1400 | 1400 | 1393 |

*Notes:* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. OLS regressions using robust standard errors with 95% confidence intervals in brackets and unadjusted p-values in parentheses estimating the effects of income rank information. The dependent variables are standardized by subtracting the control group mean from each observation and then dividing by the control group standard deviation. Misperception is defined as belief minus rank, and the difference in percentage points is divided by 100, so a misperception of 0.01 means that the believed rank is 1 percentage point higher than the actual position. Rank in the reference group corresponding to the treatment was used as a control variable but omitted from the table. Reported Bonferroni corrected p-values adjusted to the number of pairwise tests (10) between treatments concern Misperception × Treatment.

### C. Effects of relative income information for different reference groups

The treatments reveal rank information in reference groups. We did not pre-register hypotheses concerning treatment differences as we did not have strong priors. The design reveals, however, that they are primary research questions with scant prior causal evidence.

Looking at the effects of rank information on satisfaction with disposable income (Table 2), for example, rank in the national income distribution seems to matter less to people than rank in the other reference groups: the estimated effect of rank in the national income distribution (the coefficient of Treatment×Misperception, $\beta_3$, in the last column) is an order of magnitude smaller than in the other reference groups, and not statistically distinguishable from zero. Other income-related satisfaction measures provide qualitatively similar evidence, as the rank in the national income distribution is consistently insignificant in both the statistical and economic senses.

Table 4 provides results from formal tests of whether information on rank in different reference groups affects well-being differently.[29] We compare the treatment effects across the different reference groups by testing the equality of the $\beta_3$-coefficients in regressions run separately for each reference group (our baseline results, reported in Table 2). We follow the baseline specification and apply seemingly unrelated regressions (Weesie, 1999). In practice this involves stacking the data (duplicating the control group and pairing each treatment group with the same control group) and then using clustered robust standard errors to account for the interdependent samples. More details can be found in the notes of Table 4.[30] The last column of Table 4 shows the results of a joint test of whether the effects of relative income information are the same in all reference groups. The other columns provide pairwise tests across the different reference groups. We focus on the test of joint equality in the last column since we are interested mostly in whether the income rank information of different reference groups affects well-being in the same way or not.

The results in Table 4 suggest that the treatment effects are not equal between the reference groups. The $p$-value of the test against joint equality ($\beta_3$-coefficients equal across the five regressions) is 0.06 for fairness of own income, 0.02 for satisfaction with disposable income, and 0.01 for wage satisfaction. The pairwise comparisons indicate that the differences are driven by the effects for the national reference group being different from (smaller than) the more circumscribed reference groups.

Our results therefore indicate that information about rank in more circumscribed reference groups has stronger effects on well-being than information about national rank. In this sense, more circumscribed reference groups are more important.

Let us next consider the possible effects of information spillovers on the interpretation of our results. Consistent with the principles of controlled experimentation, we chose to reveal a single distribution rank to participants in information conditions. Individuals might nevertheless understand that distributional ranks are correlated, in which case information about one might cause beliefs about others to be updated, too. Belief spillovers, even if sizable, *do not* affect the interpretation of our results as providing unbiased estimates of the causal impact of *information about rank* in different reference groups. Spillovers would, however, compromise an alternative interpretation of our results as claims about the relative importance of rank itself, as opposed to information about rank.

As it turns out, there is reason to believe that the spillovers are idiosyncratic and modest in size. Consider first the extreme case in which individuals "correct" *all* their rank beliefs by the same amount: that is, discovering that one's place in the national income distribution is ten percentage

[29] We discuss the treatment effect comparison in detail only in the context of the income-related satisfaction measures, where we find clear effects overall. For the measures related to general satisfaction, we did not find significant effects overall and we also find no differences between the reference groups.

[30] An alternative would be to run regressions with interactions between $\beta_3$ and the different treatments. However, this becomes problematic as we would need to control for all the different misperceptions, measured for all reference groups, simultaneously in one regression. Given that the misperceptions in different reference groups are highly collinear, such an analysis is not feasible.

points higher than first believed will induce a rank belief update by the same amount in the national distribution, the municipal distribution (a plausible response, given their correlation), and all other distributions. (It is not essential that the update also equal 10 percentage points.) In this case, it would not matter to which treatment group the individual was randomized, and the estimated interaction coefficients would be the same, and equal to the synthesized treatment of all five reference groups. A cursory glance at our results tables reveals this is far from the case.

From a more general perspective, one might expect that where pairwise correlations are strong, there are also substantial information spillovers. Beliefs about municipal and national ranks, for example, are highly correlated (see panel (b) in Figure C2). Thus, one might expect that a significant effect on some measure of welfare in the municipal rank condition would be associated with a significant effect in the national rank condition. This, too, is not what we observe: indeed, the relative importance of municipal standing—and the relative unimportance of national standing—is a central theme of our results.

Thus, it seems reasonable to conclude that there was likely limited updating of non-treatment beliefs, and that the treatment effects of e.g. information about municipal rank do not reflect much spillover. The explanation might include bounds on cognition: misperceptions are less correlated than beliefs (0.79 for the municipal-national pair, for example, as opposed to 0.85, see panels (b) and (c) in Figure C2), which is consistent with the view that individuals are more confident/better informed about some ranks than others.

Last, there is another sense in which information could spill over. Individuals who know their own absolute income and learn something new about their rank might update their beliefs about *overall* inequality, generating changes in well-being for those who are inequality averse in the broad sense. It seems reasonable, and in line with previous evidence (Epper et al., 2024), that such individuals would also reconsider their support for redistributive policies. As we report in Table C9 in the appendix, however, no such effect is observed.

Table 4—. Test equality of the coefficient of Treatment×Misperception across the reference groups

| | **Income satisfaction** | **Fairness of income** | **Wage satisfaction** |
|---|---|---|---|
| National vs. Age | -0.838 vs. -0.027 | -0.318 vs. -0.170 | -0.312 vs. 0.322 |
| p-value | **0.006** | 0.608 | **0.043** |
| adjusted p-value | 0.050 | 1.000 | 0.299 |
| National vs. Municipality | -0.792 vs. -0.027 | -0.838 vs. -0.170 | -0.136 vs. 0.322 |
| p-value | **0.010** | **0.027** | 0.138 |
| adjusted p-value | 0.077 | 0.272 | 0.553 |
| National vs. Education | -0.921 vs. -0.027 | -0.737 vs. -0.170 | -0.769 vs. 0.322 |
| p-value | **0.002** | 0.053 | **0.000** |
| adjusted p-value | **0.023** | 0.425 | **0.004** |
| National vs. Occupation | -0.628 vs. -0.027 | -0.292 vs. -0.170 | -0.239 vs. 0.322 |
| p-value | **0.038** | 0.674 | 0.062 |
| adjusted p-value | 0.264 | 1.000 | 0.371 |
| Age vs. Municipality | -0.838 vs. -0.792 | -0.318 vs. -0.838 | -0.312 vs. -0.136 |
| p-value | 0.868 | 0.072 | 0.564 |
| adjusted p-value | 0.868 | 0.433 | 1.000 |
| Age vs. Education | -0.838 vs. -0.921 | -0.318 vs. -0.737 | -0.312 vs. -0.769 |
| p-value | 0.757 | 0.119 | 0.116 |
| adjusted p-value | 1.000 | 0.593 | 0.580 |
| Age vs. Occupation | -0.838 vs. -0.628 | -0.318 vs. -0.292 | -0.312 vs. -0.239 |
| p-value | 0.414 | 0.920 | 0.790 |
| adjusted p-value | 1.000 | 0.920 | 0.790 |
| Municipality vs. Education | -0.792 vs. -0.921 | -0.838 vs. -0.737 | -0.136 vs. -0.769 |
| p-value | 0.639 | 0.725 | **0.031** |
| adjusted p-value | 1.000 | 1.000 | 0.280 |
| Municipality vs. Occupation | -0.792 vs. -0.628 | -0.838 vs. -0.292 | -0.136 vs. -0.239 |
| p-value | 0.541 | 0.052 | 0.714 |
| adjusted p-value | 1.000 | 0.472 | 1.000 |
| Education vs. Occupation | -0.921 vs. -0.628 | -0.737 vs. -0.292 | -0.769 vs. -0.239 |
| p-value | 0.221 | 0.062 | **0.035** |
| adjusted p-value | 1.000 | 0.434 | 0.279 |
| Joint equality | **0.021** | 0.061 | **0.011** |

*Notes:* The null hypothesis of the pairwise tests is that the coefficients of Treatment×Misperception are equal between a pair of treated groups. The null hypothesis of the joint test of equality is that the coefficient of Treatment×Misperception is equal across the five regressions of the corresponding reference groups. For the pairwise tests, the $p$-value is unadjusted for a single test and the adjusted $p$-value is Holm's adjusted $p$-value for the 10 pairwise tests of each outcome. The comparisons with $p < 0.05$ are highlighted with $p$-value in bold font. The sizes of the estimates are shown for each outcome above the $p$-values. The estimates of the coefficients of Treatment×Misperception are from the regression results in Table 2. The test procedure: 1) Stack the data by duplicating the control group 4 times, so that there become 'five' control groups. 2) Each (identical) control group is paired with one treated group and used in one regression for each outcome. 3) The method, Seemingly Unrelated Regression (Weesie, 1999), is used to combine the five regression results and produce a simultaneous covariance matrix. 4) Such stacking means the five control groups are the same and the five regressions have inter-dependent samples. To account for the problem of a non-zero covariance between the estimators of the regressions, cluster robust standard error (cluster at subject-level) is used.

### D. *Additional results*

Together, the results in the last two subsections show that the acquisition of information about relative position affects various income-related measures of well-being, with the curious exception of rank in the national income distribution. We also find, however, that overall life satisfaction is less sensitive to such information. We conjecture that life satisfaction is both multi-dimensional and long term; and it is in fact natural that relative position of income affects the income-related aspects of satisfaction more than other aspects.

In this section, we report results from two sets of additional analyses that broaden and strengthen our main findings. First, we report results from a separate treatment, in which the reference group was not exogenously assigned, but rather chosen by the respondent. This approach more closely resembles real-life information acquisition, and the results from this treatment provide complementary evidence about which sort of rank information matters most to people. Second, while the main focus of our paper is on the effects of rank information on different dimensions of subjective well-being, it is important to ask about the *real effects* of our treatments. While our initial design cannot provide definitive answers, we share evidence that such effects are present.

**Choice of reference group information.** In our CHOICE treatment, individuals were invited to choose what rank information to acquire. Respondent were then informed of their rank in the chosen reference group.

Figure 4 reports the frequencies of choices, and several themes emerge. Consistent with the results described above, reference groups for which the estimated treatment effect of rank information was small were requested least often. In particular, we observed that across outcomes, information about national rank did not matter much for well-being. In a similar vein, fewer than 6% of respondents in the CHOICE treatment wanted to learn their position in the national distribution. The circumscribed reference groups were more popular choices, and at the other extreme, 45% chose to learn about occupational rank.

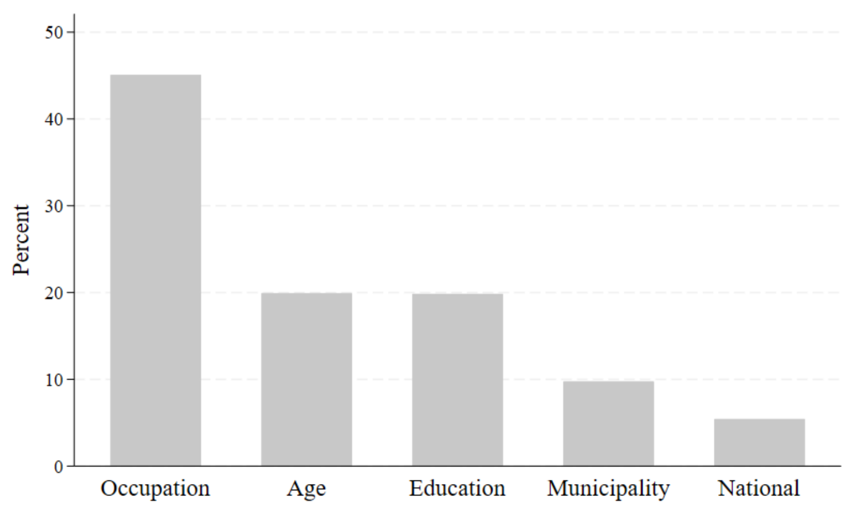



Figure 4. . Chosen reference group information

*Notes:* Figure displays the percentage of participants in treatment CHOICE choosing to learn their income rank in a given reference distribution. 1800 observations.

**Effects on real outcomes.** Let us next turn to the real effects of rank information, and report on two quite different findings, both of which support the view that our interventions had real consequences.

The first exploits a feature of our design: Recall that at the end of our survey, respondents were able to spend some or all of the compensation for participation on a charitable donation, a voluntary tax contribution and/or lottery tickets. The purpose was randomly drawn at the end of the survey, and thus we have three independent behavioral measures. We can therefore examine the effects of our information treatments on these real choices within our survey.

Table 5—. Effect of income rank information on donations to charity

| | (1) Age | (2) Municipality | (3) Education | (4) Occupation | (5) National |
|---|---|---|---|---|---|
| **Extensive margin** | | | | | |
| Treatment | -0.114 | -0.019 | -0.113 | -0.139 | -0.035 |
| | (0.146) | (0.185) | (0.169) | (0.137) | (0.191) |
| Misperception | -0.614 | -0.256 | -0.091 | -0.540 | 0.290 |
| | (0.517) | (0.461) | (0.463) | (0.405) | (0.520) |
| Treatment × Misperception | 0.312 | 0.219 | -0.034 | -0.127 | 0.318 |
| | (0.678) | (0.630) | (0.590) | (0.486) | (0.669) |
| Rank | 0.963** | 1.181*** | 0.395 | 0.027 | 1.753*** |
| | (0.297) | (0.340) | (0.306) | (0.302) | (0.344) |
| Constant | 0.741*** | 0.467 | 1.112*** | 1.306*** | 0.165 |
| | (0.183) | (0.256) | (0.202) | (0.175) | (0.260) |
| **Intensive margin** | | | | | |
| Treatment | -0.269 | -0.561 | -0.072 | 0.245 | -0.286 |
| | (0.320) | (0.399) | (0.349) | (0.294) | (0.405) |
| Misperception | 3.390*** | 3.249*** | 3.397*** | 2.351** | 3.964*** |
| | (1.013) | (0.958) | (0.895) | (0.809) | (1.002) |
| Treatment × Misperception | -4.162** | -2.944* | -2.015 | -1.046 | -3.499* |
| | (1.416) | (1.378) | (1.181) | (0.973) | (1.440) |
| Rank | 3.075*** | 3.349*** | 1.720** | 1.973** | 3.991*** |
| | (0.604) | (0.728) | (0.630) | (0.627) | (0.731) |
| Constant | 8.954*** | 8.621*** | 9.976*** | 9.623*** | 8.263*** |
| | (0.396) | (0.567) | (0.416) | (0.362) | (0.583) |
| Observations | 1508 | 1492 | 1511 | 1495 | 1490 |

*Notes:* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Logit and OLS regressions with standard errors in parentheses estimating the effects of income rank information provision on donating money (0, 5, 10, 15 EUR) to a charity using a two-part model. The top panel reports estimates from a logit regression, where the outcome is a dummy for whether the participant donated a positive amount (i.e. the extensive margin of giving). The bottom panel reports estimates from an OLS regression, where the outcome is the euro amount of donations, conditional on giving a positive amount (i.e. the intensive margin of giving). The misperception is belief minus rank in the reference group corresponding to treatment, and the difference in percentage points is divided by 100, so a misperception of 0.01 means that the believed rank is 1 percentage point higher than the actual rank. Rank is the respondent's actual rank divided by 100 in the reference group corresponding to treatment.

Table 5 reports the estimates for a two-part model of charitable donation with the same right-hand side variables as our main specification (treatment, misperception, their interaction, and rank) for each of the five reference groups. Starting with the intensive margin of giving, we note that both misperception and actual rank are significant positive predictors of donations, conditional on giving. In the case of age, municipality and national distributions, the effect of rank information is significant and offsets almost exactly the predicted effect of misperception. This is consistent with the existence of a causal influence of rank information: conditional on giving, the discovery that one's status or rank is higher than expected causes donations to rise, and the effect is substantial. The treatment effects in the education and occupation reference groups have the same sign but are somewhat smaller and not significant (at 5% level).

As Table 5 also suggests, however, the same cannot be said about the act of giving itself: rank is sometimes a significant predictor of the extensive margin of giving, but neither misperceptions nor the correction of those misperceptions matter, and the sign patterns are mixed. Our tentative conclusion is that information about position does not cause individuals to donate, but that it does cause givers to adjust their donation up or down, a meaningful real effect.

Tables C16 and C17 in the appendix, on the other hand, report the analogous results for voluntary tax contributions and the purchase of lottery tickets, with mixed results. With one exception, there is little evidence of causal effects on the extensive margin for either sort of spending. In the case of voluntary tax contributions, misperceptions are a significant and positive predictor of the intensive margin, and while the provision of correct information appears to produce partial reversal in all cases, the treatment effects are not significant. There is at best, then, the whisper of a different sort of induced pro-sociality. In the case of lottery ticket purchases, neither misperceptions nor their correction matter on either the extensive or intensive margin.

Table 6—. Effect of income rank information on earned income (log) in 2021

| | (1) Age | (2) Municipality | (3) Education | (4) Occupation | (5) National |
|---|---|---|---|---|---|
| Misperception | -0.842*** | -0.575* | -0.569** | -0.601*** | -0.500 |
| | (0.239) | (0.274) | (0.206) | (0.177) | (0.305) |
| Treatment | 0.062 | 0.090 | -0.017 | 0.070 | -0.133 |
| | (0.076) | (0.110) | (0.089) | (0.064) | (0.119) |
| Misperception × Treatment | 0.244 | 0.393 | 0.170 | 0.556* | -0.412 |
| | (0.360) | (0.392) | (0.289) | (0.233) | (0.424) |
| Constant | -0.110* | -0.126 | -0.115 | -0.081 | -0.110 |
| | (0.055) | (0.079) | (0.064) | (0.050) | (0.082) |
| Observations | 1508 | 1492 | 1511 | 1493 | 1489 |

*Notes:* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. OLS regressions with robust standard errors in parentheses estimating the effects of income rank information provision on earned income in 2021. The dependent variable is earned income consisting of cash salary items, compensation for employment-related costs and in-kind benefits. The dependent variable is standardized by subtracting the control group mean from each observation and then dividing by the control group standard deviation. Treatment is an indicator for being in the respective treatment group. The misperception is belief minus rank in the reference group corresponding to treatment, and the difference in percentage points is divided by 100, so a misperception of 0.01 means that the believed rank is 1 percentage point higher than the actual rank. Data include all individuals who completed the survey.

The second finding regarding real effects of rank information builds on a reviewer's suggestion, and exploits the opportunity to link our experimental data with register data, and follow our respondents through subsequent income and employment registers.[31] The most recent data available for research purposes (as provided by Statistics Finland) are 2021 and 2022 for income data, and 2021 for employment contract data. Utilizing these data constitutes the sternest possible test of real effects, since our participants only completed the survey in mid-2021.

Table 6 indicates that there appears to be a significant causal effect of information provision about occupational rank on standardized log of earned income in 2021. This provides an indication of another important real effect, one that isn't difficult to rationalize: someone who discovered that their standing within their occupation was lower than expected might well ask for a raise or ensure to receive all perks and bonuses, for example. Furthermore, we observe substantial, if insignificant,

[31] Our pre-registration intends to look at real-effects in register data in another project.

effects for the other narrow - that is, all but national - reference groups. The unimportance of the national reference group amplifies a recurring theme of this paper: information about national rank does not affect reports of subjective well-being and so, perhaps not surprisingly, has no obvious real effects. Last, we note, based on Table C20, that all of these effects were muted in 2022.

Finally, we have examined whether our treatments caused individuals to switch jobs. This would be a natural response to disappointment with one's income rank and implied conviction of better options elsewhere (Hirschman and Rothschild, 1973; Card et al., 2012). We define a job switch as an instance where an individual has started a new employment relationship. For this outcome, the latest available data is currently for the year 2021. As our treatments took place in summer 2021, and because job switches are likely not instantaneous, it makes sense to examine effects on behavior towards the end of the year 2021 only. Estimates for the last quarter (October-December 2021) are given in Table C21 in the appendix. The results show that receiving negative news about one's rank in the education distribution caused an increase in the likelihood of switching jobs. However, this evidence is at best suggestive, as the significant effect arises only in one reference group, and the result is not fully robust to different definitions of the time period we look at (say, Sept-Dec 2021 or Nov-Dec 2021; the estimate remains positive but is not significant for these definitions).

### *E. Robustness: Specification curve analysis*

We examine the robustness of our results using specification curve analysis as proposed by Simonsohn, Simmons and Nelson (2020).[32] We focus on the estimate of the coefficient of $Treatment \times Misperception$ ($\beta_3$) and conduct the analysis for each outcome and treatment. Table E1 in the appendix summarizes the variations in model specifications, stemming from three types of analytical decisions: (A) sample restrictions, (B) the definition of misperception, and (C) the choice of covariates.

The four independent sample restrictions which we consider here are excluding the subjects with the largest misperceptions; with incomplete answers; mismatched information between self-reported and registry data; and long response times. Misperception is operationalized in five ways to reflect the variance across studies that have a similar design to ours. For instance, it is categorized as positive, negative and no bias, and converted to three indicators in the model of Karadja, Mollerstrom and Seim (2017); it is defined as percentile in Perez-Truglia (2020), quintiles in Hoy and Mager (2021), and indicators for positive and non-positive values in Hvidberg, Kreiner and Stantcheva (2023). The last row of Table E1 reports the analytical decisions concerning covariates. The treatment dummy is included in our pre-registered main specification, but omitted in some related studies.[33] The omission of the treatment dummy assumes there is no information effect without misperception, which can affect both the size and significance of the coefficient estimate of $Treatment \times Misperception$. The other sets of covariates are the actual rank in the corresponding reference group, demographic control covariates, a set of labor market variables and a set of survey related variables. We bundle the variables and vary the five sets in the specifications. In total, the variation in analytical decisions gives 3840 specifications which we estimate for each treatment and outcome.[34]

As Figure E1a illustrates, the majority of the estimates for the effect of income rank information in educational reference group on perceived fairness of own income are negative, as we observed in

[32]We pre-registered to estimate alternative specifications to check the robustness of the results in the main specification, but not the specification curve analysis in particular.

[33]Some specifications include the indicators for the bins of misperception and their interaction with the treatment dummy, so the treatment dummy is not included for collinearity. Some specification includes only misperception and its interaction with the treatment dummy.

[34]Treatment NATIONAL is an exception as one sample restriction criterion does not apply (third criterion in first row of Table E1). We employed the computational tool developed by Young and Holsteen (2017) to obtain the distributions of the estimates for all outcomes.

main specification (see column 3 in the middle panel in Table 2). Second, the treatment dummy has an obviously large consequential impact on the statistical significance of the estimates: when the treatment dummy is excluded, the $p$-value of many estimates turn from below to above 0.05.

Third, the definition of misperception influences the effect size: the absolute effect size is smaller when the misperception is defined as dummies, as expected. Last, sample restrictions have little impact on the estimates. The other coefficient estimates are depicted in Figures E2-E7 in the appendix.

Table E2 reports the share of significant results out of all specifications, the median effect size, and the Stouffer Z-statistic. The under-the-null distribution of each effect estimate is constructed by shuffling the randomly assigned variable in our design, the treatment dummy. As seen from the inferential specification curves in Figures E1b-E11a in appendix, a large fraction of the effect estimates from the observed sample locate outside the 95% confidence interval of the under-the-null hypothesis.

The null hypotheses are rejected for all the discovered effects in all the joint tests at the 5% level.[35] Therefore, based on the specification curve analysis we conclude that the effects reported in Tables 2 and 3 are all strongly robust.

## III. Discussion

In this section, we explore one important policy implication of our results, namely, the welfare effects of "income transparency" policies, i.e. policies that reveal some information on individual incomes. It has been shown that income transparency has implications for the functioning of labor markets and for tax compliance – see Cullen (2024) for a review. Evidence presented in Reck, Slemrod and Vattø (2022) suggests that social comparisons may be important drivers of the effects of income transparency policies. We add to this discussion by analyzing the well-being implications of interventions where individuals learn their rank, but not absolute incomes or income differences e.g. to some reference point (say, the median for example). While we focus on the immediate, direct effects on subjective well-being, e.g. Cullen and Pakzad-Hurson (2023) have provided evidence of the equilibrium effects of income transparency policies, finding negative effects on equilibrium wages.

Our analysis highlights the importance of the nature of misperceptions for the implications of income transparency policies: The effects of (increased) transparency will hinge on the information content – from the individual's perspective – of the intervention, which on the other hand depends on the nature of initial misperceptions. To provide a benchmark, we first note that if the average misperception was close to zero, and if the individual effects of positive and negative surprises were roughly symmetric, the aggregate welfare effects of transparency would be negligible. Information would only generate transfers of happiness without affecting the aggregate.[36] Within our framework, then, non-negligible aggregate effects are attributable to violations of one or both conditions.

Let us consider the average treatment effect of providing information about rank in each of our reference groups. To do so, we estimate a simple model where each of our well-being measures is regressed (only) on a dummy indicating whether the individual received rank information or not. The coefficient of this dummy, then, incorporates both the implications of the average mag-

[35] The inference results remain the same when we conduct the joint tests with the treatment dummy always included in the specifications. We also conduct the joint tests with the treatment dummy included and varying the 14 covariates individually in the specifications. This makes the number of reasonable specifications amount to around 1 000 000 for each investigated effect. For computational intensity, following the practical solution suggested in Simonsohn, Simmons and Nelson (2020), we choose a random subset of the specifications (at each round of simulation, randomly 45360 out of the 1 000 000) to make statistical inferences. The analyses with the extensive list of specifications also show that all the found effects are robust in all the joint tests.

[36] Admittedly, this argument presupposes an anonymous utilitarian approach to evaluating societal welfare.

nitude of misperceptions regarding the given distribution, as well as the well-being impact of that misperception (per unit).[37]

The results are presented in Table C8. For the sake of the discussion, let us focus on the effects on perceived fairness of own income, where we find the strongest effects. The results show that rank information has positive effects on perceptions of fairness, and qualitatively similar findings are obtained for the other income-related well-being measures. Interestingly, positive treatment effects arise for rank information in the age, municipality, education, and national reference groups. The implied improvements in well-being are substantial, amounting to between 0.13 and 0.20 standard deviations.

What explains these findings? Rank information increases the well-being of pessimists who receive a positive surprise (cf. Figure 2) and typically reduces it for optimists who receive a negative surprise. Importantly, according to our results, pessimists outnumber optimists almost 9 to 1. (Recall that this pessimism, while not universal, is consistent with, for example, the work of Karadja, Mollerstrom and Seim (2017) on Sweden.) Further, we did not find much evidence of asymmetric effects of negative vs. positive surprises at the margin so the key driver behind the aggregate welfare effects indeed relates to the structure of the initial misperceptions that are undone by the transparency policy.

Note also that the effects of rank information in the national reference group are of similar magnitude as in the other reference groups. This finding stands out as apparently in contrast to our main findings where rank information in the national reference group did not appear to matter for well-being. Part of the explanation, again, relates to the nature of misperceptions: people underestimate their rank especially in the national distribution, while the perceptions of rank in the more circumscribed reference groups are somewhat more accurate and/or include a significant fraction of overestimates; cf. Table C1 and Figure 1 in the appendix. Therefore, the information treatment relating to rank in the national income distribution creates more positive surprises as compared to other types of rank information. On the other hand, misbeliefs in the occupational income distribution are much more evenly distributed around zero, which gives rise to both positive and negative surprises - and therefore the average treatment effect of increasing income transparency within the occupational distribution is small (or non-existent), even though learning one's rank may have a significant effect at the individual level.

Overall, therefore, income transparency is welfare-enhancing according to the money-related well-being measures, essentially because it provides large (if one time) benefits to the large number who believe their place in the income distribution was lower than it actually was, and because the acquisition of such knowledge has substantial effects on individual welfare. Consistent with our earlier results, on the other hand, the aggregate effects on life satisfaction are negligible.

## IV. Conclusion

Our study employs a pre-registered information provision experiment to investigate the effects of income rank information in various reference groups on individual well-being. We document, first, that income rank information has causal effects on income-related measures of well-being, such as satisfaction with own income, perceived fairness of own income, and wage satisfaction. Second, we find that the effects of income rank information are much stronger on the income-related well-being measures than on the non-income-related measures, such as life and job satisfaction, and perceived job meaningfulness.

Third, information about income rank in different reference groups affects individual well-being differently. One startling finding is that information on national income rank, which is much

[37]Note therefore that the effects may differ from what we reported in our main analysis: the main analysis reported the implications of *each unit* of misperception i.e. the coefficient ($\beta_3$) from Equation (1). If the magnitude of misperceptions differs between distributions, the aggregate effects may also differ.

emphasized in the literature, has a weak and statistically insignificant effect on all the considered well-being measures at the individual level. Information about income rank in narrower reference groups – among those with the same educational level or those living in the same municipality – have a much stronger effect on individual well-being. Thus, studies focusing on the national reference group might have underestimated the effect of relative income on well-being. We confirm the robustness of our main findings using specification curve analysis. Last, we provide suggestive evidence of real effects of rank information on two types of measures: amount of charitable donations within our survey, as well as wage income measured in register data.

We then analyze the effect on welfare at the aggregate level of unveiling individual income ranks in a specific reference group. This aggregate effect depends both on the distribution of initial misperceptions – whether individuals are positively or negatively surprised by the information they receive – and the (a)symmetry of welfare effects between the pessimists and optimists. In our data, there are far more pessimists than optimists, and by and large, the welfare effects between the two are symmetric. The two channels, together with the substantial effects on individual welfare, lead to a net positive effect on aggregate welfare. Therefore, even though relative income concerns are typically regarded as a negative externality, making relative standing more salient can improve welfare in the aggregate. Nevertheless, this is naturally only one ingredient in a welfare analysis of an income transparency policy where people privately learn their ranks, and a complete analysis would have to take into account further effects e.g. on labor markets or visibility of one's rank to others.

## REFERENCES

**Alesina, Alberto, Rafael Di Tella, and Robert MacCulloch.** 2004. "Inequality and happiness: Are Europeans and Americans different?" Journal of Public Economics, 88(9-10): 2009–2042.

**Almås, Ingvild, Alexander W. Cappelen, and Bertil Tungodden.** 2020. "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" Journal of Political Economy, 128(5): 1753–1788.

**Bartling, Björn.** 2011. "Relative performance or team evaluation? Optimal contracts for other-regarding agents." Journal of Economic Behavior & Organization, 79(3): 183–193.

**Bellemare, Charles, Sabine Kröger, and Arthur Van Soest.** 2008. "Measuring Inequity Aversion in a Heterogeneous Population Using Experimental Decisions and Subjective Probabilities." Econometrica, 76(4): 815–839.

**Bolton, Gary E., and Axel Ockenfels.** 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." American Economic Review, 90(1): 166–193.

**Bond, Timothy N., and Kevin Lang.** 2019. "The Sad Truth about Happiness Scales." Journal of Political Economy, 127(4): 1629–1640.

**Bowles, Samuel, and Wendy Carlin.** 2020. "Inequality as experienced difference: A reformulation of the Gini coefficient." Economics Letters, 186: 108789.

**Boyce, Christopher J., Gordon D.A. Brown, and Simon C. Moore.** 2010. "Money and Happiness: Rank of Income, Not Income, Affects Life Satisfaction." Psychological Science, 21(4): 471–475. PMID: 20424085.

**Brown, Gordon D. A., Jonathan Gardner, Andrew J. Oswald, and Jing Qian.** 2008. "Does Wage Rank Affect Employees' Well-being?" Industrial Relations: A Journal of Economy and Society, 47(3): 355–389.

**Bublitz, Elisabeth.** 2020. "Misperceptions of income distributions: cross-country evidence from a randomized survey experiment." Socio-Economic Review, 20(2): 435–462.

**Bursztyn, Leonardo, and Robert Jensen.** 2017. "Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure." Annual Review of Economics, 9(1): 131–153.

**Cappelen, Alexander W., Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden.** 2007. "The Pluralism of Fairness Ideals: An Experimental Approach." American Economic Review, 97(3): 818–827.

**Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez.** 2012. "Inequality at work: The effect of peer salaries on job satisfaction." American Economic Review, 102(6): 2981–3003.

**Chen, Daniel L, Martin Schonger, and Chris Wickens.** 2016. "oTree—An open-source platform for laboratory, online, and field experiments." Journal of Behavioral and Experimental Finance, 9: 88–97.

**Clark, Andrew E., and Claudia Senik.** 2010. "Who Compares to Whom? The Anatomy of Income Comparisons in Europe*." The Economic Journal, 120(544): 573–594.

**Clark, Andrew E., and Conchita D'Ambrosio.** 2015. "Chapter 13 - Attitudes to Income Inequality: Experimental and Survey Evidence." In Handbook of Income Distribution. Vol. 2 of Handbook of Income Distribution, , ed. Anthony B. Atkinson and François Bourguignon, 1147–1208. Elsevier.

**Clark, Andrew E, Niels Westergård-Nielsen, and Nicolai Kristensen.** 2009. "Economic satisfaction and income rank in small neighbourhoods." Journal of the European Economic Association, 7(2-3): 519–527.

**Clark, Andrew E, Paul Frijters, and Michael A Shields.** 2008. "Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles." Journal of Economic literature, 46(1): 95–144.

**Corneo, Giacomo, and Olivier Jeanne.** 1997. "Conspicuous consumption, snobbism and conformism." Journal of Public Economics, 66(1): 55–71.

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

**Cullen, Zoë.** 2024. "Is pay transparency good?" Journal of Economic Perspectives, 38(1): 153–180.

**Cullen, Zoë, and Ricardo Perez-Truglia.** 2022. "How much does your boss make? The effects of salary comparisons." Journal of Political Economy, 130(3): 766–822.

**Cullen, Zoë B, and Bobak Pakzad-Hurson.** 2023. "Equilibrium effects of pay transparency." Econometrica, 91(3): 765–802.

**Di Tella, Rafael, John Haisken-De New, and Robert MacCulloch.** 2010. "Happiness adaptation to income and to status in an individual panel." Journal of Economic Behavior & Organization, 76(3): 834–852.

**Dube, Arindrajit, Laura Giuliano, and Jonathan Leonard.** 2019. "Fairness and Frictions: The Impact of Unequal Raises on Quit Behavior." American Economic Review, 109(2): 620–63.

**Duesenberry, James S.** 1949. Income, Saving, and the Theory of Consumer Behavior. Cambridge, Mass.:Harvard University Press.

**Easterlin, Richard A.** 1974. "Does economic growth improve the human lot? Some empirical evidence." In Nations and households in economic growth. 89–125. Elsevier.

**Engelhardt, Carina, and Andreas Wagener.** 2018. "What do Germans think and know about income inequality? A survey experiment." Socio-Economic Review, 16(4): 743–767.

**Englmaier, Florian, and Achim Wambach.** 2010. "Optimal incentive contracts under inequity aversion." Games and economic Behavior, 69(2): 312–328.

**Epper, Thomas F, Ernst Fehr, Claus Thustrup Kreiner, Søren Leth-Petersen, Isabel Skak Olufsen, and Peer Ebbesen Skov.** 2024. "Inequality aversion predicts support for public and private redistribution." Proceedings of the National Academy of Sciences, 121(39): e2401445121.

**Fehr, Dietmar, Johanna Mollerstrom, and Ricardo Perez-Truglia.** 2022. "Your Place in the World: Relative Income and Global Inequality." American Economic Journal: Economic Policy, 14(4): 232–68.

**Fehr, Ernst, and Gary Charness.** 2023. "Social preferences: fundamental characteristics and economic consequences."

**Fehr, Ernst, and Klaus M. Schmidt.** 1999. "A Theory of Fairness, Competition, and Cooperation." The Quarterly Journal of Economics, 114(3): 817–868.

**Fehr, Ernst, and Klaus M. Schmidt.** 2006. "The economics of fairness, reciprocity and altruism–experimental evidence and new theories." Handbook of the economics of giving, altruism and reciprocity, 1: 615–691.

**Fernández-Albertos, José, and Alexander Kuo.** 2018. "Income Perception, Information, and Progressive Taxation: Evidence from a Survey Experiment." Political Science Research and Methods; Cambridge, 6(1): 83–110.

**Ferrer-i-Carbonell, Ada.** 2005. "Income and well-being: an empirical analysis of the comparison income effect." Journal of Public Economics, 89(5): 997–1019.

**Festinger, Leon.** 1954. "A theory of social comparison processes." Human relations, 7(2): 117–140.

**Frank, Robert H.** 1985. Choosing the right pond: Human behavior and the quest for status. Oxford University Press.

**Frank, Robert H.** 1989. "Frames of Reference and the Quality of Life." The American economic review, 79(2): 80–85.

**Genicot, Garance, and Debraj Ray.** 2020. "Aspirations and economic behavior." Annual Review of Economics, 12(1): 715–746.

**Godechot, Olivier, and Claudia Senik.** 2015. "Wage comparisons in and out of the firm. Evidence from a matched employer–employee French database." Journal of Economic Behavior & Organization, 117: 395–410.

**Haaland, Ingar, Christopher Roth, and Johannes Wohlfart.** 2023. "Designing Information Provision Experiments." Journal of Economic Literature, 61(1): 3–40.

**Haidt, Jonathan.** 2024. The anxious generation: How the great rewiring of childhood is causing an epidemic of mental illness. Random House.

**Hirschman, Albert O., and Michael Rothschild.** 1973. "The changing tolerance for income inequality in the course of economic development: With a mathematical appendix." The Quarterly Journal of Economics, 87(4): 544–566.

**Hoy, Christopher, and Franziska Mager.** 2021. "Why Are Relatively Poor People Not More Supportive of Redistribution? Evidence from a Randomized Survey Experiment across Ten Countries." American Economic Journal: Economic Policy, 13(4): 299–328.

**Hvidberg, Kristoffer B, Claus T Kreiner, and Stefanie Stantcheva.** 2023. "Social Positions and Fairness Views on Inequality." The Review of Economic Studies.

**Hyman, Herbert Hiram.** 1942. "The psychology of status." Archives of Psychology (Columbia University).

**Jäger, Simon, Christopher Roth, Nina Roussille, and Benjamin Schoefer.** 2024. "Worker beliefs about outside options." Q. J. Econ., 139(3): 1505–1556.

**Kaiser, Caspar, and Andrew J. Oswald.** 2022. "The scientific value of numerical measures of human feelings." Proceedings of the National Academy of Sciences, 119(42): e2210412119.

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

**Kőszegi, Botond.** 2014. "Behavioral contract theory." Journal of Economic Literature, 52(4): 1075–1118.

**Kuhn, Peter, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn.** 2011. "The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery." American Economic Review, 101(5): 2226–47.

**Kuziemko, Ilyana, Ryan W Buell, Taly Reich, and Michael I Norton.** 2014. ""Last-place aversion": Evidence and redistributive implications." The Quarterly Journal of Economics, 129(1): 105–149.

**Luttmer, Erzo F. P.** 2005. "Neighbors as Negatives: Relative Earnings and Well-Being." The Quarterly Journal of Economics, 120(3): 963–1002.

**Martinangeli, Andrea FM, and Lisa Windsteiger.** 2021. "Last word not yet spoken: a reinvestigation of last place aversion with aversion to rank reversals." Experimental Economics, 24(3): 800–820.

**McBride, Michael.** 2001. "Relative-income effects on subjective well-being in the cross-section." Journal of Economic Behavior & Organization, 45(3): 251–278.

**Mill, John Stuart.** 1859. On Liberty. Standardbook.org (Original work published in 1859.).

**Perez-Truglia, Ricardo.** 2020. "The effects of income transparency on well-being: Evidence from a natural experiment." American Economic Review, 110(4): 1019–54.

**Reck, Daniel, Joel Slemrod, and Trine Engh Vattø.** 2022. "Public disclosure of tax information: Compliance tool or social network?" J. Public Econ., 212: 104708.

**Reer, Felix, Wai Yen Tang, and Thorsten Quandt.** 2019. "Psychosocial well-being and social media engagement: The mediating roles of social comparison orientation and fear of missing out." New Media and Society, 21(7): 1486—-1505.

**Schlag, Karl, and James Tremewan.** 2021. "Simple belief elicitation: An experimental evaluation." J. Risk Uncertain., 62(2): 137–155.

**Simonsohn, Uri, Joseph P Simmons, and Leif D Nelson.** 2020. "Specification curve analysis." Nature Human Behaviour, 4(11): 1208–1214.

**Smith, Adam.** 1759. The Theory of Moral Sentiments. DD Raphael & AL Macfie. Liberty Fund.(Original work published in 1759.)[ELK].

**Twenge, Jean M., Jonathan Haidt, Jimmy Lozano, and Kevin M. Cummins.** 2022. "Specification curve analysis shows that social media use is linked to poor mental health, especially among girls." Acta Psychologica, 224: 103512.

**van Rooij, Maarten, Olivier Coibion, Dimitris Georgarakos, Bernardo Candia, and Yuriy Gorodnichenko.** 2024. "Keeping Up with the Jansens: Causal Peer Effects on Household Spending, Beliefs and Happiness."

**Veblen, Thorstein.** 1899. "The theory of the leisure class. (republished, 1934, by new york: Modern library)."

**Vogel, Erin A., Jason P. Rose, Bradley M. Okdie, Katheryn Eckles, and Brittany Franz.** 2015. "Who compares and despairs? The effect of social comparison orientation on social media use and its outcomes." Personality and Individual Differences, 86: 249–256.

**Vogel, Erin A., Jason P. Rose, Lindsay R. Roberts, and Katheryn Eckles.** 2014. "Social comparison, social media, and self-esteem." Psychology of Popular Media Culture, 3(4): 206–222.

**Weesie, Jeroen.** 1999. "On Seemingly Unrelated Estimation and the Cluster-Adjusted Sandwich Estimator." Stata Technical Bulletin, 52(10): 34–47.

**Wilkinson, Richard, and Kate Pickett.** 2010. "The spirit level." Why equality is better for everyone.

**Yamada, Katsunori, and Masayuki Sato.** 2013. "Another avenue for anatomy of income comparisons: Evidence from hypothetical choice experiments." Journal of Economic Behavior & Organization, 89: 35–57.

**Young, Cristobal, and Katherine Holsteen.** 2017. "Model Uncertainty and Robustness: A Computational Framework for Multimodel Analysis." Sociological Methods & Research, 46(1): 3–40.