Discussion Paper No. 930

CAN WE STEER INCOME COMPARISON ATTITUDES BY INFORMATION PROVISION?: EVIDENCE FROM RANDOMIZED SURVEY EXPERIMENTS IN THE US AND THE UK

Hitoshi Shigeoka
Katsunori Yamada

March 2015

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Can We Steer Income Comparison Attitudes by Information Provision?: Evidence from Randomized Survey Experiments in the US and the UK*

Hitoshi Shigeoka† Katsunori Yamada‡

Abstract

Economists have long been concerned that negative attitudes about relative income reduce social welfare. This paper investigates whether such attitudes can be mitigated by a simple information treatment. Toward this end, we conducted an original randomized online survey experiment in the US and the UK. As a baseline result, we find that UK respondents compare their incomes with others' at a much higher rate than US subjects do. Additionally, we find that our information treatment—suggesting that comparing income with others may diminish their welfare even when income levels are actually increasing—made respondents compare incomes *more*, rather than less. Interestingly, we find such effects only among UK respondents. The mechanism for this among UK respondents seems to be driven by those who are initially less comparison-conscious becoming more comparison-conscious, indicating that our information treatment gives moral "license" to make comparisons by informing that others actually do.

Keywords: income comparison, information treatment, randomized survey experiments, inter-country survey, discrete choices approach

JEL classifications: C9, D1, D3.

*Yamada acknowledges financial supports from JSPS research fund (24683006). We are grateful to Andrew McGee, Erik Kimbrough, and seminar participants at Osaka Prefecture University, Simon Fraser University, Tohoku University, Kyoto University, and Academia Sinica for their suggestions and comments. Any remaining errors are ours.

†Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A1S6, CANADA. Email: hitoshi_shigeoka@sfu.ca

‡Faculty of Economics, Kindai University, 228-3, Shin-Kami-Kosaka, Higashi-osaka City, Osaka 577-0813, JAPAN. E-mail: kyamada@kindai.ac.jp

1 Introduction

Economists have long been concerned that a tendency to compare own income with others' income may reduce social welfare.[1] Recent studies have shown that relative income comparison negatively impacts job satisfaction (e.g., Clark and Oswald 1996; Card et al. 2012), happiness (e.g., Luttmer 2005; Ferrer-i Carbonell 2005), health and longevity (e.g., Marmot 2004; Eibner and Evans 2005), and risk of suicide (Daly et al. 2013).[2] If welfare losses from income comparisons are indeed significant, any governmental interventions that alter such attitudes may have large welfare consequences.[3]

This paper investigates whether an individual's income-comparison attitudes can be mitigated by simple information treatment. Toward this end, we conduct an original randomized online survey experiment in the US and the UK. Our dataset is nationally representative in each country, which is an important virtue of our sample. About 4,500 respondents from each country are randomized into treatment and control groups, with information treatment as the intervention. We inform participants in the treatment group that people tend to compare their incomes with that of others, and that such behaviors may diminish happiness even when absolute income is actually increasing. Note that we do not provide any subjective judgment; that is, we do not suggest that people should stop comparing incomes for the sake of better social welfare.

We employ the following strategy to quantify income-comparison attitudes. In line with theoretical studies of income comparisons, such as Abel (1990), Johansson-Stenman et al. (2002), Dupor and Liu (2003), and Liu and Turnovsky (2005), we consider a utility function with a parameter that captures effects of income comparisons. We then estimate the parameter in the random utility model framework by applying obtained data from a discrete choices approach: participants choose among alternative combinations of hypothetical monthly income amounts, both for themselves and their reference persons.[4]

[1]See, for example Veblen (1899); Duesenberry (1949); Easterlin (1974, 1995, 2001); Hamermesh (1975); Akerlof and Yellen (1990); and Clark et al. (2008).

[2]Note that others' income levels *per se* can have positive effects on welfare. For example, Clark et al. (2009) argue that high income levels of neighbors may have positive impacts on life satisfaction when they serve as social capital to the community.

[3]In order to identify if governmental interventions could be beneficial, researchers have to establish a welfare criterion. See Loewenstein and Ubel (2008) and Benjamin et al. (2014b) for thoughtful discussions on this point. Our strategy in this paper is rather simple that we assume welfare loss due to income-comparisons could be characterized with a parameter of utility function.

[4]Note that income comparison is one form of social preference among many. Different aspects of social preferences include altruism, egalitarianism, reciprocity, and trust (e.g., Fehr and Falk 1999; Fehr and Schmidt 1999; Charness and Rabin 2002; Fisman et al. 2007, Andreoni and Bernheim 2009; Andreoni and Rao 2011; Horton et al. 2011; Rand 2011; Hossain and Li 2014). Our approach is different from that in the large strand of literature on social preferences using game-theoretical frameworks. Unlike these game-theoretical frameworks, we do not consider strategic interactions among respondents. This feature is actually important if we want to estimate parameters of utility functions that can be used for drawing macroeconomic-welfare implications of income-comparisons.

An alternative and traditional way of measuring income-comparison attitudes is to directly ask a question such as "How important is it for you to compare your income with other people's incomes?"[5] Despite its apparent simplicity, inter-country analysis based on such subject evaluations is problematic when these countries do not share the same rule of assessments and sense of scale (King et al. 2004). For example, it is difficult to compare a score of '3' in the US and score of '6' in the UK from subjective evaluations. Additionally, it is difficult to draw any cardinal inferences from such ordinal survey responses. The main advantage of our approach is that we estimate a parameter of a utility function, which is cardinal and thus comparable across countries.[6]

Our results can be divided into two parts. The first part of the paper documents the baseline country differences in income-comparison attitudes as measured before our information intervention (so that the responses are not contaminated by the information treatment). This baseline result is important itself, as no previous inter-country studies of income-comparison attitudes are available with non-subjective evaluation data, even limiting the study to levels of comparison attitude; here, we provide these results and extend them with the results of changes from external interventions. We find that respondents from the UK exhibit much stronger income-comparison attitudes (i.e., are much more jealous) than those from the US, possibly indicating that welfare loss due to income comparison may be more severe in the UK than in the US. The difference is large and statistically significant, and it is robust to control of observable individual characteristics, such as age, gender, education, and own income.

After documenting these baseline country differences, the second half of the paper reports the results from our information treatment. Overall, we find that information treatment makes subjects compare even *more* (i.e., become more jealous), not less. This result indicates that our information treatment guided subjects in a direction that may reduce welfare.

Notably, we find such effects for only the UK sample. This result suggests that the information treatment makes the UK participants compare themselves against others more, even though the baseline results show that UK participants compare more than US participants do. Hence, country differences are present in not only absolute levels of income-comparison attitude but also responses from external intervention. In terms of the magnitude, our information treatment changes the income-comparison attitude by one-tenth of the mean of the control group in the UK. This effect is robust to number of sample selections. To avoid respondents who answered without care or those who left the surveys open for days, we exclude those responses. In addition, we limit the sample to those who correctly answered the verification question to ensure that respondents understood the

[5]For example, the subject chooses among 0–6, where 0 corresponds to "Not at all important" and 6 corresponds to "Very important."

[6]There are some recent experimental studies examining concordance of results from choice data and those from subjective evaluations data. See, for example, Benjamin et al. (2012), Benjamin et al. (2014a), and Perez-Truglia (2015).

contents of the information provided. Such sample selections affect our estimates only negligibly.

Finally, to shed light on the mechanism of information treatment guiding participants in an undesirable direction, we examine the heterogeneous effects among UK subjects—which we find as the main effect—focusing especially on baseline income-comparison attitude as measured before information treatment. Interestingly, our results are driven by those who are initially less comparison-conscious, rather than by those who are initially jealous. This result suggests that our information treatment raises the salience of income comparison by making the recipients more aware of income comparisons, rather than making people realize its negative consequences. Consequently, our information treatment might have provided a "moral license" to compare income with others by informing that other people actually do so.

Overall, our findings suggest that there is some scope for governmental intervention to affect deep preference parameters such as income-comparison attitudes through information treatment. However, our results also imply that affecting the deep parameters in a desired direction may be difficult. Given our results, the government may do better to explicitly suggest that people should stop comparing each other because it is a "bad" thing to do. Apparently, this is a much stronger message than our current information treatment. Further discussion will be required about how strong such messages should be and what information the government should convey to alter income-comparison attitudes in a favorable direction.

A paper closely related to our study is Card et al. (2012). In that study, employees were provided with access to information on co-workers' wages, and the study examined whether this affected job satisfaction and job-seeking behaviors. The major difference of our study from that study is that we examine changes in preferences instead of changes in behavior; we made this choice because behavior is affected not only by preferences but also by many other environmental factors (e.g., peer effects).

This paper is also closely related to an emerging strand of literature that examines the effects of randomized information treatment on individual preferences. For example, some studies have examined the effect of information treatment on preferences for redistribution (see e.g., Cruces et al. 2013; Kuziemko et al. 2015; Zilinsky 2014). To the best of our knowledge, ours is the first paper to examine, in a non-laboratory setting, the effect of information treatment on individuals' deep parameters, such as income-comparison attitudes, rather than on policy preferences.[7]

Here, the long tradition of neoclassical economics takes preferences as deep parameters and assumes that they are not altered by external interventions (Stigler and Becker 1977). However,

[7]A handful of lab experiments have examined the effects of information treatment on individual preferences. For example, Benjamin et al. (2010) made ethnic identity salient for laboratory subjects and examined how it affected time and risk preferences. Chen and Li (2009) measured the effects of induced group identity on social preferences in a laboratory setting.

recent literature suggests that individuals' preferences may be altered by large negative shocks, such as early-life financial experiences (Malmendier and Nagel 2011; Giuliano and Spilimbergo 2014), financial crises (Guiso et al. 2013), changes in both macroeconomic events and household-level labor market outcomes (Krupka and Stephens 2013), conflicts or violence trauma (Voors et al. 2012; Callen et al. 2014), and natural disasters (Eckel et al. 2009; Cameron and Shah 2015; Hanaoka et al. 2014).[8] This paper is in line with recent literature suggesting that individual's deep parameters may be altered.

Finally, this paper is also related to a small but growing body of literature that uses online platforms that allow random allocation of questionnaires across subjects.[9] Most such papers use Amazon Mechanical Turk (MTurk), an online labor market in the US (e.g., Horton et al. 2011; Kuziemko et al. 2015) as an online platform. Our platform, which we describe in detail later, is another promising platform because we can balance on age, gender, and education. In our survey, subjects between the ages of 20 and 65 are selected by stratified random sampling in such a way that the cohort profiles of our sample mirrors the US and UK census data for age, gender, education, and geographic distribution.

The remainder of this paper is organized as follows. Section 2 summarizes our original randomized online survey experiments. Section 3 reports the baseline differences in income-comparison attitudes between the US and the UK. Section 4 reports the findings from information treatment. Section 5 discusses the potential explanations for the findings, and Section 6 concludes.

2 Survey Experiment

2.1 Data collection

Our original survey experiment consists of three parts.[10] The first part starts with 12 questions about socio-economic variables such as age, gender, race, educational attainments, and total pretax income in 2013. This is followed by 7 questions about respondents' subjective values. For the sake of clarity, the exact phrases for the first 12 questions were taken from the World-Value-Survey (2012) whenever available. The questions on subjective values included questions about happiness, the intensity of comparisons of income with others, information on whose income would they be most likely to compare their own with (comparison benchmark), and the predicted income of the

[8]See also Becker et al. (2012) for a review.

[9]There are a few studies in political science/economy that conducted survey experiments with random allocations of questionnaire across subjects (e.g., Horiuchi et al. 2007; Hainmueller and Hiscox 2010; Di Tella et al. 2012; Naoi and Kume 2011).

[10]See Appendix A1 for a flow chart. Survey questionnaires for the US and the UK are available as a supplemental online material at *http://www.iser.osaka-u.ac.jp/library/dp/2015/DP0930som.pdf.*

comparison benchmark. These questions were borrowed from Clark and Senik (2010). The second part uses four repeated discrete choice questions regarding environmental policy, which were taken from Viscusi et al. (2008). Data obtained from this part are not used in this paper. The final part starts with measuring income-comparison attitudes through hypothetical income questions before information treatment as a baseline. We explain how we measure income-comparison attitudes in Section 2.3. After assessing baseline attitudes, we provide the randomized information treatment to respondents as detailed in Section 2.2. Finally, we measure income-comparison attitudes again.

Our online survey experiment was carried out from March 24 to April 3 (11 days). As clearly documented in Horton et al. (2011), "[t]he validity of economics experiments depends heavily upon trust, particularly subjects' trust that the promulgated rules will be followed and that all stated facts about payment are true." As was the case in our previous original internet-based survey with Japanese subjects, reported in Yamada and Sato (2013), Nikkei Research, Inc., who maintains sample pools in the US and the UK and upholds high research ethics, was chosen to administer the survey. To maintain the quality of the data, Nikkei Research, for example, excludes so-called mindless subjects, such as those who provide the same answer to all questions in any block of discrete-choice questions. The company also excludes observations when the time to finish the survey is too short (< 4 min). In each country, subjects between the ages of 20 and 65 are selected, using stratified random sampling such that the cohort profile of the sample mirrored the national census statistics on age, gender, education, and geographical distribution.

In the invitation email sent to registered candidates, we specify that the survey is being conducted for research purposes and that the anonymity of subjects is completely secured. We also inform them that the survey should take about 15 minutes to complete. If candidates wish to participate in the survey, they are instructed to click a link in the email, which directs them to our stand-alone survey website, written in HTML. At the start of the survey, they are informed that they can withdraw their participation at any time and that they cannot return to previously answered questions. The payment for the survey is USD1.5 for US respondents and GBP1 for UK respondents. Payment is only available upon completion of the survey. The survey took respondents 13.4 minutes on average, which is an average hourly wage of roughly USD6.7 for US respondents and GSB4.5 for UK respondents. There are roughly 4,500 respondents in the US sample and about the same number in the UK sample. As we explain later, because we have two treatment groups, respondents are categorized into three groups for each country, and the sample size of each group is around 1,500 people.

We take several steps to ensure the validity of the results. First, we limit the sample to only residents in the US and the UK by asking respondents if they are US or UK residents at the beginning of the survey. Second, to avoid mindless subjects or those who left the survey open for days, we exclude such subjects in the robustness checks described in Section 4.3. To prohibit subjects from mindlessly skipping through the questions, we enforce answering for all but one question by not allowing respondents to proceed to next question unless they fill in the answers.[11] Third, we ask a verification question to ensure that people understand the contents of the information provided right after our information treatment. When a respondent answers incorrectly on the verification question, we show the same figure and instruction one more time, and ask the same verification question again. In one of the robustness checks described in Section 4.3, we limit the sample to those who correctly answered the verification question.

2.2 Information treatment

Our objective for information treatment is to alert our respondents to the existence of the hedonic treadmill in relation to income comparisons. To this end, we prepared a figure on the famous "Easterlin paradox," indicating that income growth has not been accompanied by an increase in happiness score. Then, we follow the literature in suggesting that this happens because people compare their income levels with those of others, and consequently an increase for all does not increase the happiness of anyone. We choose this figure—while being well aware that the paradox is not accepted by all academics—because the story, illustrated with only two time-series plots of income and happiness data, is one of the most simple and intuitive figures for allowing recipients to understand the possible negative consequence of income comparison.[12] Additionally, the figure itself is easy to understand and should not require a high cognitive skill, even among general subjects who are not familiar with economics.

In our survey experiment, there are two treatment groups (Treatment 1 and Treatment 2). We describe each treatment in order.

The first treatment group is only provided the Easterlin paradox figure, accompanied by the following very short descriptions of the observation.

> *[W]hile real income per capita increases sharply, happiness showed essentially no trend and has remained constant over time. From this figure, it looks as if individuals in the United States are in the "flat part of the curve" with additional income buying little, if any, extra happiness.*

[11] The only question that respondents could skip was Q7A, which asks "What do you think was your reference group's personal TOTAL income, before taxes, last year (2013)?" We allow this exception due to the potential risk of resistance to reporting other's income. However, this exception does not seem to pose any issues since there were only 11 (resp., 12) people who did not answer this question in the US (UK) sample.

[12] This "paradox" is challenged by Stevenson and Wolfers (2008, 2013), Sacks et al. (2012), and Bond and Lang (2014). However, the scientific validity itself is not a problem for our research.

Figure 1 displays a screenshot of our first treatment for the US case. We prepared UK-specific figures for the UK case. Note that this first treatment conveys only facts with data and does not explain the implications of this figure.

The second treatment group sees exactly the same figure and short description that the first treatment group sees, but we additionally add the following short description as a possible explanation for the observed relation.

> *It has been suggested by researchers that this happens because people tend to compare their income levels against the incomes of others. As long as people compare their income against the incomes of others, increasing everyone's income increases the happiness of no one.*

It is apparent here that we add two pieces of information. The first sentence points out people's tendency to compare their income with others', while the second one points out the possible negative consequences for welfare of such income comparisons. Here, it should be stressed that we do not provide a critical view of making income comparisons for the sake of greater welfare (e.g., "you should not covet"). Instead, we follow a libertarian paternalistic method of intervention (Sunstein and Thaler 2003; Thaler and Sunstein 2003), leaving the freedom of how to digest the information in the hands of recipients.

It is not certain *a priori* whether our information treatment will cause recipients to compare more or less with others (i.e., to become more or less jealous). On one hand, since we point out not only people's tendency to compare with each other but also its possible negative consequences, recipients who are made aware of the hedonic treadmill of income comparisons may stop comparing with others. On the other hand, if our information treatment results in giving a "moral license" to compare with others by informing that others actually do so (because we are living in a material world), then recipients—especially those who were initially less conscious about others—may compare more after our information treatment.

We provide the control group with information unrelated to income comparisons: a figure on the extent of Arctic sea ice in February over the last 35 years. Appendix Figure A1 is a snapshot of our information treatment for the control group. We also give some explanations for the figure that include roughly the same number of words as the treatment groups to give the same level of cognitive burden. We use responses from people who receive this placebo treatment to assess the differential response rates between treatments and control groups, and to take account for fatigue effects in the survey experiment.

After the information treatment, we ask a very simple verification question to ensure that the subjects understand the provided information.[13] For those who incorrectly answer the verification

[13]For Treatment 1 and Treatment 2, the question is "While real income per capita has increased sharply, the

question, we repeat the same question again after showing the same figure and descriptions one more time. As the content in the information treatments is simple, 76.6 percent of respondents answer the verification question correctly on the first attempt, with 87.2 percent answering correctly on the first or second attempt.[14] Importantly, the accuracy rates among different treatments are very similar: 87.6, 88.5, and 85.3 percent for Treatment 1, Treatment 2, and Control, respectively.[15] Although evidence of this is indirect, it is reassuring that our information treatment seems to provide similar cognitive burdens.

Because the content of the information provided necessarily differs by treatment group, one common concern is that attrition may differ by treatment group as well. Among those who initiated the survey, the attrition rate as a whole was 15.3 percent for the US, and 17.9 percent for the UK.[16] Among those who withdrew participation before completion, only 12.7 (resp., 13.3) percent of subjects did so after they received information treatment in the US (UK) case. Also, attrition that happened after the treatment does not seem to qualitatively differ by type of treatment. Specifically, among 106 (resp., 131) respondents who initiated participation but withdrew after the information treatment in the US (UK) case, 30.0 (27.4) percent of belonged to Treatment 1, 34.0 (35.9) percent to Treatment 2, and the rest, 35.8 (36.6) percent, to Control. The slightly lower attrition rate for Treatment 1 may reflect that the instructions were shorter by two sentences than those in Treatment 2.[17] Overall, attrition does not seem to drive our results.

2.3 How to measure income-comparison attitudes: a discrete choice approach and a random utility model

We describe how we measure the sign and magnitude of income-comparison attitudes from data of discrete choice questions. As stated earlier, income-comparison attitudes are measured twice for the same individuals, before and after information treatment. Respondents are asked to choose between two alternative situations, where each situation is a combination of hypothetical monthly income amounts for both themselves and their reference persons (comparison benchmark), keeping

trend of happiness has been [blank]" and the choices are among "increasing", "constant" (correct answer), and "decreasing". For the control group, the question is "The winter month trends of the Arctic Sea ice extent have been [blank]", and the choices are among "increasing", "constant", and "decreasing" (correct answer).

[14] We take this high accuracy rate of the small quiz as fairly good. Rand (2011) reported that at least 80 percent of experiment participants in MTurk were not merely making random selections on survey questions, which is similar to our figures.

[15] In the same vein, the accuracy rates at the first attempt are 76.8, 78.7, and 74.4 percent for for Treatment 1, Treatment 2, and the control group, respectively. We did not repeat this verification question more than twice.

[16] For example, the overall attrition rate in Kuziemko et al. (2015) using MTurk was 22 percent.

[17] In terms of the number of respondents, the difference between treatment groups 1 and 2 in those who withdrew after receiving the treatment was only 4 in the US case and 11 in the UK case. Unfortunately, we do not have information on individual characteristics for those who withdrew from the survey, and so we cannot examine differential attrition by individual covariates.

in mind that the price levels in the two situations are the same. Respondents also have the option to choose “don’t know/cannot answer,” following the suggestions by Arrow et al. (1993) and Haaijer et al. (2001).[18]

To improve the realism of our hypothetical income questions, we elicit information on each subject’s comparison benchmark and impose the answers in the screens on discrete choice questions.[19] Table 1 presents the breakdown of comparison benchmarks chosen by US and UK respondents. In the table, the top-five choices of comparison benchmark are the same between the US and the UK. The highest fraction is “I don’t compare,” at 29 percent for US respondents and 27 percent for UK respondents. This is followed by work colleagues, average people in the US/UK, close friends, and family members, in that order; each choice represents roughly 10–20 percent. These top 5 choices make up 85.3 (resp., 87.5) percent of all choices for the US (UK). Interestingly, this pattern almost replicates the pattern found in the corresponding table of Clark and Senik (2010) for 18 European countries.

Before the subjects begin responding to repeated discrete choice questions, they are shown a screen displaying the following instructions.

> *In the following screens we show your hypothetical monthly income (before tax). Also displayed in the same screen is your reference group’s monthly income (before tax). Suppose that these are the current situations of your monthly income (before tax) and your reference group’s monthly income (before tax).*

In the subsequent screens, we ask subjects to answer hypothetical discrete choice questions to choose among two situations, using the following format.

> *Comparing situation 1 and situation 2 shown in the figures, which is more preferable to you? Suppose that the price levels in the two situations are the same. Please choose from the following options.*

Each choice comprises two situations, with each situation defined by two attributes: monthly pre-tax income of respondents and that of the comparison benchmark. The income levels are chosen

[18]Arrow et al. (1993) and Haaijer et al. (2001) pointed out the importance of including a no-choice option in hypothetical choice questions. In the main analysis, we remove the no-choice selections from consideration. An alternative way is to interpret them as showing indifference between the two situations, rather than a failure to understand the survey question. Unfortunately, we have no information about the true reason why the no-choice option was chosen.

[19]The choice sets for the comparison benchmark are: 1. neighbors; 2. classmates from your school days; 3. close friends; 4. family members; 5. family members of your children’s classmates; 6. work colleagues; 7. average people in the country; 8. Friends or acquaintances other than the above; 9. others; 10. “don’t know”; 11. “I don’t compare.” For those who answer either 9, 10, or 11 in the question of comparison benchmark, we encourage an income comparison with a reference group of “someone like you,” suggesting to respondents that “in psychology, however, it has been known that people tend to compare themselves, if anything, with those with the same age group, gender, and academic background as themselves.”

on the basis of actual pre-tax monthly income quintiles in the US (resp., UK): USD900 (GBP1,000), USD1,800 (GBP1,500), USD2,900 (GBP2,000), USD4,900 (GBP3,000), and USD7,200 (GBP4,000).[20] As a snapshot of such an example, Figure 2 asks subjects to choose among situation 1, where own income is USD900 with reference group income of USD4900, and situation 2, where own income is USD1800 with reference group income of USD2900.

Such discrete income choice questions are repeated six times both before and after the information treatment (i.e., a total of 12 times) for each respondent. We prepared a total of 25 variations of choice questions, and 12 of those 25 are randomly assigned to each participant. Hence, no participant answers the same question twice. Appendix Table A1 presents all 25 variations of choice question used in this study. The details of how we constructed these hypothetical discrete choice questions are described in Appendix A2.

From respondents' choices of preferred income scenarios in our hypothetical choice questions, we estimate parameters of a utility function in the random utility model framework. We assume that respondents choose the income situation that offers them higher utility. Following Abel (1990), Johansson-Stenman et al. (2002), Dupor and Liu (2003), and Liu and Turnovsky (2005), we consider the constant relative risk aversion-type utility function U, given by

$$U = \frac{(y\bar{y}^{\gamma})^{1-\rho}}{(1-\rho)}, \tag{1}$$

where y is own income level and $\bar{y}$ is income level of others.[21] The parameter γ reflects the attitude of income comparisons. Jealousy is reflected in $\gamma < 0$, and altruism in $\gamma > 0$. When $\gamma = 0$, there is no income comparison.

The distribution of parameter γ is estimated by using a mixed logit model and maximizing the likelihood of observing subjects' choice patterns (Train 2009). We first estimate the distribution of γ for all respondents. Then, the value of γ for each respondent is obtained by finding the point in the distribution corresponding to each respondent from the choice patterns. Appendix A2 gives full technical details for the estimation method (see also Yamada and Sato 2013).

Some remarks about our approach to measuring attitudes of income comparison through a discrete choice approach are in order.[22] Our approach is different from eliciting subjective evalu-

[20]The sources of income distributions are the Current Population Survey 2013 for the US and the Survey of Personal Incomes 2011–2012 for the UK.

[21]Other theoretical macroeconomic studies that focus on preference externalities include Fershtman et al. (1996) and Corneo and Jeanne (1999).

[22]See also Solnick and Hemenway (1998), Johansson-Stenman et al. (2002), Alpizar et al. (2005), and Carlsson et al. (2007) investigated the intensity of social comparisons via hypothetical choice questions. The choice format in these studies, however, did not allow researchers to estimate a parameter of a utility function for each respondent. Carlsson et al. (2009) conducted discrete choice question on income comparisons with repeated choice questions, although their interest was not in the distribution of income-comparison attitudes.

ations, which is an alternate to the conventional way of measuring income-comparison attitudes. In the conventional approach, answers to a direct question, such as "How important is it for you to compare your income with other people's incomes?", are used to infer the intensity of income comparison (Clark and Senik 2010). Although this is a simple and convenient way of measuring attitudes, our discrete choice approach has several advantages. First, it is a drawback of the conventional method that there is no straightforward way of making meaningful comparison across individuals on the basis of subjective evaluation if respondents do not share a common rule for assessment and a common scale (e.g., a score of 3 and a score of 6 may represent equivalent feelings between to respondents).[23] In fact, King et al. (2004) point out the incomparability of answers to inter-country surveys when such subjective evaluations are employed. Second, it is difficult to draw any cardinal inferences from ordinal survey responses. In contrast, because we measure a parameter of utility function for each respondent, we can safely compare income-comparison attitudes between respondents, and it is straightforward to draw cardinal inferences about welfare from our results.

3 Results from baseline data

3.1 Summary statistics

Table 2 provides summary statistics from the US and UK samples. The questions about individual characteristics are asked at the beginning of the survey. Columns 1 and 4 of the table show that average age is 41 years-old for both samples. Both samples overrepresent white respondents (at 73 percent and 90 percent in the US and the UK, respectively). More than half of respondents are married; 7.8 (resp., 10.2) percent of respondents are immigrants to the US (UK).

We compare these summary statistics with national statistics in terms of age, gender, and education. Columns 2 and 5 present the national values, as found by the most recent census in the US (column 2) and the UK (column 5). We are trying to match the population characteristics by using stratified random sampling, and we note that the age and gender distributions are similar to those in the national statistics; this contrasts with many studies that rely on potentially biased samples from sources such as MTurk. It is rather unfortunate that those in the lowest education category in both countries are underrepresented in our data. To deal with the underrepresentation of those with lower educational attainments, we reweighted our sample so that it matches the census across 40 cells for gender (2), age bracket (4), and education categories (5). Columns 3

[23]See Daly et al. (2013) for a critical view of subjective evaluations, in particular happiness studies. Another flaw of direct questions is that it will give subjects incentives to misrepresent true responses. Many people dislike thinking of themselves as comparison-conscious, and they therefore underestimate the degree to which they care about other's income. This resembles the purchase of moral satisfaction in Kahneman and Knetsch (1992).

and 6 show the reweighted sample distribution. In general, reweighting makes quantitatively little difference to the analyses presented below. Therefore, we report the unweighted results throughout the paper (weighted results are available upon request).

3.2 Country differences in income-comparison attitudes

We first document the country-level difference in the baseline income-comparison attitudes as measured before the information treatment (hereinafter, γ_0). Our approach of using discrete choice questions can make a unique contribution to inter-country analysis on income-comparison attitudes because we do not rely on subjective evaluations in estimating such attitudes. Furthermore, even from subjective evaluation data—a parsimonious method of measuring the intensity of income comparison—there seems to have been no previous evidence for a difference in the intensity of income-comparison attitudes between the US and the UK.[24]

Figure 1 plots the distribution of baseline income-comparison attitude γ_0 for the US and the UK separately. The distribution for the UK is apparently shifted to the left from that of the US, indicating that the average γ_0 for the UK is much lower than that of the US. Recall that the higher values indicate less jealousy. This result suggests that respondents in the UK are more jealous than those in the US.

We can obtain the same results by regression. In column 1 of Table 3, our measures of income-comparison attitude γ_0 are regressed on the UK dummy, which takes a value of one for respondents in the UK sample. The intercept is the average value for the US, and the use of the UK dummy captures the difference between the US and the UK sample. Column 1 shows that the averages for γ_0 are 0.056 and -0.129 ($= -0.185 + 0.056$) in the US and the UK, respectively. The value for the US is close to zero, suggesting that subjects in the US do not care much about others' income.[25] In contrast, the average γ_0 for the UK has an opposite sign and a much larger magnitude, suggesting that UK respondents are much more jealous than US respondents. The UK dummy of -0.185 is statistically significant at the 1% level.[26]

To see whether the between-country difference in income-comparison attitudes is simply a

[24]Clark and Senik (2010) use subjective evaluation data on attitudes of income comparison from European Social Survey (round3) but information of the US is missing. Also World Values Surveys 2010–2012 (question V96) ask about the perception of income equality and egalitarianism. However, the question, albeit capturing an aspect of social preferences, is not about the intensity of income comparisons

[25]This result is consistent with results in Alesina et al. (2004) showing that people in the US do not care about inequality as much as people in European countries. Our results echo recent finding by Luttmer and Singhal (2011), suggesting that part of the preference for redistribution can be explained by cultural determinants that may not change much over time.

[26]We confirmed that this US–UK difference does not stem from the difference in absolute income levels chosen in the discrete income choice questions across countries. Specifically, we re-estimated γ with replacing the UK income level by the corresponding US income level (e.g. replacing GBP1,000 by USD900, GBP1,500 by USD1,800 and so on); the obtained estimates were quantitatively very similar (results are available upon request).

consequence of the sample differences between the US and the UK, we included individual controls in addition to the UK dummy. Column 2 in Table 3 shows that the coefficient of the UK dummy is barely affected, and it is still statistically significant at the 1% level. Notably, other individual characteristics do not well predict income-comparison attitudes. An exception is ethnicity/race. Respondents who identify as Asian seem to be more jealous than those who identify as white (which is the omitted category). Nonetheless, the result gives some assurance that individual covariates do not have much predictive power for deep parameters such as income-comparison attitude.

While we cannot completely exclude the possibility that our results on the UK dummy are driven by other unobserved individual characteristics, this result may serve as a supportive evidence of a well-known endogenous growth model by Cole et al. (1992) which attributes the difference in the economic growth rates between the US and the UK, two countries with very similar economic fundamentals, to the difference in social preferences. Our study is the first evidence that document the US–UK differences in income-comparison attitudes measured in the discrete choice approaches.

As a separate exercise, we also confirm that our measures of income-comparison attitudes are indeed in line with an alternate and parsimonious measure of intensity of income comparisons. In our survey, we have one traditional self-reported measure concerning intensity of the comparison with others. The exact question is "How important is it for you to compare your income with other people's incomes?", and the answer takes values from 1 to 5, where 1 indicates "don't compare", and 5 indicates "compare intensively". Thus the higher the value is, the more they compare to others. Column 3 in Table 3 adds this variable into column 2. It is reassuring that this variable is highly correlated with our income comparison measures in the expected sign. Nonetheless, it is again noteworthy that subjective evaluation data contains above-mentioned drawbacks when researchers want to use it for the comparison purposes. Note also that this relationship is just a simple correlation as both variables are choice variables.

4 Results from information treatments

After establishing the between-country difference in our measures on income-comparison attitude, we now move onto our second set of results, those on information treatments.

4.1 Balance check

Before describing the results, we verify that randomization was correctly performed. One common way to check is to compare the initial individual characteristics across the groups and make sure that these variables are balanced. Table 4 provides the results of checking for balance, presented separately for the US and UK samples. The means for each treatment group and the control group are shown together with the p-values for the null hypothesis that the means are the same across the three groups.[27]

The results show that overall both the US and UK samples are balanced. Out of 32 variables examined in Table 4, only five variables are statistically significantly different across the groups ($p < 0.10$).[28] Also, it is reassuring that the baseline γ_0 is balanced across groups in each country. In the regression analysis below, some specifications control for these individual characteristics. As shown below, our estimates are barely affected by imposing these controls, which are shown later, further reconfirming that randomization was successful.

4.2 Estimation

Since this is a randomized survey, the econometric model is very straightforward. We simply estimate the following equation:

$$\gamma_i = \alpha + \beta_1 Treat1_i + \beta_2 Treat2_i + X_i^{'}\pi + \epsilon_i \tag{2}$$

where γ_i is the income-comparison attitude of individual i as measured after information treatment. $Treat1_i$ and $Treat2_i$ are dummy variables indicating membership in Treatment 1 and Treatment 2, respectively. These treatment dummies measure the differences in γ between the treatment groups and the control group (which is the omitted category). $X_i^{'}$ is a vector of pre-determined individual controls. We always include the baseline γ_0 as measured before the information treatment to increase the precision; this helps because the baseline γ_0 is highly correlated with γ measured after information treatment.[29] In some specifications, we also add age, gender, and education to further gain efficiency. Since baseline γ_0 and individual characteristics are uncorrelated with treatment status, inclusion of these variables should not affect our estimates. We also report the difference in β_1 and β_1 with associated standard errors to examine whether the effects of two treatments are different.

4.3 Basic results

Table 5 provides the results for estimating (2) separately for the US and the UK. Odd-numbered columns show the estimates without individual controls other than baseline γ_0. Even-numbered

[27] This presentation of the balance check follows Table 2 in Muralidharan and Sundararaman (2011).

[28] Note that we define "low-educated" as having educational attainment less than or equal to "some college" in the US, and less than level 2 qualification in the UK.

[29] Alternatively, we can take the first difference of γ before and after the information treatment. We do not take this approach because the first-difference approach arbitrarily forces the coefficient on γ_0 to be -1. Since the baseline γ_0 is balanced, the results from the first-difference approach are quantitatively similar (results available upon request).

columns show estimates with individual controls included to increase efficiency.[30] The estimates barely changes in both specifications, suggesting that randomization works well. First two columns show results for the US and the next two columns show the results for the UK. Column 1 in Table 5 demonstrates that none of the treatments have much effect on income-comparison attitudes in the US sample. Adding the covariates in column 2 has no effect on the estimated treatment effects. The magnitudes of the estimates are very small and far from statistically significant, suggesting that neither information treatment has an impact on income-comparison attitudes among US respondents.

In contrast, columns 3 and 4 show that although Treatment 1 has no impact, Treatment 2 is negative and marginally statistically significant ($p < 0.10$), suggesting that Treatment 2 makes the UK respondents more jealous. The coefficient of -0.012 on Treatment 2 dummy translates into the worsening of income-comparison attitudes by one tenth of the mean of the control group (-0.127).[31] The last two rows in column 4 show that the difference between the two treatments is only marginally significant. As shown in Section 5.1, however, the results mask a heterogeneous effect among UK respondents, and a proportion of UK respondents are indeed significantly influenced by the second treatment.

Three things are worth mentioning. First, it is interesting that Treatment 1 did not have any impact on either group. This result suggests that simply showing the observed data on income and happiness is not sufficient to affect attitudes toward income comparison. While the context is different, this result echoes the recent finding of Kuziemko et al. (2015) that displaying data with a highly skewed income distribution was not sufficient to affect preferences for redistribution. Second, Treatment 2, which add only two sentences to Treatment 1, affects the income-comparison attitudes of some UK respondents, but the effect is in an undesired direction. In Section 5.1, we discuss in more detail why we think that Treatment 2 directed those respondents in such a direction. Third, it is striking that our information treatment makes only UK respondents more jealous, even though the UK respondents are already more jealous of others at the baseline than US respondents. Hence, country differences lie not just in absolute levels of income-comparison attitudes but also in responses to external intervention. The US participants do not appear to care about others' income at all and thus the information treatment has no impact, while the UK participants are much more malleable.[32]

[30] Specifically, we added a dummy for age (in years), gender, ethnicity/race dummies, a marriage dummy, an immigrant dummy, education dummies, employment dummies (unemployed, full-time worker, and chief earner), a dummy for being a parent, and log income (in USD).

[31] The reweighted estimates are slightly bigger in magnitude. The estimate on a Treatment 2 dummy becomes -0.018 with standard error of 0.008. The difference between the estimates of Treatment 1 and Treatment 2 becomes -0.017 with standard error of 0.009.

[32] These results are consistent with those of Kuziemko et al. (2015) demonstrating that most policy preferences among MTurk subjects in the US are relatively fixed.

4.4 Robustness

We verify the robustness of our results here. The effect of Treatment 2 on the UK respondents is robust to a number of sample selections. Table 6 summarizes the results. To ease comparison, column 1 in Table 6 repeats the estimates from column 4 in Table 5. Since adding controls does not have any appreciable effects on our estimates, we report the estimates with controls in the rest of the paper to improve efficiency.

First, responses from mindless subjects were excluded. Specifically, we excluded observations from those whose time to take the survey was below the 5th percentile or above the 95th. Column 2 in Table 6 shows that the estimates for Treatment 2 are hardly changed by this.

Next, we limit the sample to those who correctly answer the verification questions. While the correctness in the verification question is slightly correlated with educational attainment (not shown), our results are robust to such sample selections. Columns 3 and 4 in Table 6 present the estimates with limiting the sample to those who answer correctly on either the first or second verification question (88.3 percent of the UK sample) and to those who answer correctly on the first verification question (79.2 percent of the UK sample), respectively. The magnitudes of the estimates are essentially unchanged, although column 4 for Treatment 2 loses statistical significance due to the smaller sample size.

5 Discussion

5.1 Heterogeneous effects of information treatment

Here, by investigating heterogeneity in the treatment effects, we discover some potential mechanisms that could explain why Treatment 2 makes the UK respondents more jealous, rather than less, in making income comparisons.[33] Of course, other explanations may account for the observation, and we do not view our explanations as definitive.

As mentioned repeatedly, Treatment 1 and Treatment 2 differ by only two sentences. The first of the two sentences added in Treatment 2 points out people's tendency to compare their income with others', while the second sentence points out the possible negative consequences of such income comparisons on welfare. Given our empirical results, we conjecture that our subjects weighted the first sentence more than the second and as a result they become more jealous after

[33] We also examined the heterogeneous effects among the UK subjects on other dimensions. We found that Treatment 2 has larger effects among female, highly educated, and young subjects. However, except for male-female differences, these were not significant, though age was statistically significant at the 10 percent level. We also investigated the effects through choice of comparison benchmark, but we did not obtain any meaningful differences, partly due to the smaller sample size once we stratified the sample by comparison benchmark (all results available upon request).

Treatment 2. Namely, the first sentence, "It has been suggested by researchers that this happens because people tend to compare their income levels against the incomes of others," heightens the salience of income comparison and ends up directing the focus towards others' income levels. Consequently, our information treatment might have provided the moral license to compare by informing that others actually do so. If this is indeed the case, we are likely to find larger effects of information treatment among the subjects who did not care much about others before information treatment.

To investigate this possibility, we first examine the heterogeneous effects according to the baseline γ_0, which is measured before our information treatment. Consistent with our conjecture, columns 1 and 2 in Table 7 show that our results on Treatment 2 are mostly driven by those who did not initially care about others' incomes. While the estimated effect of Treatment 2 among those who are initially more jealous (baseline γ_0 is below median) is -0.004 and far from statistically significant, the same estimate among those who are initially less jealous (baseline γ_0 is above median) is -0.020 and significant ($p < 0.05$). The latter estimate is large and corresponds to a roughly 20 percent change in income-comparison attitudes from the control mean. Unfortunately, the difference between the two estimates is not statistically significant at the conventional level ($\chi_2(1) = 1.62, p = 0.203$), and thus we cannot exclude the possibility that the two estimates are the same.

Second, as Duesenberry (1949) and others argue, the income comparison can be asymmetric depending on income positions. Thus, we examine whether the effects differ by perception of relative position to others' income. In our survey, we ask subjects to answer the level of own income (y_{survey}), as well as the level of expected income of the comparison benchmark ($\bar{y}_{survey}$). Then, we divide the sample into three subgroups: own income is lower than that of comparison benchmarks ($y_{survey} < \bar{y}_{survey}$), own income is higher than that of comparison benchmarks ($y_{survey} > \bar{y}_{survey}$), and own income is the same level as that of the comparison benchmarks ($y_{survey} = \bar{y}_{survey}$).[34]

Columns 1, 4, and 5 in Table 8 show that our results are driven mainly by those who stated that their own income levels are greater than those of their comparison benchmarks (column 5). Importantly, we control for own income in all the specifications. This result suggests that while their superior income position initially makes them relatively indifferent to others' income, information treatment directs them to be more conscious of others' income. In fact, as shown in third to last row in Table 8, the control mean of γ among those who are initially optimistic about their income position (-0.111) is smaller (i.e., less jealous) than both those who are pessimistic

[34]Note that $\bar{y}_{survey}$ is collected before information treatment, although we are well aware that $\bar{y}_{survey}$ is still a choice (and endogenous) variable. Alternatively, the reference income can be the average income of individuals in the same regions, which some other studies use (e.g., Luttmer 2005). However, in our study, Table 2 shows that only 1.2% percent of subjects in the UK choose neighbor as the reference group. Of course, this fraction may be a lower bound since the reference groups in other categories such as close friends may include neighbors.

(-0.124) and those who stated the same income as others (-0.137).[35]

We further investigate whether the distance between own income and that of others matters, as suggested by Ferrer-i Carbonell (2005). Specifically, for each subgroup, we estimate the following specification, which is a variant of the main specification (2):

$$\begin{aligned}\gamma_i =& \alpha + \beta_1 Treat1_i + \beta_2 Treat2_i + X_i^{'}\pi \\ & + \theta_1\{Treat1_i \times (\ln y_{survey} - \ln \bar{y}_{survey})\} + \theta_2\{Treat2_i \times (\ln y_{survey} - \ln \bar{y}_{survey})\} + \epsilon_i,\end{aligned} \tag{3}$$

where we add the interactions of each treatment with the difference between $\ln y_{survey}$ and $\ln \bar{y}_{survey}$.

Interestingly, column 6 shows that Treatment 2 makes those who are initially optimistic about their income position ($y_{survey} > \bar{y}_{survey}$) more comparison-conscious as the distance between own income and that of others becomes greater. While it is speculative, this result is consistent with that of Fehr and Schmidt (1999) on loss-aversion. Column 2 shows that such a pattern is not observed among whose stated income is lower than that of the comparison benchmark.

These results do not depend on the specification of the difference between own income and others' income. Instead of taking the difference in log income, we also take the ratio of own income to others' income, $y_{survey}/\bar{y}_{survey}$, as in Watson and McLanahan (2011). Column 7 in Table 8 demonstrates that the estimates using this measure show a very similar pattern to those in column 6, where difference in log income is used.

In sum, although we acknowledge that each piece of evidence is not sufficient on its own, the weight of the evidence taken all together supports the hypothesis that Treatment 2 makes the income comparisons more salient to our UK respondents, and that those who initially pay less attention to others' income levels became more comparison-conscious and jealous. Consequently, we suggest that we may have provided a form of moral license to compare with others by informing that other people actually do so. Such effects are exacerbated by the magnitude of perceived difference between own income and that of peers, but the effects are limited to those who think that their own income level is higher than others', possibly reflecting loss aversion.

5.2 "Paternalistic" intervention

The aim of this experiment is to examine whether government can alter income-comparison attitudes to improve welfare, and, if so, whether this is easy. For the first question, our results indicate that it is indeed possible to change income-comparison attitudes by providing simple

[35]The empirical evidence on the direction of income comparison is mixed. While Ferrer-i Carbonell (2005) finds evidence that income comparison is upward looking, McBride (2001) finds evidence that it is the other way around. Our evidence is more consistent with that of Ferrer-i Carbonell (2005) , since those whose income was lower than others tended to compare more (upward) than those whose income was higher than others (downward).

pieces of information.

However, for the second question, our respondents were actually guided in a direction of decreased welfare. Importantly, we did not tell the participants in our information treatments that they should stop comparing with each other for the sake of better welfare. In this sense, our approach is similar to a paternalistic mode of intervention (Thaler and Sunstein 2003; Sunstein and Thaler 2003), as we leave the freedom of how to digest the information in the hands of recipients: when for some reasons individuals cannot choose the best option that would provide them with the highest welfare among available options, authorities should steer their choices in directions that will improve own welfare without limiting freedom of choice. Note that our intervention is intended to directly affect preferences—not behaviors and choices—unlike previous studies of paternalistic intervention such as Thaler and Sunstein (2003) and Sunstein and Thaler (2003). Also, our intervention happens to be asymmetrically paternalistic (Camerer et al. 2003) in the sense that those who are already income-conscious are not affected, while those who are initially less jealous are more affected.

Thus, our results may serve as one illustration of a point discussed in Sunstein and Thaler (2003), which documents the difficulty in guiding agents in a desired direction by information treatment.[36] Given our results on information treatments, governments may want to take the further step of explicitly warning of the risk of the hedonic treadmill of income comparisons, doing so to mitigate the welfare loss from such attitudes. The degree of directness and explicitness, as well as what information the government should provide, is beyond the scope of this paper but warrants further investigation.

6 Conclusion

This paper investigates whether we can alter income-comparison attitudes by providing simple but intuitive information suggesting that making income comparisons may reduce welfare. To this end, we randomly assigned respondents to online surveys to different information treatments. Specifically, we showed a figure suggesting that people's tendency to compare incomes is a cause of happiness remaining unchanged, even as we are getting richer on average. We found that providing such information made respondents in the UK even more jealous, but we did not find any effect on US respondents. Our findings suggest that there is some scope for governmental intervention to affect deep preferences, such as income-comparison attitudes. However, our results also imply that affecting deep parameters in a desired direction may be difficult.

[36] Thaler and Benartzi (2004) provide some examples of successful governmental "paternalistic" interventions. For example, see Madrian and Shea (2001), Choi et al. (2002), Downs et al. (2013), and Wisdom et al. (2010). Note once again that these examples aim at changing choices, rather than preferences.

There are a couple of limitations to our study. First, we cannot examine the lasting effect of our information treatment due to data limitation. It is possible that such information treatment can alter income-comparison attitudes in the short term only.[37] However, it seems unlikely to see long-term effects where there are no short-term effects. Second, although our studies indicate the possibility that government can probably affect income-comparison attitudes, we cannot show *how* to direct the income-comparison attitudes in a desired way. Providing more direct and stronger information—stressing that making income comparisons is, indeed, a bad thing to do—could be more effective than our information treatment. These questions leave room for future research.

[37]Few papers tested the duration of effects from informational survey experiments. The evidence is mixed. For example, Druckman and Nelson (2003) found that their results disappear within ten days while Kuziemko et al. (2015) found that their effects lasted one month after the information treatment.

References

Abel, A. B., “Asset Prices under Habit Formation and Catching Up with the Joneses,” *American Economic Review* 80 (May 1990), 38–42.

Akerlof, G. A. and J. L. Yellen, “The Fair Wage-Effort Hypothesis and Unemployment,” *The Quarterly Journal of Economics* 105 (May 1990), 255–83.

Alesina, A., R. Di Tella and R. MacCulloch, “Inequality and happiness: are Europeans and Americans different?,” *Journal of Public Economics* 88 (August 2004), 2009–2042.

Alpizar, F., F. Carlsson and O. Johansson-Stenman, “How much do we care about absolute versus relative income and consumption?,” *Journal of Economic Behavior & Organization* 56 (March 2005), 405–421.

Andreoni, J. and B. D. Bernheim, “Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects,” *Econometrica* 77 (September 2009), 1607–1636.

Andreoni, J. and J. M. Rao, “The power of asking: How communication affects selfishness, empathy, and altruism,” *Journal of Public Economics* 95 (August 2011), 513–520.

Arrow, K., R. Solow, P. R. Portney, E. E. Leamer, R. Radner and H. Schuman, “Report of the NOAA Panel on Contingent Valuation,” Technical Report, National Oceanic and Atmospheric Administration, 1993.

Becker, A., T. Deckers, T. Dohmen, A. Falk and F. Kosse, “The Relationship Between Economic Preferences and Psychological Personality Measures,” *Annual Review of Economics* 4 (July 2012), 453–478.

Benjamin, D. J., J. J. Choi and A. J. Strickland, “Social Identity and Preferences,” *American Economic Review* 100 (September 2010), 1913–28.

Benjamin, D. J., O. Heffetz, M. S. Kimball and A. Rees-Jones, “What Do You Think Would Make You Happier? What Do You Think You Would Choose?,” *American Economic Review* 102 (August 2012), 2083–2110.

———, “Can Marginal Rates of Substitution Be Inferred from Happiness Data? Evidence from Residency Choices,” *American Economic Review* 104 (November 2014a), 3498–3528.

Benjamin, D. J., O. Heffetz, M. S. Kimball and N. Szembrot, “Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference,” *American Economic Review* 104 (September 2014b), 2698–2735.

Bond, T. N. and K. Lang, "The Sad Truth About Happiness Scales," NBER Working Papers 19950, National Bureau of Economic Research, Inc, March 2014.

Callen, M., M. Isaqzadeh, J. D. Long and C. Sprenger, "Violence and Risk Preference: Experimental Evidence from Afghanistan," *American Economic Review* 104 (January 2014), 123–48.

Camerer, C., S. Issacharoff, G. Loewenstein, T. O 'Donoghue and M. Rabin, "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism'," *University of Pennsylvania Law Review* 151 (2003), 1211–1254.

Cameron, L. and M. Shah, "Risk-Taking Behavior in the Wake of Natural Disasters," *Journal of Human Resources* Forthcoming (2015).

Card, D., A. Mas, E. Moretti and E. Saez, "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction," *American Economic Review* 102 (October 2012), 2981–3003.

Carlsson, F., G. Gupta and O. Johansson-Stenman, "Keeping up with the Vaishyas? Caste and relative standing in India," *Oxford Economic Papers* 61 (January 2009), 52–73.

Carlsson, F., O. Johansson-Stenman and P. Martinsson, "Do You Enjoy Having More than Others? Survey Evidence of Positional Goods," *Economica* 74 (November 2007), 586–598.

Charness, G. and M. Rabin, "Understanding Social Preferences With Simple Tests," *The Quarterly Journal of Economics* 117 (August 2002), 817–869.

Chen, Y. and S. X. Li, "Group Identity and Social Preferences," *American Economic Review* 99 (March 2009), 431–57.

Choi, J. J., D. Laibson, B. C. Madrian and A. Metrick, "Defined Contribution Pensions: Plan Rules, Participant Choices, and the Path of Least Resistance," in *Tax Policy and the Economy, Volume 16* NBER Chapters (National Bureau of Economic Research, Inc, 2002), 67–114.

Clark, A. and C. Senik, "Who compares to whom? The anatomy of income comparisons in Europe," *Economic Journal* 120 (2010), 573–594.

Clark, A. E., P. Frijters and M. A. Shields, "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles," *Journal of Economic Literature* 46 (March 2008), 95–144.

Clark, A. E., N. Kristensen and N. Westergard-Nielsen, "Economic Satisfaction and Income Rank in Small Neighbourhoods," *Journal of the European Economic Association* 7 (04-05 2009), 519–527.

Clark, A. E. and A. J. Oswald, "Satisfaction and comparison income," *Journal of Public Economics* 61 (September 1996), 359–381.

Cole, H. L., G. J. Mailath and A. Postlewaite, "Social Norms, Savings Behavior, and Growth," *Journal of Political Economy* 100 (December 1992), 1092–1125.

Corneo, G. and O. Jeanne, "Social Organization in an Endogenous Growth Model," *International Economic Review* 40 (August 1999), 711–25.

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

Daly, M. C., D. J. Wilson and N. J. Johnson, "Relative Status and Well-Being: Evidence from U.S. Suicide Deaths," *The Review of Economics and Statistics* 95 (December 2013), 1480–1500.

Di Tella, R., S. Galiani and E. Schargrodsky, "Reality versus propaganda in the formation of beliefs about privatization," *Journal of Public Economics* 96 (2012), 553–567.

Downs, J. S., J. Wisdom, B. Wansink, and G. Loewenstein, "Supplementing Menu Labeling With Calorie Recommendations to Test for Facilitation Effects," *American Journal of Public Health* 103 (September 2013), 1604–1609.

Druckman, J. N. and K. R. Nelson, "Framing and Deliberation: How Citizens' Conversations Limit Elite Influence," *American Journal of Political Science* 47 (October 2003), 729–745.

Duesenberry, J. S., *Income, Saving and the Theory of Consumer Behavior* (Harvard University Press, 1949).

Dupor, B. and W.-F. Liu, "Jealousy and Equilibrium Overconsumption," *American Economic Review* 93 (March 2003), 423–428.

Easterlin, R. A., "Does Economic Growth Improve the Human Lot? Some Empirical Evidence," *In: David, P.A., Reder, M.W. (Eds.), Nations and Households in Economic Growth: Essays in Honour of Moses Abramowitz* (1974).

———, “Will raising the incomes of all increase the happiness of all?,” *Journal of Economic Behavior & Organization* 27 (June 1995), 35–47.

———, “Income and Happiness: Towards an Unified Theory,” *Economic Journal* 111 (July 2001), 465–84.

Eckel, C. C., M. A. El-Gamal and R. K. Wilson, “Risk loving after the storm: A Bayesian-Network study of Hurricane Katrina evacuees,” *Journal of Economic Behavior & Organization* 69 (February 2009), 110–124.

Eibner, C. and W. N. Evans, “Relative Deprivation, Poor Health Habits, and Mortality,” *Journal of Human Resources* 40 (2005).

Fehr, E. and A. Falk, “Wage Rigidity in a Competitive Incomplete Contract Market,” *Journal of Political Economy* 107 (February 1999), 106–134.

Fehr, E. and K. M. Schmidt, “A Theory Of Fairness, Competition, And Cooperation,” *The Quarterly Journal of Economics* 114 (August 1999), 817–868.

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

Fershtman, C., K. M. Murphy and Y. Weiss, “Social Status, Education, and Growth,” *Journal of Political Economy* 104 (February 1996), 108–32.

Fisman, R., S. Kariv and D. Markovits, “Individual Preferences for Giving,” *American Economic Review* 97 (December 2007), 1858–1876.

Giuliano, P. and A. Spilimbergo, “Growing Up in a Recession: Beliefs and the Macroeconomy,” *Review of Economic Studies* 81 (2014), 787–817.

Guiso, L., P. Sapienza and L. Zingales, “Time Varying Risk Aversion,” NBER Working Papers 19284, National Bureau of Economic Research, Inc, August 2013.

Haaijer, R., W. Kamakura and M. Wedel, “Mixed Logit With Repeated Choices: Households’ Choices Of Appliance Efficiency Level,” *International Journal of Market Research* 43 (2001), 93–106.

Hainmueller, J. and M. J. Hiscox, “Attitudes toward highly skilled and low-skilled immigration: evidence from a survey experiment,” *American Political Science Review* 104 (February 2010), 61–84.

Hamermesh, D. S., “Interdependence in the Labour Market,” *Economica* 42 (November 1975), 420–29.

Hanaoka, C., H. Shigeoka and Y. Watanabe, “Do Risk Preferences Change? Evidence from Panel Data Before and After the Great East Japan Earthquake,” *mimeo* (2014).

Horiuchi, Y., K. Imai and N. Taniguchi, “Designing and analyzing randomized experiments: application to a Japanese election survey experiment,” *American Journal of Political Science* 51 (2007), 669–687.

Horton, J., D. Rand and R. Zeckhauser, “The online laboratory: conducting experiments in a real labor market,” *Experimental Economics* 14 (September 2011), 399–425.

Hossain, T. and K. K. Li, “Crowding Out in the Labor Market: A Pro-Social Setting is Necessary,” *Management Science* 60 (2014), 1148–1160.

Johansson-Stenman, O., F. Carlsson and D. Daruvala, “Measuring Future Grandparents" Preferences for Equality and Relative Standing,” *Economic Journal* 112 (April 2002), 362–383.

Kahneman, D. and J. L. Knetsch, “Valuing public goods: The purchase of moral satisfaction,” *Journal of Environmental Economics and Management* 22 (January 1992), 57–70.

King, G., C. J. L. Murray, J. A. Salomon and A. Tandon, “Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research,” *American Political Science Review* 98 (2004), 191–207.

Krupka, E. L. and M. Stephens, “The stability of measured time preferences,” *Journal of Economic Behavior & Organization* 85 (2013), 11–19.

Kuziemko, I., M. I. Norton, E. Saez and S. Stantcheva, “How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments,” *American Economic Review* forthcoming (2015).

Liu, W.-F. and S. Turnovsky, “Consumption externalities, production externalities, and long-run macroeconomic efficiency,” *Journal of Public Economics* 89 (June 2005), 1097–1129.

Loewenstein, G. and P. A. Ubel, “Hedonic adaptation and the role of decision and experience utility in public policy,” *Journal of Public Economics* 92 (August 2008), 1795–1810.

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

Luttmer, E. F. P. and M. Singhal, "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy* 3 (February 2011), 157–79.

Madrian, B. C. and D. F. Shea, "THE POWER OF SUGGESTION: INERTIA IN 401(k) PARTICIPATION AND SAVINGS BEHAVIOR," *The Quarterly Journal of Economics* 116 (November 2001), 1149–1187.

Malmendier, U. and S. Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics* 126 (2011), 373–416.

Marmot, M., *The Status Syndrome: How Social Standing Affects Our Health and Longevity* (New York: Times Books, 2004).

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

Muralidharan, K. and V. Sundararaman, "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy* 119 (2011), 39–77.

Naoi, M. and I. Kume, "Explaining Mass Support for Agricultural Protectionism: Evidence from a Survey Experiment During the Global Recession," *International Organization* 65 (October 2011), 771–795.

Perez-Truglia, R., "A Samuelsonian Validation Test for Happiness Data," SSRN WP 1658747, 2015.

Rand, D. G., "The promise of mechanical turk: how online labor markets can help theorists run behavioral experiments," *Journal of Theoretical Biology* 299 (2011), 172–179.

Sacks, D. W., B. Stevenson and J. Wolfers, "The new stylized facts about income and subjective well-being," *Emotion* 12 (December 2012), 1181–1187.

Solnick, S. J. and D. Hemenway, "Is more always better?: A survey on positional concerns," *Journal of Economic Behavior & Organization* 37 (November 1998), 373–383.

Stevenson, B. and J. Wolfers, "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox," (Spring 2008), 1–87.

———, "Subjective Well-Being and Income: Is There Any Evidence of Satiation?," NBER Working Papers 18992, National Bureau of Economic Research, Inc, April 2013.

Stigler, G. J. and G. S. Becker, "De Gustibus Non Est Disputandum," *American Economic Review* 67 (March 1977), 76–90.

Sunstein, C. R. and R. H. Thaler, "Libertarian paternalism is not an oxymoron," *The University of Chicago Law Review* 70 (2003), 1159–1202.

Thaler, R. H. and S. Benartzi, "Save More Tomorrow (TM): Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy* 112 (February 2004), S164–S187.

Thaler, R. H. and C. R. Sunstein, "Libertarian Paternalism," *American Economic Review* 93 (May 2003), 175–179.

Train, K., *Discrete Choice Methods with Simulation, 2nd edition* (Cambridge University Press, Cambridge, 2009).

Veblen, T., *The Theory of the Leisure Class. reprinted 1957* (George Allen Unwin, London, 1899).

Viscusi, W. K., J. Huber and J. Bell, "Estimating discount rates for environmental quality from utility-based choice experiments," *Journal of Risk and Uncertainty* 37 (2008), 199–220.

Voors, M. J., E. E. M. Nillesen, P. Verwimp, E. H. Bulte, R. Lensink and D. P. V. Soest, "Violent Conflict and Behavior: A Field Experiment in Burundi," *American Economic Review* 102 (April 2012), 941–64.

Watson, T. and S. McLanahan, "Marriage Meets the Joneses: Relative Income, Identity, and Marital Status," *Journal of Human Resources* 46 (2011), 482–517.

Wisdom, J., J. S. Downs and G. Loewenstein, "Promoting Healthy Choices: Information versus Convenience," *American Economic Journal: Applied Economics* 2 (April 2010), 164–78.

World-Value-Survey, *WVS 2010-2012 Wave, revised master* (2012).

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

Zilinsky, J., "Perceptions of Fairness and Preferences for Redistribution after Information Exposure: Evidence from an On-line Experiment," *mimeo* (2014).

Figure 1: Information treatment for treatment group 1 (US case)

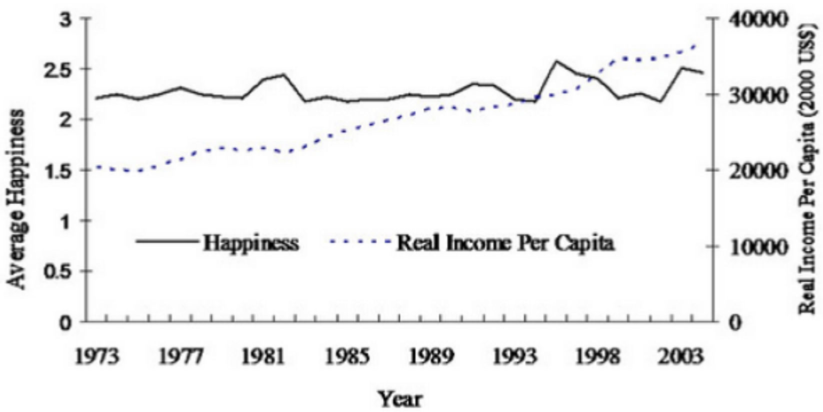
Survey on socio-economic attitudes

start end

Now we would like you to notice a scientific finding.

In the figure below, two series of data over the period 1973-2004 are reported. One is Real Income Per Capita in the United States. The other is so-called Happiness Data.

Happiness data is collected in the General Social Survey organized by National Opinion Research Center, the University of Chicago. Happiness is the average reply to the following question: *"Taken all together, how would you say things are these days? Would you say that you are...?"* The responses are coded as (3) Very Happy, (2) Pretty Happy, and (1) Not too Happy.



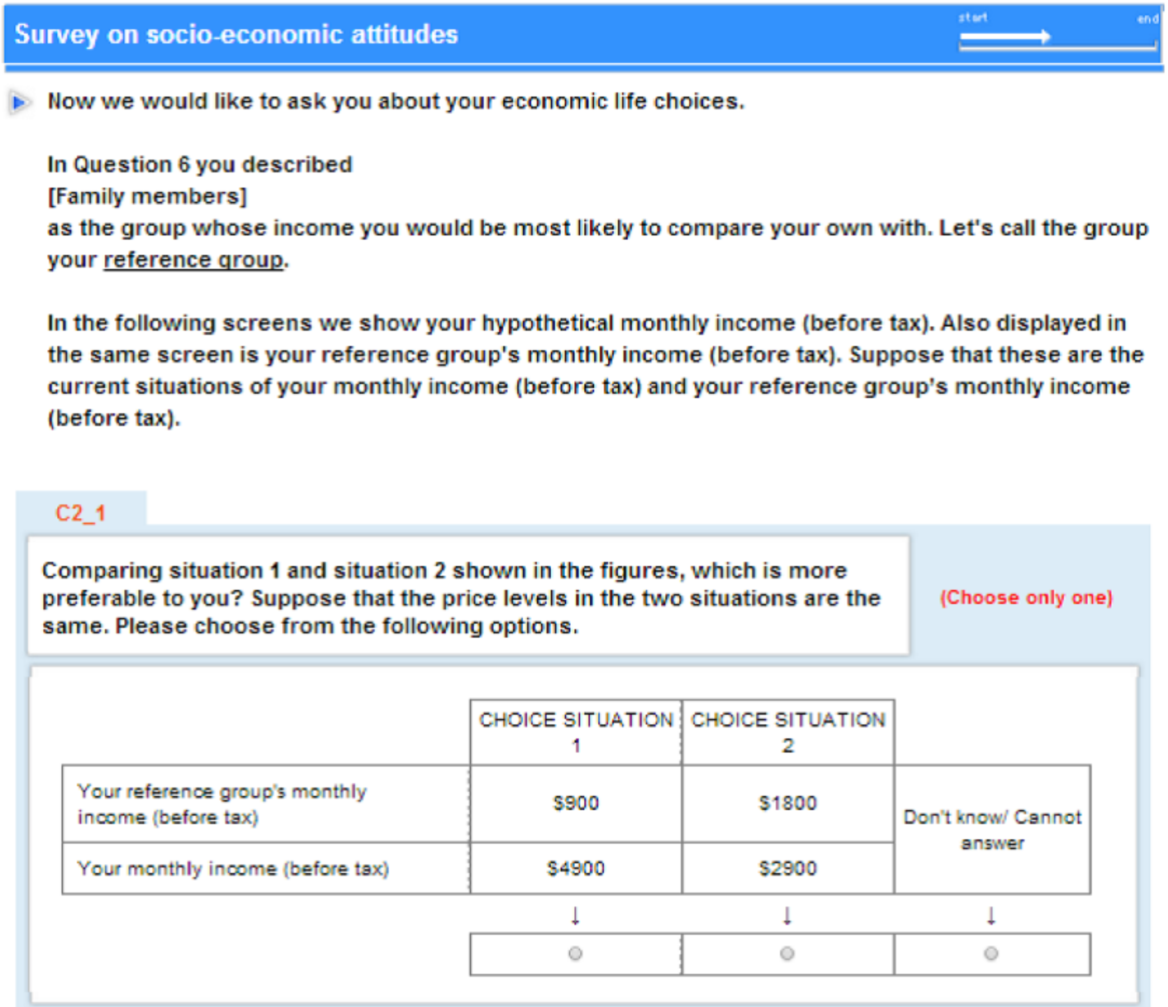
Source: World Database of Happiness and Penn World Tables.

As the graph shows, while real income per capita increases sharply, happiness shows essentially no trend and has remained constant over time.

From this figure, it looks as if individuals in the United States are in the "flat part of the curve" with additional income buying little, if any, extra happiness.

Note: For the UK case, the US figures in USD are changed to the UK figures in GBP, and all the terms “United States” in the sentences are changed to “United Kingdom”.

Figure 2: Snap shot of a discrete choice problem (the US case)



Survey on socio-economic attitudes

start end

Now we would like to ask you about your economic life choices.

In Question 6 you described
[Family members]
as the group whose income you would be most likely to compare your own with. Let's call the group your reference group.

In the following screens we show your hypothetical monthly income (before tax). Also displayed in the same screen is your reference group's monthly income (before tax). Suppose that these are the current situations of your monthly income (before tax) and your reference group's monthly income (before tax).

C2_1

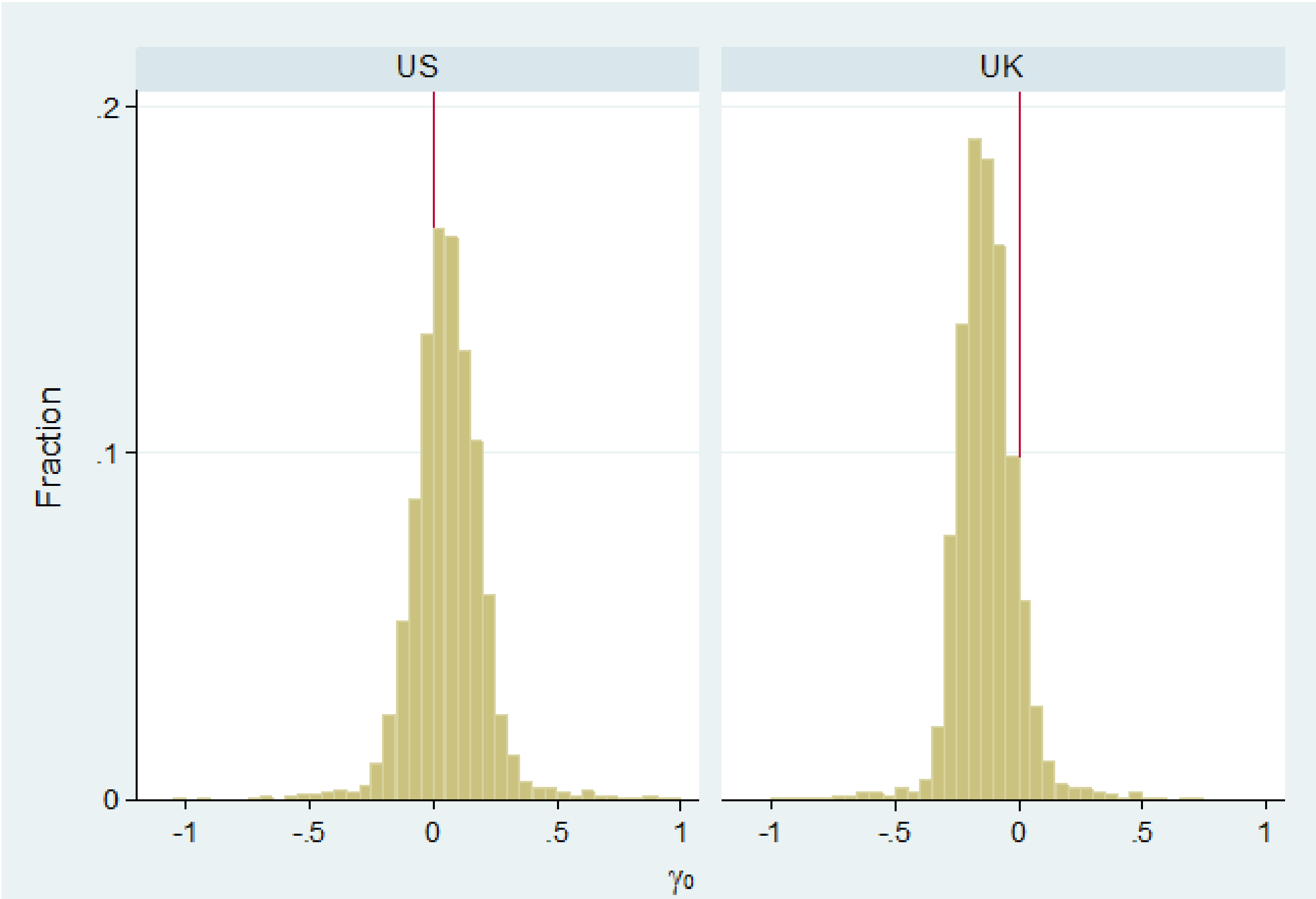
Comparing situation 1 and situation 2 shown in the figures, which is more preferable to you? Suppose that the price levels in the two situations are the same. Please choose from the following options.

(Choose only one)

	CHOICE SITUATION 1	CHOICE SITUATION 2	
Your reference group's monthly income (before tax)	$900	$1800	Don't know/ Cannot answer
Your monthly income (before tax)	$4900	$2900	
	↓	↓	↓
	○	○	○

Note: For the UK, the figures in USD are converted to GBP equivalents. On top of this question, the comparison benchmark (i.e., who to compare with) chosen by each subject is shown with the aim of making the comparison more realistic.

Figure 3: Distribution of γ_0 (baseline)



Note: The bin size is 0.05 in both graphs. The vertical line corresponds to the value of zero. The total sample size is 4,587 for the US and 4,500 for UK. For the US, the mean, standard deviation, minimum, and maximum are 0.056, 0.146, -1.019, and 0.965, respectively. For the UK, the corresponding values are -0.130, 0.126, -0.974, and 0.737, respectively. See the main text and Appendix 2 for a description of how to estimate the distribution of parameter γ.

Table 1: Comparison benchmark

		Neighbor	Classmates of your school days	Close friends	Family members	Family members of your children's classmates	Work colleagues
US							
	Observations	109	173	544	540	14	803
	Percentage	2.40%	3.80%	11.90%	11.80%	0.30%	17.50%
UK							
	Observations	52	201	689	395	27	890
	Percentage	1.20%	4.50%	15.30%	8.80%	0.60%	19.80%
			Average people in the US/UK	Friend or acquaintance	Others	Don't know	Don't compare
US							
	Observations		687	102	52	227	1,336
	Percentage		15.00%	2.20%	1.10%	4.90%	29.10%
UK							
	Observations		765	100	47	139	1,195
	Percentage		17.00%	2.20%	1.00%	3.10%	26.60%

Note: The comparison benchmark is the reference group to whom each subject compares their income. The exact question in the survey is "Whose income would you be most likely to compare your own with? Please choose one of the groups". Total sample size is 4,587 for the US and 4,500 for the UK.

Table 2: Summary Statistics

Variable	US			UK		
	No weight	Census	Weight	No weight	Census	Weight
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.521	0.495	0.495	0.496	0.483	0.482
Age (in years)	41.423		41.474	41.051		41.277
	(13.013)		(12.789)	(12.722)		(12.717)
Age (<24)	0.118	0.153	0.115	0.114	0.188	0.115
Age (25–34)	0.240	0.215	0.225	0.242	0.208	0.227
Age (35–49)	0.318	0.330	0.346	0.346	0.332	0.362
Age (50–64)	0.324	0.300	0.314	0.298	0.271	0.296
Educ 1	0.029	0.123	0.120	0.067	0.164	0.169
Educ 2	0.236	0.300	0.300	0.104	0.165	0.164
Educ 3	0.404	0.295	0.289	0.202	0.187	0.179
Educ 4	0.230	0.190	0.195	0.263	0.158	0.148
Educ 5	0.101	0.092	0.096	0.365	0.324	0.341
White	0.733		0.727	0.900		0.898
Black	0.097		0.098	0.022		0.022
Hispanic	0.099		0.113	0.002		0.002
Asian	0.044		0.036	0.052		0.048
Other race	0.028		0.027	0.024		0.031
Married	0.510		0.513	0.614		0.623
Immigrant	0.078		0.079	0.102		0.097
Unemployed	0.118		0.145	0.085		0.101
Full-time worker	0.409		0.373	0.423		0.400
Chief earner	0.616		0.592	0.585		0.566
Have kids	0.571		0.591	0.599		0.616
Log income (USD)	10.259		10.161	10.109		10.058
	(0.975)		(0.989)	(0.915)		(0.913)
Observations	4,587	-	4,587	4,500	-	4,500

Note: All figures, except for age (in years) and log income (in USD), are percentages for each category. Educ 1, 2, 3, 4 and 5 correspond to "Less than high school graduate", "High school graduate", "Some college or associate's degree", "Bachelor's degree", and " Advanced degree" in the US. In the UK, Educ 1, 2, 3, 4 and 5 correspond to "No qualifications", "Level 1 qualifications", "Level 2 qualifications", "Level 3 qualifications", and "Level 4 qualifications and above", respectively.

Table 3: Individual characteristics and γ_0 (baseline)

Outcome: γ_0 (baseline)	(1)	(2)	(3)
UK	-0.185***	-0.186***	-0.185***
	(0.003)	(0.003)	(0.003)
Male		0.005	0.005
		(0.003)	(0.003)
Age (in years)		-0.000	-0.000
		(0.000)	(0.000)
Black		-0.009	-0.008
		(0.006)	(0.006)
Hispanic		-0.008	-0.007
		(0.007)	(0.007)
Asian		-0.023***	-0.022***
		(0.007)	(0.007)
Other race		0.015	0.015
		(0.009)	(0.009)
Married		0.002	0.002
		(0.003)	(0.003)
Immigrants		0.006	0.008
		(0.006)	(0.006)
Low educated		0.002	0.002
		(0.003)	(0.003)
Unemployed		-0.002	-0.002
		(0.005)	(0.005)
Fulltime worker		-0.004	-0.003
		(0.004)	(0.004)
Chief earner		-0.002	-0.002
		(0.004)	(0.004)
Have kids		-0.006*	-0.005
		(0.003)	(0.003)
Log income (USD)		-0.000	0.000
		(0.002)	(0.002)
Intensity of comparison			-0.004***
			(0.001)
Constant	0.056***	0.064***	0.071***
	(0.002)	(0.019)	(0.019)
R-square	0.316	0.317	0.318
Observations	9,087	9,087	9,087

Note: All explanatory variables, except for age (in years) and log income (in USD), are dummy variables. The reference group for race is white. "Low educated" is less than or equal to "some college or associate's degree" in the US, and less than or equal to "level 2 qualifications" in the UK. "Intensity of comparison" can take values from 1 (don't compare) to 5 (compare intensively). Significance levels are $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table 4: Sample balance across treatments

	US				UK			
	Treatment 1	Treatment 2	Control	*p-value* (equality of all groups)	Treatment 1	Treatment 2	Control	*p-value* (equality of all groups)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
γ_0 (baseline)	0.050	0.061	0.056	0.106	-0.131	-0.127	-0.131	0.651
	(0.153)	(0.146)	(0.139)		(0.124)	(0.130)	(0.123)	
Male	0.515	0.523	0.525	0.860	0.501	0.502	0.484	0.545
	(0.500)	(0.500)	(0.500)		(0.500)	(0.500)	(0.500)	
Age (in years)	41.394	41.599	41.274	0.781	41.013	41.192	40.950	0.865
	(13.035)	(12.964)	(13.048)		(12.473)	(12.906)	(12.792)	
White	0.723	0.741	0.733	0.520	0.916	0.892	0.890	0.036**
	(0.448)	(0.438)	(0.443)		(0.278)	(0.310)	(0.312)	
Black	0.099	0.098	0.095	0.909	0.015	0.028	0.023	0.070*
	(0.298)	(0.298)	(0.293)		(0.123)	(0.164)	(0.150)	
Hispanic	0.096	0.103	0.098	0.807	0.003	0.000	0.005	0.037**
	(0.295)	(0.304)	(0.297)		(0.052)	0.000	(0.068)	
Asian	0.051	0.034	0.046	0.061*	0.045	0.054	0.057	0.314
	(0.220)	(0.182)	(0.210)		(0.208)	(0.226)	(0.233)	
Other race	0.025	0.015	0.021	0.137	0.021	0.026	0.024	0.599
	(0.156)	(0.121)	(0.143)		(0.142)	(0.160)	(0.154)	
Married	0.511	0.518	0.502	0.678	0.615	0.613	0.614	0.992
	(0.500)	(0.500)	(0.500)		(0.487)	(0.487)	(0.487)	
Immigrant	0.077	0.071	0.086	0.300	0.099	0.095	0.112	0.294
	(0.266)	(0.258)	(0.281)		(0.299)	(0.293)	(0.315)	
Low educated	0.689	0.666	0.653	0.530	0.366	0.381	0.371	0.717
	(0.463)	(0.472)	(0.476)		(0.482)	(0.486)	(0.483)	
Unemployed	0.103	0.134	0.115	0.028**	0.088	0.083	0.083	0.841
	(0.304)	(0.341)	(0.319)		(0.284)	(0.276)	(0.276)	
Full-time worker	0.403	0.406	0.418	0.659	0.429	0.415	0.424	0.723
	(0.491)	(0.491)	(0.493)		(0.495)	(0.493)	(0.494)	
Chief earner	0.618	0.608	0.623	0.673	0.604	0.582	0.568	0.131
	(0.486)	(0.488)	(0.485)		(0.489)	(0.493)	(0.495)	
Parent	0.577	0.574	0.561	0.630	0.612	0.583	0.601	0.261
	(0.494)	(0.495)	(0.496)		(0.487)	(0.493)	(0.490)	
Log income (USD)	10.259	10.267	10.252	0.914	10.116	10.086	10.125	0.470
	(0.953)	(0.977)	(0.995)		(0.947)	(0.901)	(0.897)	
Observations	1,488	1,555	1,544		1,498	1,487	1,515	

Note: Columns (4) and (8) provide the p-values of the joint test of equality. See Table 3 for abbreviations. Significance levels are $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table 5: Basic results

Outcome: γ	US		UK	
	(1)	(2)	(3)	(4)
Treatment1	0.001	0.001	-0.001	-0.001
	(0.006)	(0.006)	(0.007)	(0.007)
Treatment2	-0.003	-0.003	-0.012*	-0.012*
	(0.006)	(0.006)	(0.007)	(0.007)
Controls	No	Yes	No	Yes
Control mean	0.025	0.025	-0.127	-0.127
R-square	0.001	0.005	0.003	0.008
Observations	4,587	4,587	4,500	4,500
Treatment1 − Treatment2	-0.004	-0.005	-0.011	-0.011*
	(0.006)	(0.006)	(0.007)	(0.007)

Standard errors are reported in parentheses. All regressions control for baseline γ_0, even those labeled as including "no" controls. Controls for covariates further include a dummy for sex, age (in years), each race dummy, a dummy for married, a dummy for immigrant, each education category dummy, a dummy for unemployed, a dummy for fulltime worker, a dummy for chief earner, and a dummy for being a parent as well as log income (in USD). Control mean is the mean value of γ for the control group. The last two rows report the difference between the estimates of Treatment 1 and Treatment 2, with the standard errors in parentheses. Significance levels are $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table 6: Robustness checks (UK sample)

Outcome: γ	Baseline (whole sample)	*Time*	*Quiz 1*	*Quiz 2*
	(1)	(2)	(3)	(4)
Treatment1	-0.001	-0.002	0.002	0.007
	(0.007)	(0.007)	(0.007)	(0.007)
Treatment2	-0.012*	-0.012*	-0.012*	-0.010
	(0.007)	(0.007)	(0.007)	(0.007)
Controls	Yes	Yes	Yes	Yes
Control mean	-0.127	-0.127	-0.127	-0.127
R-square	0.008	0.007	0.010	0.015
Observations	4,500	4,088	3,974	3,567
Treatment1 − Treatment2	-0.011*	-0.010	-0.014**	-0.018**
	(0.007)	(0.007)	(0.007)	(0.007)

Standard errors are reported in parentheses. See Table 5 for list of controls. Control mean is the mean value of γ for the control group. The last two rows report the difference between the estimates of Treatment1 and Treatment2, with the corresponding standard errors in parentheses. Column (2) (*Time*) excludes observations whose duration time is below the 5th percentile or above the 95th. Columns (3) (*Quiz 1*) and (4) (*Quiz 2*) limit the sample to those who correctly answered either the first or second verification questions (88.3 percent of the sample) and those who answered correctly on the first verification question (79.2 percent of the sample). Significance levels are $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table 7: By Baseline γ_0 (UK sample)

Outcome: γ	Baseline γ_0	
	Below median (more jealous)	Above median (less jealous)
	(1)	(2)
Treatment1	0.003	-0.005
	(0.009)	(0.010)
Treatment2	-0.004	-0.020**
	(0.009)	(0.010)
Controls	Yes	Yes
Control mean	-0.151	-0.102
R-square	0.013	0.014
Observations	2,250	2,250

Note: Standard errors are reported in parentheses. See Table 5 for list of controls. Control mean is the mean value of γ for the control group. For columns (1) and (2), note that the lower the baseline γ_0, the more jealous the subject initially is. For columns (3) – (5), others' income is the subjects' expectation about the income of their comparison benchmark. Significance levels are $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Table 8: Own vs. others' income (UK sample)

	Own vs. others' income						
	$y_{survey} < \bar{y}_{survey}$			$y_{survey} = \bar{y}_{survey}$	$y_{survey} > \bar{y}_{survey}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment1	0.004	-0.008	0.016	-0.004	-0.022	-0.031	-0.040
	(0.009)	(0.013)	(0.017)	(0.011)	(0.021)	(0.028)	(0.025)
Treatment2	-0.009	-0.010	-0.006	-0.008	-0.037*	0.014	0.017
	(0.009)	(0.013)	(0.017)	(0.011)	(0.021)	(0.028)	(0.027)
Treatment1 $\times(\ln y_{survey} - \ln \bar{y}_{survey})$		-0.012				0.016	
		(0.010)				(0.034)	
Treatment2 $\times(\ln y_{survey} - \ln \bar{y}_{survey})$		-0.002				-0.099***	
		(0.010)				(0.036)	
Treatment1 $\times(y_{survey}/\bar{y}_{survey})$			-0.026				0.009
			(0.031)				(0.008)
Treatment2 $\times(y_{survey}/\bar{y}_{survey})$			-0.006				-0.028***
			(0.031)				(0.009)
Control mean	-0.124	-0.124	-0.124	-0.137	-0.111	-0.111	-0.111
R-squared	0.017	0.018	0.017	0.010	0.038	0.054	0.061
Observations	2,417	2,405	2,405	1,598	485	485	485

Note: y_{survey} is own income, and $\bar{y}_{survey}$ is the subjects' expectation of income for their comparison benchmark. Standard errors are reported in parentheses. See Table 5 for a list of controls. Control mean is the mean value of γ for the control group. Significance levels are $*p < 0.10$, $**p < 0.05$, and $***p < 0.01$.

Appendix A1 Flow chart of questionnaire

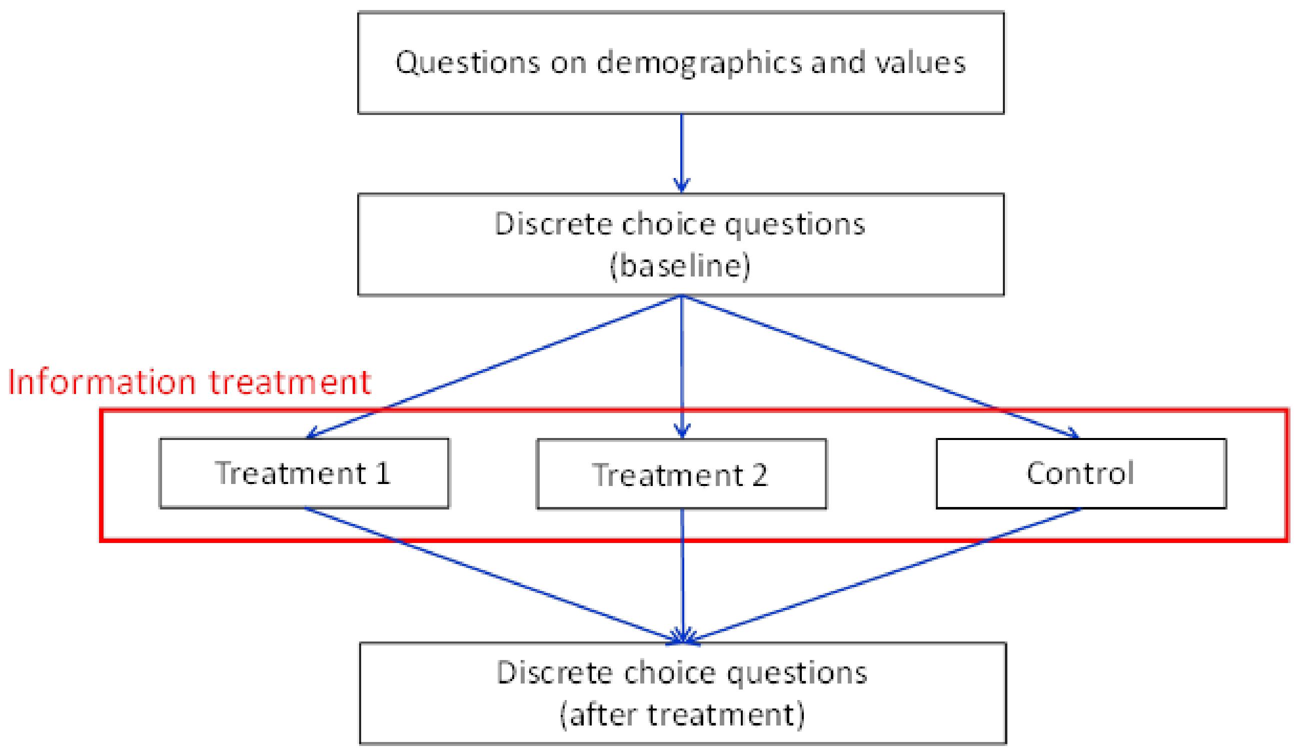
The structure and the flow of the experiment is as follows.

1. Background socioeconomic questions including age, gender, race, education, income, and work status

2. First discrete choice questions, to measure baseline income-comparison attitudes before information treatment

3. Randomized treatment, providing information on income and happiness shown solely to the treatment group and unrelated information to control group (as shown in Figure A1)

4. Second discrete choice questions, to measure income-comparison attitudes after information treatment



Appendix A2 Technical appendix

Appendix A2.1 Construction of hypothetical discrete choice question

First, we explain how we constructed our choice questions. Using the monthly income distributions in the US and the UK as bases, five levels of attributes are set. The two *attributes* (own income and reference income) and five possible income levels for each attribute provide 25 potential variations in the income situation scenario. In the literature of choice experiments, these scenarios are called *alternatives*.

Researchers have to make their own choices about which alternatives to use in survey questions and which to discard. Following Louviere et al. (2000), we conducted *orthogonal planning* in choosing the alternatives to be used in choice questions. This method effectively pairs multidimensional and multiple-level attributes in alternatives and provides an experimental plan with the greatest amount of information using the least number of observations. Further, by employing orthogonal planning, we can avoid multicollinearity problems in the regressions of the random utility model, which is explained later, because the independent variables in the regressions become orthogonal. We used SPSS Conjoint (ver. 20.0) for orthogonal design of alternatives in this study.

Next, we constructed *choice sets* consisting of pairs of alternatives and the no-choice option. By the requirement of orthogonal design, we generated two different alternative vectors, each of which consists of 25 pairs of own income levels and reference income levels.[1] Finally, choice sets, with a no-choice option, are created by pairing two alternatives, one of which is taken from one alternative vector and the other of which is chosen from the other alternative vector. The pairing strategy is at the discretion of the researchers, but all the variations must be used and the same alternative cannot be used twice. Because the orthogonality in the alternative matrix is maintained for each row permutation, we can arbitrarily pair alternatives to meet the requirement.

As documented in Huber and Zwerina (1996) and Viscusi et al. (2008), it is ideal if the choice design can be paired so as to balance the utility of each alternative. One difficulty in choice experiments of income comparisons, however, is that an increase (resp., decrease) in a reference group's income does not necessarily mean that there is a decrease (increase) in one's own utility level; as such, we did not exclude the possibility of altruistic preference. Given these constraints, our best strategy for pairing alternatives is as follows.

Suppose we have the scenario $S = (x, y)$, where x denotes the level of one's own income and y is others' income. Then, qualitatively, candidates of paired scenarios consist of the following eight variations: $(x, y+)$, $(x, y-)$, $(x+, y)$, $(x+, y+)$, $(x+, y-)$, $(x-, y+)$, $(x-, y)$, and $(x-, y-)$, where

[1] As such, in the social average task, 25 alternatives out of 25 potential variations *had to be used* to meet the requirement of orthogonal design.

$x+$ means some value greater than x and $x-$ means some value smaller than x, and similarly for y. We do not exclude the possibility of altruism a priori, so there are no a priori dominant choices for S from these eight alternatives. We then made pairs such that these eight situations appear as evenly as possible. Using the procedures discussed here, we were able to efficiently obtain parameter estimates. Table A1 shows the set of questions we used in the survey. Each respondent answered 6 randomly assigned questions out of the 25 total questions.

Appendix A2.2 Random utility model and empirical method

Here, we introduce the econometric foundation of how subjects' choice data can be used to estimate the sign of coefficients and intensity of income comparisons. To analyze decisions in hypothetical choice experiments, we use a random utility model framework. It is assumed that subjects choose an income situation because they obtain higher utility from that situation than from the other available situations. When there are two situations (A and B, for example as in this study), and if they chose A rather than B, then the choice data is recorded as 1 for A and 0 for B, along with the levels of the explanatory variables (own income level and reference income level in this study). These pieces of information constitute the observation for regression analyses.

Now, we assume N subjects who answer $T(\geq 1)$ repeated choice questions. The utility of subject n when s/he chooses situation i at question $t \in T$, U_{itn} consists of observable components of the explanatory variables V_{itn} and unobservable components ϵ_{itn}, so that utility can be viewed as $U_{itn} = V_{itn} + \epsilon_{itn}$. Utilities from observable components are assumed to be linear combinations of each variable as $V_{itn} = \sum_{k=1}^{K} \beta_k X_{ik}$, where $k = 1, \cdots, K(K \geq 2)$ represents the variety of explanatory variable ($K = 2$ in this study), X_k denotes the levels of the kth explanatory variables, and β_k measures the marginal utility of each variable. In the following analysis, the vector $\beta \equiv (\beta_1, \cdots, \beta_K)$ that maximizes the log likelihood function of observed choice patterns by subject is the estimator for mixed logit model regressions. Following McFadden (1974), ϵ_{itn} is distributed according to an independent and identical distribution of extreme value type 1 (IIDEV1) with variance σ^2.

The logit formula of choice probability P_{itn} for subject n choosing situation i from the set of situations S_t (the choice set) in question $t \in T$ can be written as

$$P_{itn} = prob(U_{itn} > U_{jtn}, \forall j \neq i \in S_t) = prob(\epsilon_{jtn} - \epsilon_{itn} < V_{itn} - V_{jtn}, \forall j \neq i \in S_t).$$

In McFadden (1974), it was shown that $P_{itn} = \exp(\lambda V_{itn}) / \sum_{j \in S} \exp(\lambda V_{jtn})$, where $\lambda = \pi/\sqrt{6}\sigma$ is a scale parameter.

Finally, a dummy variable d_{itn} is defined, taking a value of 1 if subject n chooses situation i

for question $t \in T$, and 0 otherwise. Together with the logit formula of choice probability P_{itn}, the log likelihood function of repeated choices observed in experiments can be written as

$$LL(\beta) = \sum_n \sum_t \sum_{i \in S_t} d_{itn} \ln P_{itn}.$$

In this paper, we consider a case where the independence of irrelevant alternatives (IIA) does not hold.[2] We then obtain distributions $f(\beta_i)$ of some parameters in β across subjects by the following method. We specify that $f(\beta)$ is a normal distribution function with its parameter set as θ, following Train (2009). The choice probability function P_{itn}^{ML} for the mixed logit model can be written as

$$P_{itn}^{ML} = \int P_{itn}(\beta) f(\beta|\theta) d\beta,$$

where P_{itn} is the logit choice probability in the conditional logit model given β. The value of θ can be obtained via simulation to maximize the simulated log likelihood function[3]

$$SLL(\theta) = \sum_n \sum_t \sum_{i \in S_t} d_{itn} \ln P_{itn}^{ML}.$$

Next, we specify the shape of the utility function for our purpose here. Individuals derive utility not only from their own income $X_1 = y$ but also from the income of those in the comparison benchmark $X_2 = \bar{y}$. From textbook assumptions, we suppose that subjects value attribute y positively. We consider the constant relative risk-aversion-type utility function as

$$V = \frac{(y\bar{y}^{\gamma})^{1-\rho}}{(1-\rho)}, \tag{1}$$

where $\rho > 0$. The parameter γ regulates the intensity and sign of relative utility and is the central topic of this study. If $\gamma < 0$, then the individual has jealousy. If $\gamma > 0$, then the individual has an altruistic preference, whereas if $\gamma = 0$, there is no income comparison. We take the logarithms of both sides in Equation 1 to obtain

$$\ln V_{ni} = (1-\rho) \ln y_{ni} + (1-\rho)\gamma \ln \bar{y}_{ni} - \ln(1-\rho). \tag{2}$$

Using maximum-likelihood estimation, we obtain the point estimates for $\ln y$ and distributions of the coefficients for $\ln \bar{y}$.[4] Finally, we obtained individual parameters of income comparisons γ by dividing the individual coefficient of $(1-\rho)\gamma$ by the average effect of $(1-\rho)$. The individual

[2]We assume that the error term is independently and identically distributed, as in the conditional logit model.
[3]See Section 6 of Train (2009) for details.
[4]See Revelt and Train (1998) for details on this procedure.

coefficient of $(1-\rho)\gamma$ can be obtained by applying the inverse Bayesian formula to the estimated distribution of the coefficient and the choice pattern of the individual subject (Train 2009).

References

Huber, J. and K. Zwerina, "The importance of utility balance in efficient choice designs," *Journal of Marketing Research* 33 (1996), 307–317.

Louviere, J. J., D. A. Hensher and J. D. Swait, *Stated Choice Methods: Analysis and Applications* (Cambridge University Press, 2000).

McFadden, D., *Conditional Logit Analysis of Qualitative Choice Behavior in P. Zarembka (ed.) Frontiers in Econometrics* (Academic Press, 1974).

Revelt, D. and K. Train, "Mixed Logit With Repeated Choices: Households' Choices Of Appliance Efficiency Level," *The Review of Economics and Statistics* 80 (November 1998), 647–657.

Train, K., *Discrete Choice Methods with Simulation, 2nd edition* (Cambridge University Press, Cambridge, 2009).

Viscusi, W. K., J. Huber and J. Bell, "Estimating discount rates for environmental quality from utility-based choice experiments," *Journal of Risk and Uncertainty* 37 (2008), 199–220.

Figure 1: A1: Information treatment for control group (US case)

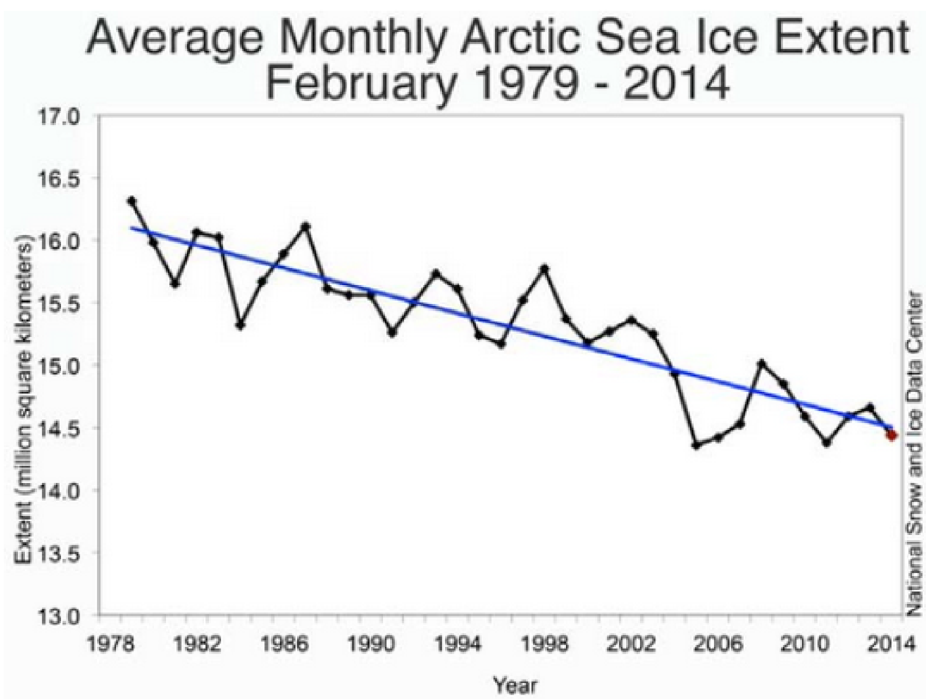
Survey on socio-economic attitudes

start end

Now we would like you to notice a scientific finding.

In the figure below, a time series of data concerning the Arctic Sea ice extent in February over the period 1979-2014 is presented.

According to a recent report by the National Snow & Ice Data Center, University of Colorado, on March 3, 2014, the Arctic sea ice extent in February 2014 measured 14.44 million square kilometers (5.58 million square miles). This is the fourth lowest February ice extent in the satellite data record, and is 910,000 square kilometers (350,000 square miles) below the 1981 to 2010 average.



Source: National Snow & Ice Data Center

<u>The sea ice extent trend through February 2014 is -3.0% per decade relative to the 1981 to 2010 average, a rate of -46,100 square kilometers (-17,800 square miles) per year. These winter month trends have been fairly consistent.</u>

Table 1: A1: Choice sets of discrete income choice questions

	Situation 1		Situation 2	
	Own income	Ref. income	Own income	Ref. income
Q 1	4	3	3	1
Q 2	5	5	3	5
Q 3	5	4	4	5
Q 4	5	1	5	3
Q 5	5	3	5	4
Q 6	5	2	5	1
Q 7	2	3	4	4
Q 8	4	4	5	2
Q 9	3	3	4	3
Q 10	4	2	1	1
Q 11	3	2	2	2
Q 12	4	1	3	2
Q 13	3	5	3	4
Q 14	4	5	5	5
Q 15	1	4	2	3
Q 16	2	5	1	4
Q 17	2	1	4	2
Q 18	1	5	3	3
Q 19	2	2	4	1
Q 20	1	2	2	5
Q 21	1	3	1	2
Q 22	3	4	2	4
Q 23	3	1	1	5
Q 24	1	1	1	3
Q 25	2	4	2	1

Note: 1= USD900 (GBP1,000), 2= USD1,800 (GBP1,500), 3= USD2,900 (GBP2,000), 4= USD4,900 (GBP3,000), and 5= USD7,200 (GBP4,000). Each subject answers 12 randomly assigned questions (6 questions before information treatment and 6 after information treatment) out of the 25 possible questions.