NBER WORKING PAPER SERIES

INCOME-COMPARISON ATTITUDES IN THE US AND THE UK: EVIDENCE FROM DISCRETE-CHOICE EXPERIMENTS

Hitoshi Shigeoka
Katsunori Yamada

Working Paper 21998
http://www.nber.org/papers/w21998

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2016

Previously titled as "Can We Steer Income-comparison Attitudes by Information Provision?: Evidence from Randomized Survey Experiments in the US and the UK." Yamada acknowledges financial support from JSPS (24683006). We are grateful to Glenn Harrison, Hideki Hashimoto, Frank Heinemann, Chen-Ying Huang, Hidehiko Ichimura, Michihiro Kandori, Erik Kimbrough, Andrew McGee, Rosemarie Nagel, Krishna Pendakur, Yasuyuki Sawada, Yasutora Watanabe, and seminar participants at Academia Sinica, Barcelona GSE Summer Forum, GRIPS, HKSTU, Hokkaido University, Kobe University, Kyoto University, National Taiwan University, Nangyang Technological University, Osaka Prefecture University, Osaka University, Simon Fraser University, Tohoku University, University of Tokyo, and Yokohama National University for their comments. Any remaining errors are ours. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Income-comparison Attitudes in the US and the UK: Evidence from Discrete-choice Experiments
Hitoshi Shigeoka and Katsunori Yamada
NBER Working Paper No. 21998
February 2016
JEL No. C9,D1,D3

ABSTRACT

Economists have long been aware of utility externalities such as a tendency to compare own income with others'. If welfare losses from income comparisons are significant, any governmental interventions that alter such attitudes may have large welfare consequences. We conduct an original online survey of discrete-choice questions to estimate such attitudes in the US and the UK. We find that the UK respondents compare incomes more than US respondents do. We then manipulate our respondents with simple information to examine whether the attitudes can be altered. Our information treatment suggesting that comparing income with others may diminish welfare even when income levels increase makes UK respondents compare incomes more rather than less. Interestingly, US respondents are not affected at all. The mechanism behind the UK results seems to be that our treatment gives moral license to make income comparisons by providing information that others do so.

Hitoshi Shigeoka
Department of Economics
Simon Fraser University
8888 University Drive, WMC 4653
Barnaby, BC V5A 1S6
and NBER
hitoshi_shigeoka@sfu.ca

Katsunori Yamada
Faculty of Economics
Kindai University
228-3 Shin-Kami-Kosaka
Higashi-osaka City, Osaka
Japan
kyamada@kindai.ac.jp

1 Introduction

Economists have long been aware of utility externalities such as a tendency to compare own income with others' (e.g., Duesenberry 1949; Akerlof and Yellen 1990). If welfare losses from income comparisons are significant, any governmental interventions that alter such attitudes may have large welfare consequences.[1] Indeed, recent studies have accumulated empirical support that relative income 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).

There are two objectives in this paper. Firstly, we measure baseline income-comparison attitudes for the US and the UK. Toward this end, we conduct an original online survey using nationally representative sample of each country, which is an important virtue of our study. Drawing macroeconomic pictures of income comparisons between countries are beneficial: An elaborate work by Cole et al. (1992) shows that differences in social preference can explain differences in the macroeconomic performances of two countries with very similar economic fundamentals, such as the US and the UK. Our work on these two countries offers a unique opportunity to test the presumptions of Cole et al. (1992).

Estimation of income-comparison attitudes for different countries in a way that allows direct inter-country analysis is challenging. A 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?"[2] Despite its apparent simplicity, inter-country analysis based on such subjective evaluations can be problematic when these countries do not share the same assessment rules or 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.

Our novelty is that we estimate income-comparison attitudes for different countries in a way that makes comparison direct and meaningful.[3] Concretely, we first set up a behavioral model

[1] In order to investigate if governmental interventions could be beneficial, researchers have to establish a welfare criterion. See Loewenstein and Ubel (2008) and Benjamin et al. (2014) for thoughtful discussions on this point. Our strategy in this paper is rather simple, in that we assume that welfare loss due to income comparisons can be characterized with a utility function parameter. It is possible that income comparisons motivate people to become more competitive, thereby increasing effort toward improving income level. We argue that these positive effects of income comparison are general equilibrium effects. In this paper, however, we focus on partial equilibrium aspects of income comparisons, keeping factors fixed other than the immediate welfare consequences due to income comparisons.

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

[3] Clark and Senik (2010) use subjective evaluation data on attitudes of income comparison from the European Social Survey (Round 3), but information for the US is missing. Question V96 on the World Values Surveys 2010–2012 asks about the perception of income equality and egalitarianism, but while this question captures one aspect

of income comparisons. In the model, utility function is linked to constant relative risk aversion (CRRA) (Gali 1994; Abel 1990; Dupor and Liu 2003) and others' incomes work as negative externalities.[4] We then estimate this parameter of utility externality with data from hypothetical discrete choices questions, and respondents choose among two alternative combinations of hypothetical monthly income scenarios for themselves and their reference persons. Since we estimate a utility function parameter, it is cardinal and thus comparable between countries.

After measuring baseline income-comparison attitudes for the two countries, our second objective is to investigate whether such attitudes can be altered by a simple information provision. Our information provision is conducted on an experimental basis within the same survey. We randomly assign respondents into treatment groups and a control group. We inform treatment groups that people tend to compare their incomes with those of others, and that such behavior may diminish happiness even when absolute income is increasing. After the information provision, income-comparison attitudes are again measured with discrete-choice questions as in the baseline measurement. We then examine whether the intervention affects income comparisons attitudes in a desired direction.

Our results are as follows. First, for the baseline country differences before information intervention, we find that respondents from the UK exhibit much stronger income-comparison attitudes (i.e., are more jealous) than those from the US, possibly indicating that welfare loss from 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 for observable individual characteristics, such as age, gender, education, and own income. While we cannot control for other institutional and cultural differences between two countries that are not observed in the data, we provide the first evidence of differences in income-comparison attitudes between the US and the UK measured by a discrete-choice approach. Our results are broadly consistent with Alesina et al. (2004), who documented differences between the US and the UK in attitudes toward income redistribution.

Secondly, we find that information treatment *strengthens* the tendency for UK respondents to compare incomes (i.e., become more jealous). This result indicates that our information treatment guided UK respondents in a direction that may reduce welfare. Interestingly, US respondents are not affected by the intervention. Recall that at baseline UK respondents compare more than do US respondents, so UK respondents are more jealous than US respondents from the beginning

of social preferences, it does not address the intensity of income comparison.

[4]This setting implies that we have to exclude strategic components among respondents, so the respondents did not interact in our survey. Also note that income comparison is one of many forms of social preference. Other aspects 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 of discrete-choice experiments differs from that of most literature on social preferences using game-theoretical frameworks.

and become even more jealous after the information intervention. Hence, country differences are present in not only absolute levels of income-comparison attitude but also responses to an 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 a 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 contents of the information provided. Such sample selections hardly affect our estimates.

Finally, we divide the UK sample by the baseline income-comparison attitudes to shed light on the mechanism behind information treatment guiding UK respondents in an undesirable direction. We find that our results are driven by those who are initially less comparison-conscious, not by those who are initially jealous, implying that our information treatment raises the salience of income comparison rather than making people realize its negative consequences. Consequently, our information treatment might have provided moral license for comparing income with others by informing that other people do so.

Overall, our findings suggest that income-comparisons attitudes are relatively stable against our information manipulation, but also that there is some scope for governmental intervention among a certain type of population. Our results also imply that affecting income comparison attitudes in a desired direction can be difficult. Because our information intervention is not decisive, given our results, the government may better urge people to stop comparing incomes for the sake of better social welfare (i.e., "you should not covet").[5] 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.

Our work is closely related to the literature on utility externality in macroeconomies. Abel (1990), Gali (1994), and Dupor and Liu (2003) consider CRRA-type utility functions with comparison externalities, and investigate the asset pricing implications. There are also studies considering income comparisons expressed in the form of "degree of positionality" and investigating macroeconomic implications of such externalities on social welfare (Aronsson and Lofgren 2008), public goods provision (Aronsson and Johansson-Stenman 2008, 2014b), optimal taxes (Aronsson and Johansson-Stenman 2010), and the sustainability of economies (Aronsson and Johansson-Stenman 2014a). Finally, Smets and Wouters (2003, 2007) consider

[5] In this sense, our approach is similar to a paternalistic mode of intervention (Thaler and Sunstein 2003; Sunstein and Thaler 2003), as we leave interpretation of the information to its recipients.

comparisons in the form of external habit in the framework of dynamic stochastic general equilibrium models.

This paper is also 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; Zilinsky 2014; Kuziemko et al. 2015).[6] Cruces et al. (2013) provided information on true income ranking and examined its effect on the support for redistribution. Kuziemko et al. (2015) also provided information on income inequality in the US, and investigated its impact on a wide variety of redistributive policy outcomes.[7]

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 respondents. 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. In our survey, respondents between the ages of 20 and 65 are selected by stratified random sampling such that the cohort profiles of our sample intend to mirror 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 our findings from the information treatment. Section 5 discusses alternative behavioral models, and Section 6 concludes.

2 Survey Experiment

2.1 Data collection

Our original survey experiment consists of three parts. See Appendix A1 for a flow chart.[8] The first part starts with twelve questions about socio-economic variables such as age, gender, race, educational attainments, and total pre-tax income in 2013. For the sake of clarity, the exact phrases for these first twelve questions were taken from the World Values Survey (2012) whenever available.

[6] A few lab experiments have also examined the effects of information treatment on individual preferences. For example, Benjamin et al. (2010) made ethnic identity salient for laboratory respondents 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. Also, using non-laboratory data, recent literature suggests that individual 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), or changes in macroeconomic events and household-level labor market outcomes (Krupka and Stephens 2013).

[7] As in these previous studies, it is not our goal to determine whether changes in income-comparison attitudes are the consequences of changes in deep preference themselves or in information sets with the deep preference fixed.

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

This is followed by seven questions about respondents' subjective values. These 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 (hereafter, the "comparison benchmark"), and the predicted income of the 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 discrete-choice questions as our 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 to examine whether the information treatment impacts such attitudes.

Our online survey experiment was carried out from March 24 to April 3, 2014 (11 days). As clearly documented in Horton et al. (2011), "[t]he validity of economics experiments depends heavily upon trust, particularly respondents' 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 respondents, 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 excludes so-called "mindless respondents," 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).

For each country, respondents between the ages of 20 and 65 are selected, using stratified random sampling such that the cohort profile of the sample intended to mirror national census statistics on age, gender, education, and geographical distribution. 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 and one control group, respondents are categorized into three groups for each country, so the sample size of each group is around 1,500 people.

In the invitation email sent to registered candidates, we specify that the survey is being conducted for research purposes and that the anonymity of respondents is secured. Note that the title of the survey is "Survey on socio-economic attitudes," which is very generic and does not suggest anything about income comparisons. We also inform respondents that the survey should take about 15 min 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. 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 completing 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 min on average, and thus payment is equivalent to an average hourly wage of roughly USD6.7 for US respondents and GBP4.5 for UK respondents.

We take several steps to ensure the validity of the results. First, we limit the sample to only residents of the US and the UK by asking respondents if they are US or UK residents at the beginning of the survey. Second, to prohibit respondents from mindlessly skipping through the questions, we enforce answering all but one question by not allowing respondents to proceed to the next question unless they provide an answer.[9] Third, to avoid mindless respondents or those who left the survey open for days, we exclude such respondents in one of the robustness checks in Section 4.4. Fourth, 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. We also limit the sample to those who correctly answered the verification question in Section 4.4.

2.2 Information treatment

Our objective for information treatment is to alert our respondents to the existence of the hedonic treadmill of income comparisons. To this end, we prepared a figure on the famous "Easterlin paradox," which indicates that income growth is not accompanied by an increase in happiness score. Then, we follow the literature that suggests 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.[10] Additionally, the figure itself is easy to understand and should not require high cognitive skill, even among general respondents who are unfamiliar 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 provided only the Easterlin paradox figure, accompanied by the following very short descriptions of the observation.

[9]Respondents could skip question 7A, 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 respondents' expectations of others' income. However, this exception does not seem to pose any issues since there were only 11 (12) people who did not answer this question in the US (UK) sample.

[10]This "paradox" is challenged by Stevenson and Wolfers (2008), and Sacks et al. (2012). However, the scientific validity itself is not an issue for our research.

[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.

Figure 1 displays a screenshot of our first treatment for the US case. We used 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 two sentences 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 of such income comparisons. It is, hence, 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 the 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 raising the salience of income comparisons and hence giving a moral license to compare with others by informing that others do so, 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 treatment groups and the control group, and to account for fatigue effects in the survey experiment.

After the information treatment, we ask a simple verification question to ensure that the respondents understand the provided information. For Treatment 1 and Treatment 2, the question

is "While real income per capita has increased sharply, the 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). For those who incorrectly answer the verification 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% of respondents answer the verification question correctly on the first attempt, with 87.2% answering correctly after the second attempt.[11] Importantly, the accuracy rates among treatments are very similar: the accuracy rates for the first attempt are 76.8%, 78.7%, and 74.4% for Treatment 1, Treatment 2, and the control group, respectively, and the accuracy rates after the second attempt are 87.6%, 88.5%, and 85.3%. We did not repeat this verification question more than once. Although this evidence 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. Among those who initiated the survey, the attrition rate as a whole was 15.3% for the US and 17.9% for the UK, below the 22% attrition rate using MTurk in Kuziemko et al. (2015). Among those who withdrew participation before completion, only 12.7% (13.3%) of respondents did so after they received the information treatment in the US (UK) case. Also, attrition that occurred after the treatment does not seem to qualitatively differ by type of treatment. Specifically, among 106 (131) respondents who initiated participation but withdrew after the information treatment in the US (UK) case, 30.0% (27.4%) belonged to Treatment 1, 34.0% (35.9%) to Treatment 2, and the rest, 35.8% (36.6%), to the control group. The slightly lower attrition rate for Treatment 1 may reflect that the instructions were shorter by two sentences than those in Treatment 2.[12] Unfortunately, we do not have information on individual characteristics for those who withdrew from the survey, meaning we cannot examine differential attrition by individual covariates, but attrition does not seem to drive our overall results.

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

[12] 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.

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). 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).[13]

To enhance the reality in making choices over hypothetical income scenarios, we elicit information on each respondent's comparison benchmark and impose the answers in the screens on discrete-choice questions.[14] We are aware that there are both advantages and disadvantages to making respondents choose their comparison benchmarks preceding discrete-choice questions. While doing so increases the reality of income comparisons in hypothetical questions, it also forces the respondent to picture comparison situations, which may induce respondents to think about income comparisons more strongly than they normally do. We choose to emphasize reality, because mindless responses could be an issue for an online survey. Also, as a robustness check, we add dummies for each comparison benchmark to control for the choice of comparison benchmark in the regression below. The results are virtually unchanged.

Table 1 presents a 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% for US respondents and 27% 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%.[15] These top-five choices

[13]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.

[14]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 US (UK); 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, we encourage an income comparison with a reference group of "someone like you," suggesting to those 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."

[15]One may think that respondents who choose family members as their comparison benchmark may try to maximize the total income rather than to compare income levels. In fact, their comparison attitudes are slightly lower (i.e., less jealous) than those who chose other categories as the comparison benchmarks. The difference, however, is not statistically significant (results available upon request). Furthermore, to mitigate such concerns, we

make up 85.3% (87.5%) 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 respondents 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 subsequent screens, we ask respondents to answer each hypothetical discrete-choice question by choosing between two situations with each situation defined by two attributes: monthly pre-tax income of respondents and that of the comparison benchmark.

> *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.*

As a snapshot of such an example, Figure 2 asks respondents to choose among situation 1, where own income is USD900 and reference group income is USD4900, and situation 2, where own income is USD1800 and reference group income is USD2900. These income levels are chosen on the basis of actual pre-tax monthly income quintiles in the US (UK): USD900 (GBP1,000), USD1,800 (GBP1,500), USD2,900 (GBP2,000), USD4,900 (GBP3,000), and USD7,200 (GBP4,000).[16]

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. Importantly, following Louviere et al. (2000), we maintained orthogonality in choosing alternatives choices in discrete-choice questions. This method effectively pairs multi-dimensional and multiple-level attributes in alternatives, and provides the greatest amount of information using the least number of observations. 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 the relative importance of own and reference person's income, that is, utility function

exclude those respondents (11.80% in the US and 8.80% in the UK) from our main analyses, but as shown later the estimates are not affected.

[16]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.

parameters, in the random utility model framework introduced by McFadden (1974). Specifically, we assume that respondents choose the income situation that offers them higher utility. Using choice data from the income scenarios, we estimate the trade-offs between marginal utility from own income and that from income of the reference person. The trade-offs are linked to a parameter of the constant relative risk aversion (CRRA)-type utility function V^m with which the effects of income comparisons, γ, appear as a negative externality of others' status. V^m is given by

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

where y is own income level and $\bar{y}$ is income level of others.

The parameter γ reflects the attitude of income comparisons. Jealousy is reflected as $\gamma < 0$, and altruism as $\gamma > 0$. When $\gamma = 0$, there is no income comparison.

The distribution of the relative importance of own and reference person's income, or equivalently γ, is estimated by using the mixed logit model framework (Train 2009). We first estimate the economy-wide distribution of γ of all respondents in such a way that the likelihood of observing all the choice patterns is maximized. Then, the value of γ for each respondent is obtained by finding the point in the distribution corresponding to each respondent from one's own choice patterns (i.e., Bayesian reverse formula). We repeat this procedure separately before and after the information treatment. Appendix A2 gives full technical details for the estimation method.

Finally, we document the advantages of our discrete-choice approach over the conventional approach with subjective evaluation data, where the intensity of income comparison is measured from a direct question such as "How important is it for you to compare your income with other people's incomes?".[17] First, there is no straightforward way of making meaningful comparison across individuals on the basis of subjective evaluation unless respondents share a common rule for assessment and a common scale (e.g., a score of 3 from one person and a score of 6 from another person may represent equivalent feelings between respondents).[18] In fact, King et al. (2004) point out the incomparability of answers to subjective evaluations in inter-country surveys. Second, many people dislike thinking of themselves as comparison-conscious, and they therefore underestimate the degree to which they care about others' income (social-desirability bias). Third, it is difficult to draw any cardinal inferences from ordinal survey responses to subjective evaluations. In contrast, because we measure a utility function parameter for each respondent, we can safely

[17] See also Solnick and Hemenway (1998), Johansson-Stenman et al. (2002), Alpizar et al. (2005), Carlsson et al. (2007), Mujcic and Frijters (2013), and Mujcic and Frijters (2015) for studies investigating the intensity of social comparisons via hypothetical choice questions. Carlsson et al. (2009) conducted discrete-choice experiment on income-comparisons with repeated choice questions just as the current study, although their interest was not in the distribution of income-comparison attitudes.

[18] See Daly et al. (2013) for a critical view of subjective evaluations, in particular happiness studies.

compare income-comparison attitudes between respondents, and it is straightforward to draw cardinal inferences about welfare from our results. One caveat to our approach is that estimation of income-comparison attitudes depends on the specification of our behavioral model. We address different types of utility functions in Section 5.

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 for both samples. Both samples over-represent white respondents (at 73% and 90% in the US and the UK, respectively). More than half of respondents are married, and 7.8% (10.2%) 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, a common issue when using online surveys, including MTurk (see e.g., Horton et al. 2011; Kuziemko et al. 2015.) 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 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 this 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 differences in the intensity of

income-comparison attitudes between the US and the UK.

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 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 respondents in the US do not care much about others' income. 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. Also a two-sample Kolmogorov–Smirnov test for equality of distribution functions between the US and the UK sample is rejected at the 1% level.

To examine whether the between-country difference in income-comparison attitudes is simply a consequence of sample differences between the US and the UK, we include 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).[19] Column 3 further adds dummies for each comparison benchmark to control for the choice of reference group. The coefficient on the UK dummy is unchanged.[20]

We also confirm that this US–UK difference does not stem from differences in absolute income levels chosen in the discrete-choice questions across countries. Specifically, we re-estimate γ_0, 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).

These results are consistent with Alesina et al. (2004), who show that Americans do not care

[19]This result echoes the finding in Yamada and Sato (2013), which examined income-comparison attitudes of Japanese respondents and obtained $\gamma = -0.458$. Nonetheless, caution is needed for the direct comparisons of this study and Yamada and Sato (2013), as Yamada and Sato (2013) imposed reference persons in the income comparisons to be solely the average in Japanese society, rather than allowing respondents to choose the comparison benchmark.

[20]The estimate for the UK dummy after excluding respondents who chose family members as their comparison benchmark is almost identical (-0.185 with $p < 0.01$).

about inequality as much as Europeans do. They suggest the following explanation: compared with Europeans, Americans may perceive their country as a more mobile society, where people have more opportunities to move up the income ladder, and so they care less about current inequality level. Likewise, the Americans may care less about the others' income if they perceive that social mobility is high in the US. However, we cannot isolate the role of economic and institutional environments (including social mobility) from cultural determinants across two countries (Luttmer and Singhal 2011; Alesina and Fuchs-Schuendeln 2007) from our data, so this argument is at best speculative.

Also these results may serve as supportive evidence of a well-known endogenous growth model by Cole et al. (1992) which attributes differences in the economic growth rates between the US and the UK, two countries with very similar economic fundamentals, to differences in social preferences. It is beyond the scope of this study to understand the underlying causes of differences in income-comparison attitudes across two counties, but our study is the first evidence of such US–UK differences in income-comparison attitudes as measured in discrete-choice approaches.

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 presenting the main results, we verify that randomization was correctly performed. One common check is comparing predetermined individual characteristics across the groups to ensure 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.

The results show that overall both the US and UK samples are balanced. Out of 32 variables examined in Table 4, only five variables showed a statistically significant difference across the groups ($p < 0.10$).[21] 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, but our estimates are barely affected by adding these controls. These results further confirm that randomization was successful.

[21]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.

4.2 Estimation

Since the information treatments are randomized, our econometric model is very straightforward. We simply estimate the 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 predetermined 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 as measured after information treatment.[22] In some specifications, we also add individual controls presented in Table 4 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_2 with associated standard errors to examine whether the effects of the two treatments are different.

4.3 Baseline 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 columns show estimates with individual controls included to increase efficiency.[23] The estimates barely change in both specifications, suggesting that randomization works well. The 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

[22] 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).

[23] Specifically, we added age (in years), log income (in USD/GBP), and dummies for gender, ethnicity/race, marital status, immigrant status, education, employment (unemployed, full-time worker, and chief earner), and parental status.

negative and marginally statistically significant ($p < 0.10$), suggesting that Treatment 2 makes the UK respondents more jealous. The coefficient of -0.012 on the Treatment 2 dummy translates into the worsening of income-comparison attitudes by one-tenth of the mean of the control group (-0.127).[24] The last two rows in column 4 show that the difference between the two treatments is only marginally statistically significant. As shown in Section 4.5, however, the results mask a heterogeneous effect among UK respondents, and a fraction of UK respondents are indeed significantly influenced by Treatment 2.

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 adds 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 4.5, we discuss in more detail why we think 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 respondents do not appear to care about others' income at all and thus the information treatment has no impact, while the UK respondents are much more malleable. These results are consistent with those of Kuziemko et al. (2015), which demonstrate that most policy preferences among MTurk respondents in the US are relatively fixed.

4.4 Robustness of UK results

In this subsection we verify the robustness of our results on the UK respondents. The effect of Treatment 2 is robust to a number of sample selections. Table 6 summarizes the results. To ease comparisons, 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.

First, responses from mindless respondents 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.

[24] The reweighted estimates are slightly larger in magnitude. The estimate on the 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.

Next, we limit the sample to those who correctly answered the verification questions. While the correctness in the verification question is weakly 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% of the UK sample) and to those who answer correctly on the first verification question (79.2% 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.

4.5 Potential mechanism

Here we aim at understanding a potential mechanism that could explain why Treatment 2 makes UK respondents more jealous, rather than less, when making income comparisons. Of course, other explanations may account for this 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 respondents weighted the first sentence more than the second, and as a result they become more jealous after Treatment 2. Consequently, our information treatment might have heightened the salience of income comparison and ended up directing the focus toward others' income levels. If this is indeed the case, we are likely to find larger effects of information treatment among those respondents who did not care much about income of others before the information treatment.[25]

We divide the UK sample by the baseline income-comparison attitudes γ_0 measured before the information treatment to shed light on the mechanism behind why the information treatment guided UK respondents in an undesirable direction.[26] 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% change

[25]We also examined heterogeneous effects among the UK respondents on other dimensions such as gender, educational attainments, and age. We did not observe systematic differences in the effects of treatment along these dimensions. We also investigated differential effects by choice of comparison benchmark, but obtained no meaningful differences, partly due to the smaller sample size once we stratified the sample by comparison benchmark (all results available upon request).

[26]We replicate the same exercise for the US respondents. We do not find any effects even after dividing the US sample by γ_0.

in income-comparison attitudes from the control mean. Our information treatment thus ends up making those who initially pay less attention to others' income levels become more comparison-conscious. 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.

In sum, our findings suggest that income-comparison attitudes are relatively stable, but there is some scope for governmental intervention to affect such attitudes in certain population groups (e.g., people who are initially less jealous).

5 Supplemental analyses

We established our main results on the baseline differences of income-comparison attitudes between the US and the UK in Section 3, and the differential responses to information intervention in Section 4. Here, we further present supplemental results considering different types of behavioral models.

As Duesenberry (1949) documented, comparison attitudes can differ depending on whether the person's relative position is above or below a reference point. Fehr and Schmidt (1999) consider such effects in a game-theoretic situation, constructing the egalitarian model. The utility function in this case should read

$$U = y - A\max\{\bar{y} - y, 0\} - B\max\{y - \bar{y}, 0\},$$

where $B < A$ and $0 < B < 1$. The parameter A stands for envy, while B is for the emotion of guilt. While our experimental setting is not designed to estimate parameters of such behavioral models, our setting indeed allows researchers to do so.[27]

Before applying this behavior model to our data, some remarks on the following supplemental analyses are in order. First, information shown to the respondents in our discrete-choice questions do not directly match with Fehr and Schmidt (1999), because the respondents need to calculate differences between y and $\bar{y}$. Second, we lose orthogonality in variations of choice questions since orthogonality is maintained based on y and $\bar{y}$ rather than $y - \bar{y}$.

[27] Another candidate for a utility function, not examined here, is the Stone–Geary type, where $U^{SG} = U^{SG}(y - S, \bar{y} - S)$. The idea is that the marginal rate of substitution (MRS) between own income levels and income of others can differ depending on the subsistence level S. It is not, however, technically possible to estimate differential MRSs depending on subsistence levels via the mixed logit model. This is because in the mixed logit model when we subtract some value (S) from the attributes, the coefficients obtained from the random utility model become the same as when original figures of attributes are used for the estimation. Hence, while it is intriguing to consider subsistence levels along with income comparisons, we need a different experimental setting, such as one extended to allow for different levels of S, to estimate the utility function.

Finally and most importantly, under the egalitarian model, we cannot directly compare A and B derived from mixed logit models across different sample pools (e.g., US vs. UK). This is a crucial flaw for our main purpose of cross-country analyses. Technically, the estimates from the mixed logit models are divided by *scale parameters*, which are different between the US and the UK. Since these scale parameters are unknown to researchers, we cannot compare the effects of jealousy and the sense of guilty between the countries. In contrast, we can obtain *true* estimates of γ in our behavioral model, because scale parameters in the numerator and denominator cancel when obtaining γ (see, for example, Train (2009) and footnote 17 of Yamada and Sato (2013) for more details on this point).

Nonetheless, we estimated the egalitarian model for each country. In both the US and the UK, we find that $A > 0$ and $B > 0$, as the theory presumes. However, we find that $A < B$, which contradicts the prediction of the egalitarian model (results available upon request).

Secondly, asymmetrical effects of income comparisons by income positions could apply not only to the baseline estimation as above, but also to the responses to information treatment. Given the significant responses of UK respondents to the information treatment, we examine whether the effects among the UK respondents differ by perception of relative position of their own income in the society. In our survey, before discrete-choice questions, we ask respondents to provide the level of own income (y_{survey}), as well as the level of expected income of the comparison benchmark ($\bar{y}_{survey}$). Note that these variables are for the real world, not for the hypothetical situations in the experiment.

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 ($y_{survey} = \bar{y}_{survey}$). Table 8 shows the results for reapplication of Equation (2) to each subgroup. We should view these results with considerable caution, because while y_{survey} and $\bar{y}_{survey}$ are collected before information treatment, both are reported by the respondents themselves, and thus are potentially endogenous to our outcomes.[28]

Interestingly, columns 1, 3, and 4 in Table 8 indicate that our results are driven mainly by those who stated that their own income levels are greater than those of their comparison benchmarks (column 4). Importantly, we control for own income in all the specifications. This result suggests that while their perceived higher income position initially makes them relatively indifferent to others' income, information intervention directs them to be more conscious of others' income. In

[28]Instead of using $\bar{y}_{survey}$ as reported by respondents themselves as reference income, another candidate for the reference income can be the average income of individuals in same regions, which some other studies use (e.g., Luttmer 2005). However, Table 2 shows that only 1.2% of respondents in the UK choose neighbor as the reference group in our study. Of course, this fraction may be a lower bound since the reference groups in other categories such as close friends could also be neighbors.

fact, as shown in third-to-last row in Table 8, the control mean of γ_0 among those who think they are in the higher-income position (-0.111) is smaller (i.e., less jealous) than both those who think they are in the lower-income position (-0.124) and those who stated the same income as others (-0.137).[29]

We further investigate the suggestion by Ferrer-i Carbonell (2005) regarding the importance of not only the relative position of one's income to others, but also the distance between own income and that of others. Specifically, for each subgroup, we estimate the following specification, which is a variant of the main specification (2):

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

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

Interestingly, column 5 shows that Treatment 2 makes those who are initially in the superior income position ($y_{survey} > \bar{y}_{survey}$) more comparison-conscious as the distance between own income and that of others becomes greater (i.e., running away from the Jones). Column 2 shows that such a pattern is not observed among those whose stated income is lower than that of the comparison benchmark.

6 Conclusion

Economists have long been concerned that negative attitudes toward relative income may reduce social welfare. This paper investigates whether we can alter such income-comparison attitudes by providing simple but intuitive information pointing out the potential harm of being comparison-conscious. To this end, we conduct an original randomized online survey experiment in the US and the UK. First, we find that UK respondents compare their incomes with others' more intensely than US respondents do. Second, we find that providing information on income comparisons makes a certain group of UK respondents even more jealous, but we do not find a similar effect on US respondents or other UK respondents. Our findings suggest that income-comparison attitudes are relatively stable, but also that there is some scope for governmental intervention to affect such attitudes among at least some population segments. However, our respondents are guided in a direction of decreased welfare, implying that changing these attitudes in a desired direction by a

[29]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).

simple message may be difficult.

There are some limitations to our study. First, we cannot examine the lasting effect of our information treatment due to data limitations. It is possible that such information treatment can alter income-comparison attitudes in the short term only. A few papers have tested the duration of effects from informational survey experiments, but the evidence is mixed. For example, Druckman and Nelson (2003) find that their results disappear within ten days after the information treatment, while Kuziemko et al. (2015) find that their effects last one month. Second, although our study indicates the possibility that government can probably affect income-comparison attitudes, we cannot show *how* to direct income-comparison attitudes in a desired way. Providing more direct and stronger information—stressing that making income comparisons is, indeed, problematic—could be more effective than our information treatment. These questions leave room for future research.

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.

Alesina, A. and N. Fuchs-Schuendeln, "Good bye Lenin (or not?): The Effect of Communism on People's Preferences," *American Economic Review* 97 (September 2007), 1507–1528.

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.

Aronsson, T. and O. Johansson-Stenman, "When the Joneses' consumption hurts: Optimal public good provision and nonlinear income taxation," *Journal of Public Economics* 92 (June 2008), 986–997.

———, "Positional Concerns In An Olg Model: Optimal Labor And Capital Income Taxation," *International Economic Review* 51 (November 2010), 1071–1095.

———, "Genuine Saving and Conspicuous Consumption," Working Papers in Economics 605, University of Gothenburg, Department of Economics, November 2014a.

———, "State-variable public goods and social comparisons," *Journal of Environmental Economics and Management* 68 (2014b), 390–410.

Aronsson, T. and K.-G. Lofgren, "Welfare equivalent NNP and habit formation," *Economics Letters* 98 (January 2008), 84–88.

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.

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 N. Szembrot, "Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference," *American Economic Review* 104 (September 2014), 2698–2735.

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.

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. 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.

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.

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.

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.

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

Gali, J., "Keeping Up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices," *Journal of Money, Credit and Banking* 26 (February 1994), 1–8.

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.

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.

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* 105 (2015), 1478–1508.

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.

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

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.

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.

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

Mujcic, R. and P. Frijters, "Economic choices and status: measuring preferences for income rank," *Oxford Economic Papers* 65 (January 2013), 47–73.

———, "Conspicuous consumption, conspicuous health, and optimal taxation," *Journal of Economic Behavior & Organization* 111 (2015), 59–70.

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.

Smets, F. and R. Wouters, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association* 1 (September 2003), 1123–1175.

———, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review* 97 (June 2007), 586–606.

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," *Brookings Papers on Economic Activity* 39 (Spring 2008), 1–102.

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).

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.

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.

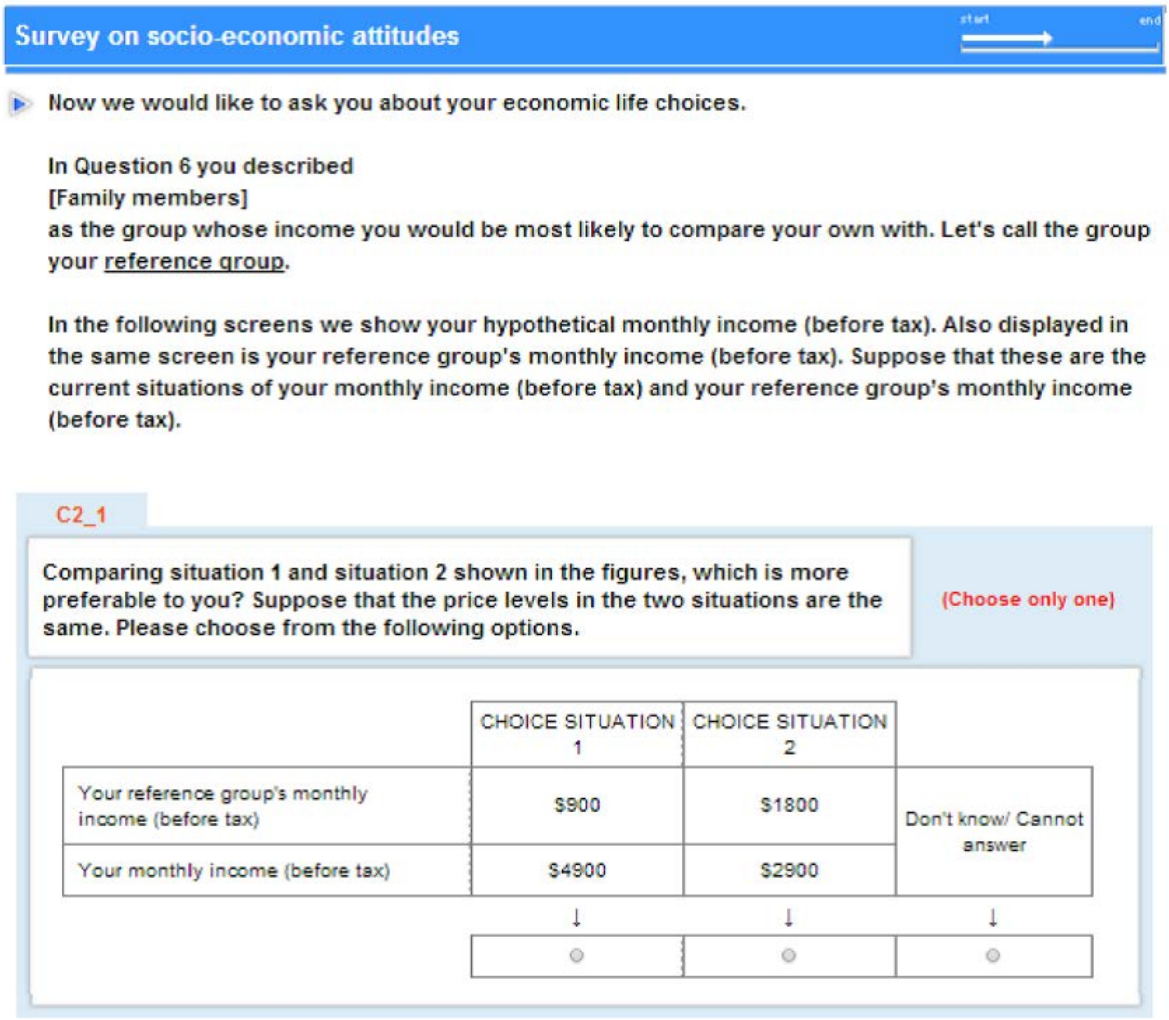
Source: World Database of Happiness and Penn World Tables.

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.

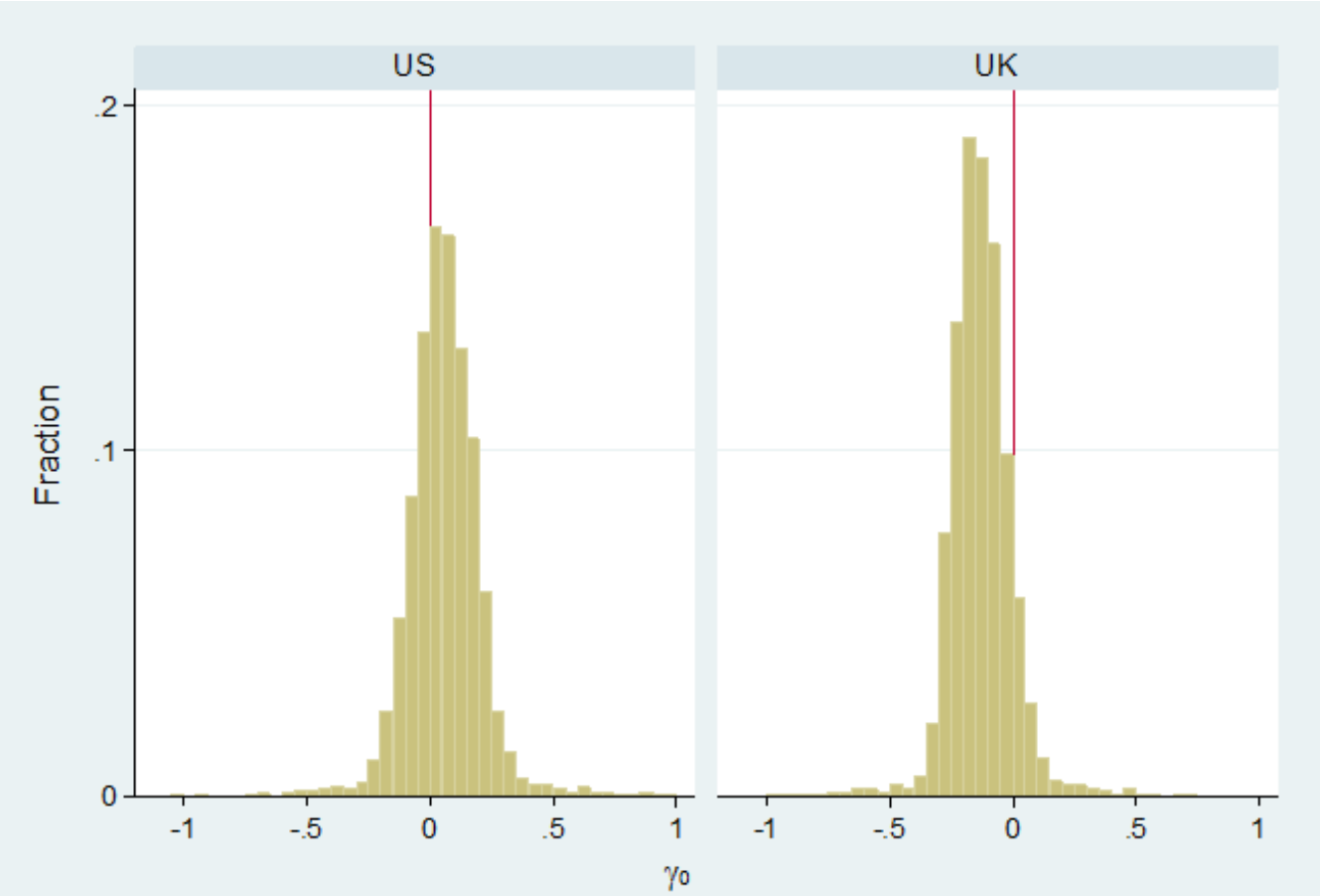
Note: For the UK case, the US figures in USD are changed to the UK figures in GBP, and all the terms "United States" in the sentences are changed to "United Kingdom".

Figure 2: Snapshot of a discrete-choice question (the US case)



Note: For the UK, the figures are in GBP and come from actual income distribution in the UK. On top of this question, the comparison benchmark (i.e., who to compare with) chosen by each respondent is shown with the aim of making the comparison more realistic. In this example, "family members" is chosen as a comparison benchmark group.

Figure 3: Distribution of γ_0 (baseline)



Note: The bin size is 0.05 in both graphs. The vertical line corresponds to a value of zero. The total sample size is 4,587 for the US and 4,500 for the 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 −0.129, 0.126, −0.974, and 0.737, respectively. See Section 2.3 and Appendix A2 for a description of how we 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 with whom each respondent 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/GBP)	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/GBP), 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***	-0.186***
	(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	-0.009
		(0.006)	(0.006)
Hispanic		-0.008	-0.009
		(0.007)	(0.007)
Asian		-0.023***	-0.023***
		(0.007)	(0.007)
Other race		0.015	0.016*
		(0.009)	(0.009)
Married		0.002	0.002
		(0.003)	(0.003)
Immigrants		0.006	0.006
		(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*	-0.006*
		(0.003)	(0.003)
Log income (USD/GBP)		-0.000	-0.000
		(0.002)	(0.002)
Constant	0.056***	0.064***	0.033
	(0.002)	(0.019)	(0.022)
Control for comparison benchmark	No	No	Yes
R-square	0.316	0.317	0.320
Observations	9,087	9,087	9,087

Note: All explanatory variables, except for age (in years) and log income (in USD/GBP), 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. 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/GBP)	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: Baseline 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)

Note: 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 age (in years), log income (USD/GBP), and dummies for gender, ethnicity/race, marital status, immigrant status, education, employment (unemployed, full-time worker, and chief earner), and parental status. 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)

Note: 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. The last two rows report the difference between the estimates of Treatment 1 and Treatment 2, 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% of the sample) and those who correctly answered the first verification question (79.2% 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	By 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 a list of controls. Control mean is the mean value of γ for the control group. Note that the lower the baseline γ_0, the more jealous the respondent initially is. 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)
Treatment1	0.004	-0.008	-0.004	-0.022	-0.031
	(0.009)	(0.013)	(0.011)	(0.021)	(0.028)
Treatment2	-0.009	-0.010	-0.008	-0.037*	0.014
	(0.009)	(0.013)	(0.011)	(0.021)	(0.028)
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)
Control mean	-0.124	-0.124	-0.137	-0.111	-0.111
R-squared	0.017	0.018	0.010	0.038	0.054
N	2,417	2,405	1,598	485	485

Note: y_{survey} is own income, and $\bar{y}_{survey}$ is the respondents' 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 $***p < 0.01$.

Appendicies (not for publication)

A1 Flow chart of questionnaire

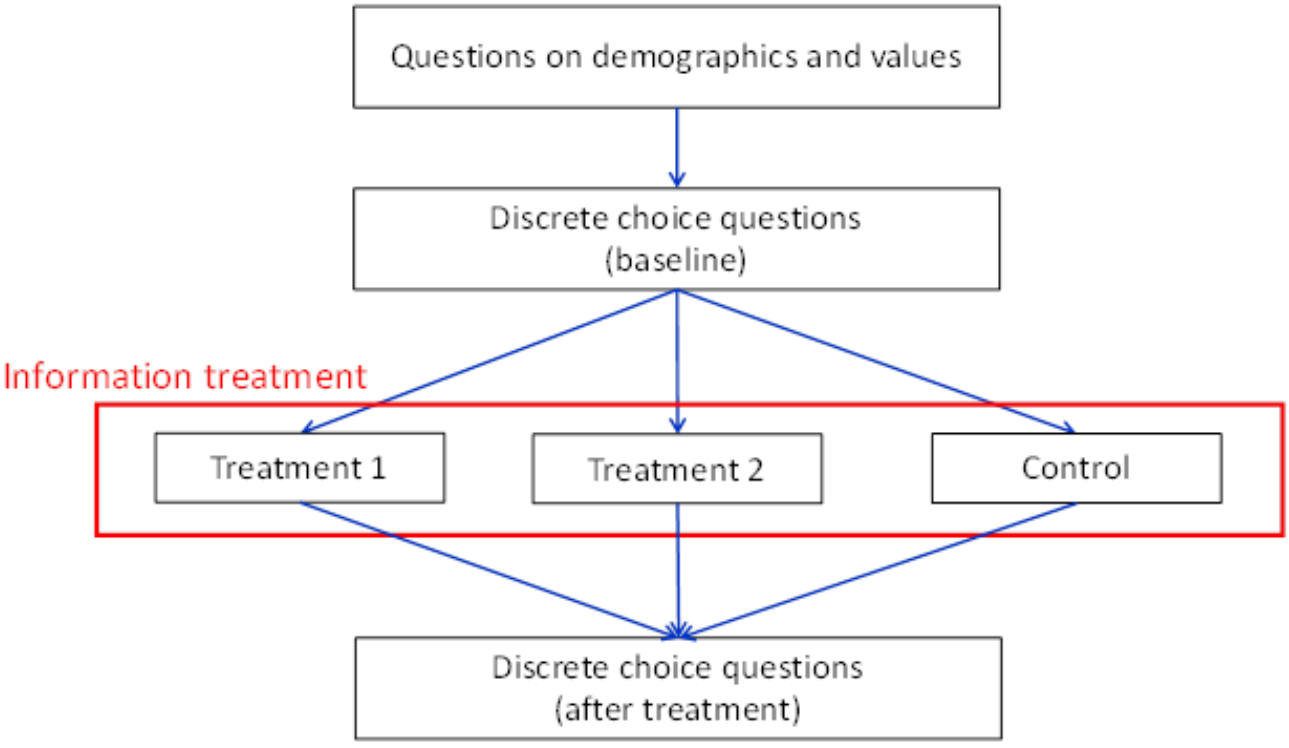
The structure and the flow of the experiment are as follows:

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

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

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

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



A2 Technical Appendix

A2.1 Construction of hypothetical discrete-choice questions

First, we explain how we construct our discrete-choice questions. Five attribute levels are set considering monthly income distributions in the US and the UK. 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 conduct *orthogonal planning* in choosing the alternatives to be used in choice questions. This method effectively pairs multi-dimensional 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 regressions of the random utility model (explained later), because the independent variables in the regressions become orthogonal. We use SPSS Conjoint (ver. 20.0) for orthogonal design of alternatives in this study.

Next, we construct *choice sets* consisting of pairs of alternatives and the no-choice option. According to the requirements of orthogonal design, we generate two 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 alternatives, one taken from one alternative vector and the other from the other alternative vector. The pairing strategy is at the discretion of the researchers, but all variations must be exploited and same alternatives cannot be used twice. Because 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 (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 do not exclude the possibility of altruistic preference. Given these constraints, our best strategy for pairing alternatives is as follows.

Suppose the scenario $Z = (x, y)$, where x denotes the level of one's own income and y is others' income. Then, qualitatively, candidates for 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 $x+$

[1] As such 25 alternatives out of 25 potential variations had to be used to meet the requirement of orthogonal design.

means some value greater than x and $x-$ means some value smaller than x. Since we do not exclude the possibility of altruism, there are no *a priori* dominant choices for Z from these eight alternatives. We then make pairs such that these eight situations appear as evenly as possible. Using the procedures discussed here, we are able to efficiently obtain parameter estimates. Appendix Table A1 shows the set of questions we use in the survey. Each respondent answers 12 randomly assigned questions (6 questions before information treatment and 6 after information treatment) out of the 25 possible questions.

A2.2 Random utility model and empirical method

Here, we introduce the econometric foundation of how respondents' 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 respondents 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 choose 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 respondents who answer $T(\geq 1)$ repeated questions. The utility of respondent n when choosing 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, 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 the respondent 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 respondent n choosing situation i from the set of situations S_t (the choice set) in question $t \in T$ can be written as

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

McFadden (1974) shows 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 respondent 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 respondents 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]

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

In our study, we consider the case where $K = 2$ and 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 respondents value attribute y positively. With a linear utility function of the random utility model, experimental respondents derive utility according to

$$V = \beta_1 y + \beta_2 \bar{y}, \tag{1}$$

where the ratio of marginal utility from $\bar{y}$ to that from y, β_2/β_1, captures the trade-offs between own and reference group income.

Using maximum-likelihood estimation, we obtain point estimates for β_1 and the distribution of β_2.[4] The individual coefficient of $\beta_{2,i}$ can be obtained by applying the inverse Bayesian formula to the estimated distribution of the coefficient and the choice pattern of the individual respondent (Train 2009). Finally, we obtain individual trade-off between own and reference group income by dividing the individual coefficient of $\beta_{2,i}$ by the average effect of β_1.

Now we relate the ratio of β_2/β_1 to macroeconomics theory. Following Abel (1990), Gali (1994),

[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.

and Dupor and Liu (2003), consider the constant relative risk-aversion-type utility function as

$$V^m = \frac{(y\bar{y}^\gamma)^{1-\rho}}{(1-\rho)}, \tag{2}$$

where $\rho > 0$ and the parameter γ regulates the intensity and sign of income-comparisons. 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. With this utility function, the ratio of marginal utility from the reference income and that from own income becomes

$$\frac{\partial V^m/\partial \bar{y}}{\partial V^m/\partial y} = \gamma \frac{y}{\bar{y}}. \tag{3}$$

Hence, the ratio of marginal utility, β_2/β_1, as obtained in the random utility framework corresponds to a parameter γ of the constant relative risk-aversion-type utility function in the economy of symmetric equilibrium where $y = \bar{y}$.

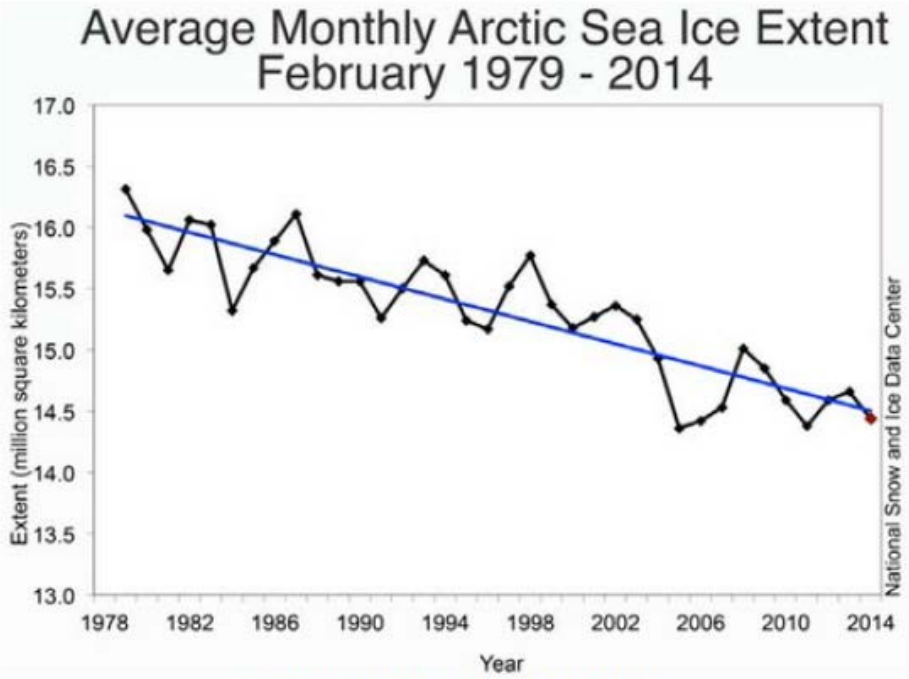
Figure A1: Information treatment for control group (US case)

Survey on socio-economic attitudes

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

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.

Table A1: Choice sets for 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 respondent answers 12 randomly assigned questions (6 questions before information treatment and 6 after information treatment) out of the 25 possible questions.

References

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

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

Gali, J., "Keeping Up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices," *Journal of Money, Credit and Banking* 26 (February 1994), 1–8.

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.