

IZA DP No. 9189

When Experienced and Decision Utility Concur: The Case of Income Comparisons

Andrew E. Clark
Claudia Senik
Katsunori Yamada

July 2015

When Experienced and Decision Utility Concur: The Case of Income Comparisons

Andrew E. Clark

*Paris School of Economics
and IZA*

Claudia Senik

*Paris School of Economics
and IZA*

Katsunori Yamada

Kindai University

Discussion Paper No. 9189
July 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

When Experienced and Decision Utility Concur: The Case of Income Comparisons^{*}

While there is now something of a consensus in the literature on the economics of happiness that income comparisons to others help determine subjective wellbeing, debate continues over the relative importance of own and reference-group income, in particular in research on the Easterlin paradox. The variety of results in this domain have produced some scepticism regarding happiness analysis, and in particular with respect to the measurement of reference-group income. We here use data from an original Internet survey in Japan to compare the results from happiness regressions to those from hypothetical-choice experiments. The trade-off between own and others' income (showing the importance of absolute and relative income) is similar in these two sets of results. This kind of validation of experienced utility via direct comparison with decision utility remains rare in this literature.

JEL Classification: D31, D63, I3, J31

Keywords: satisfaction, income comparisons, reference-group income,
discrete-choice experiments

Corresponding author:

Andrew E. Clark
Paris School of Economics
48 Boulevard Jourdan
75014 Paris
France
E-mail: Andrew.Clark@ens.fr

^{*} We would like to thank the Associate Editor and the three anonymous referees for their constructive and detailed advice, which has substantially improved the paper. We also would like to thank Hitoshi Shigeoka for his comments. Financial support from CEPREMAP, the Global COE program of Osaka University, and the Grants-in-Aid for Scientific Research of the Japan Society for Promotion of Science (24683006) is greatly appreciated.

1 Introduction

Work on subjective well-being and income comparisons in economics has increased exponentially since the first labour-market based contributions of Cappelli and Sherer (1988) and Clark and Oswald (1996). While the broad idea that relative standing matters is now arguably accepted by many, it is still pretty much open season on the key questions of “How much”. Using macro-level data, where the assumption is that GDP per capita is a good proxy for reference income, Easterlin (1995, 2001) suggests that own and reference-group income are equally important, while Stevenson and Wolfers (2008) conclude that own income predominates.

The discussion of the relative importance of own and reference-group income will likely continue to rage in subjective well-being research (see Clark et al. 2008). However, a killer blow may be dealt if it can be shown that many of the approaches used to define comparison income in happiness regressions are artificial, and hence unreliable.¹ One alternative to researcher-defined reference-group income is to exploit instead respondent-supplied measures of reference-group income: Mayraz et al. (2009), de la Garza et al. (2010), and Goerke and Pannenberg (2013) all use such subjective information on cardinal reference-group income and conclude that income comparisons do indeed affect reported satisfaction.² Another is to conclude that subjective well-being regressions in general are unreliable, and instead estimate preferences from observed choice behaviour.

We here contribute to this strand of research using data from a Japanese survey that includes information on both subjective well-being and self-reported reference income. It is common in non-experimental work to calculate “others’ income” as some conditional or unconditional cell-mean, with the cells being defined by neighborhood, workplace or demographic type. We here do so as well and show that two such cell-mean measures (one from within the dataset, the other matched in from external data) fit the well-being data roughly just as well as does self-reported reference income. Simple questions on others’ earnings then do seem to capture viable information on the income in the actual reference group to which our respondents compare.

One novel feature of the survey that we use is that it also includes discrete-choice experiments regarding income comparisons. We will show here that the estimated own income and reference-group income trade-off in happiness regressions is close to that resulting

¹See Manski (1993) and Sloane and Williams (2000).

²Knight et al. (2009) and Senik (2009) also use respondent-supplied reference-group incomes, but on ordinal scales, making the coefficients less straightforward to interpret.

from the choice experiments in the same dataset. These discrete-choice experiments were the subject of Yamada and Sato (2013). This validation of happiness regressions via an alternate estimation strategy is still rare in the literature, although there have been a small number of recent salient contributions.

One notable piece of work in this extent is Oswald and Wu (2010), who look at the life satisfaction scores (on a 1-4 scale) in the US Behavioral Risk Factor Surveillance System data from 2005-2008. They run satisfaction regressions on individual demographics and 49 State dummies to provide a State-by-State picture of well-being. Their objective measure of well-being is a weighted sum for each State of variables such as temperature, wind speed, sunshine, National Parks, hazardous waste sites, commuting time and violent crime (with the weights in the sum coming from the estimated coefficients in regional wage and house price equations). The objective ranking of US States obtained this way is shown to be strongly correlated with that from the life satisfaction analysis.

Some recent work by Daniel Benjamin and co-authors has compared decisions and subjective well-being scores (as we do here) to see whether individuals make decisions based only on happiness, or rather on happiness and something else as well. Benjamin et al. (2012) consider a sequence of hypothetical pairwise-choice scenarios. In the example they highlight (page 2087), individuals are asked to decide between two new jobs. The jobs are identical except with respect to work hours and pay, as follows.

- Option 1: A job paying \$80,000 per year. The hours for this job are reasonable, and you would be able to get about 7.5 hours of sleep on the average work night.
- Option 2: A job paying \$140,000 per year. However, this job requires you to go to work at unusual hours, and you would only be able to sleep around 6 hours on the average work night.

Individuals are first asked between these two options, taking all things together, which do you think would give you a happier life as a whole? They were then asked if you were limited to these two options, which do you think you would choose? One of the authors' main findings is that some individuals do not choose the option in the second question that they said would make them happier in the first question (see their Table 2), although only around ten per cent do not do so. A follow-up paper (Benjamin et al. 2014) takes something of the same tack, but this time looks at actual (and not hypothetical) choices. The actual choice here is a major one: the choice of residency submitted by medical students to the National Resident Matching Program. The students were asked

to score their top four choices in terms of well-being (both during residency itself, and for the rest of their lives). They also scored other aspects of the residency (prestige, stress, career value, etc.). The argument developed by Benjamin et al. (2014) is that if choice reflects well-being only, then the trade-offs between the aspects of the residency (stress vs. prestige, say) should be the same in choice equations and subjective well-being equations. The key result here is that, although the subjective well-being scores are correlated with the ranking of actual choices (see their Figure 2), the marginal rates of substitution between aspects in actual choice are different from those in subjective well-being regressions. The authors conclude that we should then be wary of using the results from the empirical analysis of happiness or life satisfaction questions to inform us about the trade-offs that occur in actual choices.

Perez-Truglia (2015) also sounds a warning note. He considers the concordance of SWB data and decisions by estimating consumption functions from SWB data and then looking at the difference between predicted consumption and actual consumption. He finds that, while life satisfaction predicts real consumption well, measures of economic satisfaction (which is what we have here) did not yield accurate consumption predictions. Somewhat along the same lines, Ferrer-i Carbonell et al. (2011) compare the estimates on job characteristics in choice equations using vignettes to those on the same characteristics in determining the respondent's own job satisfaction. They find statistically significant differences between the way in which individuals evaluate their own job and the hypothetical jobs.

Last, Akay et al. (2015) compare the income-leisure trade-off in labour-supply models estimated on BHPS data to the income-leisure trade-off in life satisfaction equations. Heterogeneity is introduced in labour supply via a preference shifter in leisure which varies with a number of demographic variables. The income-leisure trade-off is found to be similar in size in both approaches.

The small existing literature has then come to mixed conclusions regarding the determinants of choice and subjective well-being. We here contribute to this work in the specific context of income comparisons (asking how individuals trade off absolute and relative income), which are one of the most frequently-cited explanations of the Easterlin Paradox. We will find that the trade-off between own and others' income in hypothetical choices is similar to that in satisfaction regressions.

The remainder of the paper is organized as follows. Section 2 describes our Japanese data. Section 3 then presents the results from happiness regressions and compares the

income-comparison results to those in Yamada and Sato (2013). Last, Section 4 concludes.

2 Data description

2.1 Survey strategy and data construction

Our data comes from an original Internet-based survey run by a Japanese consumer monitoring company, NIKKEI Research Inc., under the direction of one of the current authors. To ensure the reliability of the research data, individuals who register with NIKKEI Research Inc. for consumer monitoring are subject to monthly screenings so that their individual information remains up to date and double registrations are avoided. The incentives for respondent participation in this particular project were provided by cash vouchers.³ The invitation email stated that the incentive would be paid on a lottery basis: 800 winners among those who completed the survey would receive 500 JPY per person. Subjects were ensured anonymity. Those who wished to participate were instructed to follow the link in the email which directed them to our stand-alone survey website.

The aim was to have a minimum of 10,000 survey respondents. Considering the response rates in seven similar academic choice-experiment surveys run by the same company in 2008 and 2009, invitation emails for this survey were sent to 60,482 subjects (out of over 160,000 registered in the NIKKEI Database).⁴ The subjects, aged between 20 and 65, were chosen via stratified random sampling, so that the sample age profile mirrored Japanese census statistics. We could not match by education, as subjects do not have to declare this information when they register. The email specified that the survey was being carried out for research purposes and followed the disclosure requirements for research involving human beings provided with incentives as set forth by the University of Osaka ethics committee. Those who wished to participate were instructed to follow the link in the email which directed them to our stand-alone survey website. The survey was open for one week (February 18th-25th 2010), and 14,370 subjects completed the survey.

To construct the analysis sample, we dropped a certain number of observations. The

³We used cash vouchers rather than point vouchers that can be redeemed against specific goods. The latter is considered to lead to potential bias, as particular respondent types (those interested in the good for which the points are destined) are more likely to respond to the survey. If these characteristics are not orthogonal to the right-hand side variables under consideration (here, comparison effects), the estimated coefficients will be biased. Our use of cash incentives avoids this problem.

⁴Broadband Internet access is pervasive in Japan.

average time taken to complete the survey, including the hypothetical discrete-choice questions on preferences over combinations of own and others' income that we describe below, was just over 9 minutes, with a median value of just over 6 minutes. We drop observations with no time records (90) or time taken to complete the survey of over 60 minutes (106); we also exclude subjects with a completion time of under four minutes (968), following feedback from a within-company pilot test supplied by NIKKEI. In the context of weeding out spurious responses, we also pay attention to the answer patterns supplied to the hypothetical-choice questions. One easy way for subjects to finish the survey quickly is to provide the same answers to all of these choice questions, which are arguably the most difficult part of the survey. We therefore drop individuals who gave the same number for all five of the consecutive questions in the choice tasks (2,218).⁵ Observations were also dropped if they were either missing information for some of the variables used in the empirical analysis below (219), or contained inconsistencies, such as retirement before the age of 55 (1). Finally, we excluded observations from respondents who report their personal annual pre-tax income in 2009 to be over 12 million JPY (565).⁶ This cut-off point, 12 million JPY, is more than three standard deviations over average personal pre-tax annual income. After data cleaning our final sample consists of 10,203 respondents, and is the same sample as analyzed in Yamada and Sato (2013).⁷

The survey includes standard demographic information. Annual pre-tax personal income in 2009 is recorded via a list of 11 categories, where category 1 refers to an annual figure of below 2 million Yen and category 11 to an annual income of over 50 million Yen. Income is calculated as the mid-point of each of the nine intermediate categories, with *ad hoc* values of 1.5 and 55 million Yen being assigned to the two extreme categories. Sample descriptive statistics as compared to national statistics appear in Table 1. There is evident over-sampling of the more-educated. As Yamada and Sato (2013) show that the income-comparison results are robust to adjustments for education over-sampling, we mainly analyze the data without adjusting the sampling weights. Average own pretax annual income for men in the analysis sample, as described above, is 4.9 million Yen (around 54 000 US Dollars) in Table 1; the figure for women is notably lower at 2.7 million Yen. All of the income variables are entered in log form in the regression equations.

More importantly for the purposes of the current paper, the survey also includes infor-

⁵This is known as the “long string” indicator of careless responding (Meade and Craig 2012)

⁶At the February 21st 2010 exchange rate, 12 million Japanese Yen is around 132,000 US Dollars.

⁷It is worth noting, however, that all of our main results listed below are qualitatively similar in the overall and analysis samples.

mation on both subjective well-being and subjective reference-group income. In addition, survey respondents participate in choice experiments regarding income comparisons. Income satisfaction comes from the question “*How satisfied are you with the amount of your current income?*”, with responses on a one-to-five scale (where 1 corresponds to “Extremely dissatisfied” and 5 to “Extremely satisfied”). Income satisfaction in our sample is fairly low: 21% of respondents report the lowest level of 1, and 41% a score of 2. Only 7% of respondents report a score of 4 or 5. These low satisfaction scores correspond to existing work underlining lower satisfaction figures given by Japanese respondents in general (Abdallah et al. 2008), and that the answers to income satisfaction questions are less positive than those to other kinds of satisfaction questions (see Appendix A in Clark 2001 for the domains of job satisfaction in the BHPS, for example). Income satisfaction will be our dependent variable in the regressions below, where we relate income satisfaction to both own and others’ income.⁸

2.2 Three types of comparison-income measures

For the income satisfaction regressions, we have available three different measures of the income to which individuals may compare their own income. One of these is a self-reported measure of the income that “people like you” earn; the other two are constructed as cell means, one from within the database and the other from an external data source.

The first comparison-income measure comes from the respondents themselves, who are asked: “*About how much do you suppose was the average personal income (before taxes) in 2009, for people of the same age, sex, and education as you?*” The answers here are on the same 11-point scale as used for own income above.⁹ The figures in Table 1 show that both men and women believe that they earn less than do comparable others. Men think that others with the same characteristics earn on average around 20% more than

⁸It is more common in the literature to consider general measures of subjective well-being, such as life satisfaction. The only subjective variable we have available in our survey refers to income satisfaction. It is well-known, however, that domain satisfaction variables (such as income) and overall life satisfaction evaluations are strongly correlated. In addition, there is existing work on income comparisons which has appealed to measures of financial satisfaction similar to that which we use here (see Clark et al. 2009a). Last, income satisfaction may be a particularly salient well-being measure when we are looking at the effect of income comparisons.

⁹Information on what individuals believe others earn is arguably important for their feelings of deprivation or satisfaction. The disadvantage is that this question captures the reference income of only one particular group defined in a *Leyden-group* fashion as a cell mean. This group may not be the most salient for the income comparisons that individuals carry out. We argue that this measure of self-reported reference income will still capture some part of the effect of income comparisons, as it is likely correlated with the real income to which the individual compares.

they do, while the comparable figure for women is a remarkable 64%.

The second measure is given by cell means from within the internet survey. To be compatible with the other measures used here, this is calculated using cells defined either by age, education, and gender, or age, education, gender and labour-force status (as described below). Average cell-mean internal income over all respondents equals average own income within the survey, by construction (see Table 1).

The last measure of reference-group income is external, corresponding to average income across the following labour-force statuses: (i) employed in the private sector, (ii) professional jobs, (iii) civil servants, (iv) temporary workers, (v) other public sector (but not civil servants), (vi) unemployed and receiving benefits, and (vii) self-employed. This income data comes from the Basic Survey on Wage Structure (BSWS) released by the Japanese Ministry of Health, Labour and Welfare (<http://www.mhlw.go.jp/english/database/dbl/>). For statuses (i) and (iii), we construct cell-average income by sex, age and education. For (ii), we use information on the income of doctors and chartered accountants: the cell averages here are by age and gender, but not education (which varies only little within this professional group). For group (v), we have no direct information. These positions are public sector but inferior to those in group (iii). We thus assume that the income of those in group (v) is proportional to that in group (iii). Income in temporary jobs is inferred from information on average days worked per month, working hours per day, and hourly wage by age and gender. For the unemployed, we have income information from unemployment insurance. We assume for simplicity that the unemployed have no other sources of income. Last, we use information on pre-tax income by age of the self-employed. This external measure of comparison income is then defined by age, education, gender and labour-force status cells.

2.3 The discrete-choice question on income comparisons

One of the most novel aspects of the survey data that we use here is that it includes the responses to three types of hypothetical-choice questions regarding income comparisons, as analyzed in Yamada and Sato (2013). We here replicate the result from one of these three questions, regarding comparisons to the country average, which we then use to compare the answers to the experienced utility (income satisfaction) and decision utility (hypothetical choice) analyses.

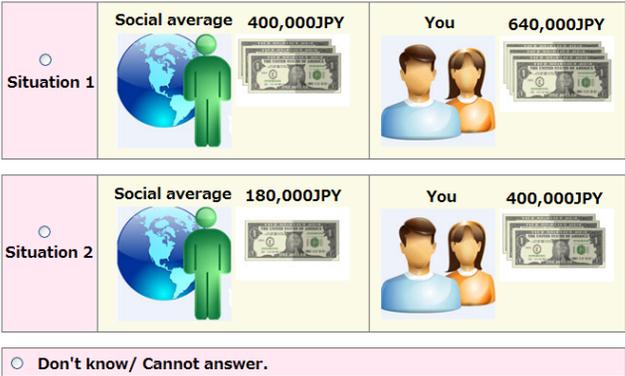
Before the subjects began the questionnaire section including the hypothetical choice questions regarding income comparisons, they were shown an instruction screen with the

following information:

The following figures show your hypothetical monthly income (before tax). The figures also display the average monthly income (before tax) in Japan overall. We would like you to imagine that these figures refer to your actual monthly income (before tax) and the average monthly income (before tax) in Japan.

In the screens that followed, respondents were asked hypothetical discrete-choice questions relating to a variety of figures for both own income and average income in Japan, as illustrated below.¹⁰ The number of banknotes in the corresponding images varies according to the particular monthly income figures shown.

Which of situations 1 and 2 shown in the figures below do you prefer? The price levels in the two scenarios are the same. Please choose from the following options.



In the figure, each situation is defined by two *attributes*, the individual’s own monthly pre-tax income and that of the reference group. As well as expressing a preference for one or the other situation, individuals could also indicate “Don’t know / Cannot answer.” In our regression analyses of hypothetical choice we do not use observations in which the Don’t know option was chosen. Taking into account the actual monthly income distribution in Japan, the attribute levels were chosen from the following list: 180,000 JPY, 240,000 JPY, 400,000 JPY, 640,000 JPY, and 900,000 JPY. The combination of the two attributes (own income and reference income) and five possible income levels for each attribute yield 25 different income situation scenarios. In the literature, these different scenarios are called *alternatives*. To construct the choice question, the two alternatives

¹⁰In the actual survey the information in the figures appeared in Japanese.

were paired (as shown in the figure above) using *orthogonal planning*.¹¹ Each respondent reported their preferences for five randomly-chosen alternatives out of the 25 total possible combinations in turn. Respondents were not allowed to go back and change their answers to any given question once they had made their choice.

2.4 The estimation of income-comparison effects from discrete-choice data

We here introduce the econometric description of how respondents' choice data can be used to estimate the importance of income comparisons in individual well-being. We analyze the decisions in hypothetical-choice experiments in a random utility model framework. It is here assumed that subjects choose one income situation over any other because they would obtain the highest utility in the first situation. When there are two situations (A and B, for example as in the current paper), and individuals choose A over B, then the choice data is recorded as 1 for A and 0 for B, along with the levels of the explanatory variables (here, the levels of own and reference income). The choice and the explanatory variables make up the observation in the regression analyses.

Assume that there are N subjects who answer $T (\geq 1)$ repeated-choice questions. The utility of respondent n who chooses situation i for question $t \in T$, U_{itn} , consists of the observable components from the explanatory variables, V_{itn} , and unobservable components, ϵ_{itn} , so that utility can be written as $U_{itn} = V_{itn} + \epsilon_{itn}$. Utility from the observable components is assumed to come from a linear combination of the explanatory variables, $V_{itn} = \sum_{k=1}^K \beta_k X_{ik}$, where $k = 1, \dots, K (K \geq 2)$ represents the type of explanatory variable ($K = 2$ here), X_k denotes the levels of the k th explanatory variables, and β_k the marginal utility of each variable. In the analysis below, the vector $\beta \equiv (\beta_1, \dots, \beta_K)$ that maximizes the log-likelihood of the observed choice patterns for each subject is the mixed-logit regression estimator. Following McFadden (1974), ϵ_{itn} is considered to be independent and identically distributed according to an extreme-value-type 1 (IIDEV1) distribution with variance σ^2 .

The logit choice-probability formula P_{itn} for subject n choosing situation i from the

¹¹See Louviere et al. (2000) for details of the construction of choice questions via the orthogonal planning method. See also Huber and Zwerina (1996) and Viscusi et al. (2008) for the design of choice questions that aims to balance the utility of each alternative.

set of situations S_t (the choice set) in question $t \in T$ can be written as

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

McFadden (1974) shows that $P_{itn} = \exp(\omega V_{itn}) / \sum_{j \in S} \exp(\omega V_{jtn})$, where $\omega = \pi / \sqrt{6}\sigma$ is a scale parameter.¹²

Finally, we define a dummy variable d_{itn} for subject n choosing 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 the observed repeated choices can be written as

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

The error term is assumed to be independently and identically distributed. If we also assume that the independence of irrelevant alternatives (IIA) holds, we obtain the conditional logit model; in non-IIA situations, we can appeal to the mixed logit model where we can obtain the distributions $f(\beta_i)$ of some parameters in β across subjects.

In the mixed logit model, $f(\beta)$ is specified as being normally distributed with a parameter set of θ , following Train (2009). The choice probability function P_{itn}^{ML} is here

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

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 function¹³

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

We now specify the shape of the utility function over own income, $X_1 = y$, and that of those in the reference group, $X_2 = \bar{y}$. We make the typical assumption that individuals value y positively. We consider the constant relative risk-aversion-type utility function as

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

¹²There are of course many other individual characteristics that affect the utility of each alternative. In the formula for P_{itn} above, these appear in both U_{itn} and U_{jtn} and thus cancel each other out. The regression here includes only the attributes of the choices as explanatory variables.

¹³See Section 6 of Train (2009) for details.

where $\rho > 0$. The parameter γ reflects the intensity and sign of income comparisons, and is the main focus of the current paper. When $\gamma < 0$, the individual is jealous of others, in that higher values of \bar{y} reduce the individual’s utility. On the contrary, if $\gamma > 0$, then the individual has altruistic preferences (or assigns a positive signal value to others’ income, as in Clark et al. (2009b)). Last, when $\gamma = 0$ others’ incomes do not affect individual utility. Taking logarithms of both sides of Equation 1, we have

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

Under the conditional logit model, we can use maximum-likelihood estimation to obtain the point estimates for $\ln y$ and $\ln \bar{y}$; with the mixed logit model we have a point estimate for $\ln y$ and an estimated distribution of the effects of $\ln \bar{y}$.¹⁴ In the former, the average effect of income comparisons γ is calculated by dividing the estimate of $(1 - \rho)\gamma$ by that of $(1 - \rho)$.¹⁵ Following mixed logit estimation, the individual value of γ_i can be obtained by applying the inverse Bayesian formula to the estimated coefficient distribution and the individual’s choice pattern (Train 2009).

Table 2 presents the results from the discrete-choice experiment for our study sample of 10,203 subjects. As each individual answered five repeated questions, the number of observations (48,172) is nearly five times the number of experimental subjects (10,203).

The coefficient on *Own income* in column (1) of Table 2, referring to the conditional logit model, shows that own income affects utility positively and significantly. The negative coefficient on *Reference income* reveals that our Japanese respondents are jealous of others’ income. The ratio of these two estimated coefficients yields the relative-utility parameter, γ . From column (1), we have $\gamma = -0.458$. Column (2) of table 2 reports the result of the mixed model estimation. $\gamma = -0.676$ is the ratio of the two estimated income coefficients: that’s math we say a few lines above. The last column of Table 2 refers to a robustness test which we shall describe below.

3 Satisfaction and Income Comparisons in Japan

We now estimate the effect of income comparison in experienced utility (i.e. satisfaction) data in the same vein as Eq. (2). When we estimated the extent of income comparisons

¹⁴See Revelt and Train (1998) for the details regarding this procedure.

¹⁵Note that we cannot estimate the true value of ρ in a random utility model, as the estimated coefficient on $\ln y_{ni}$ and the constant term are divided by the scale parameter ω . We do not know the value of this latter parameter. For more technical details, see footnote 17 in Yamada and Sato (2013).

in Section 2.4 above, no demographic variables were introduced in Table 2’s conditional logit model. To make sure that we compare identical specifications across decision and experienced utility, our main analyses using income satisfaction as the dependent variables have only own income and reference-group income as right-hand side variables. We will then carry out some robustness checks including standard demographic variables such as age and gender as additional controls.¹⁶

3.1 Results

The results of the income-satisfaction regressions appear in Table 3. These estimations are OLS, except for column (3) which is an ordered probit. In columns (1)–(4) reference-group income is the self-reported level given by respondents. Column (1) shows the results from a baseline regression of income satisfaction on own income and reference-group income, and column (2) those when we use post-stratification sampling weights.¹⁷ Column (3) is the ordered probit version of column (1). Column (4) adds gender, age, age-squared, education dummies (5), marital-status dummies (2), region dummies (8), and the individual values of γ_i obtained from the mixed-logit estimation of discrete choice data to the specification in column (1).¹⁸ We control for γ_i as a measure of individual fixed effects or persistent unobservable characteristics in our cross-section data. All income variables are entered in logarithms.

The estimates in column (1) reveal a positive relationship between income satisfaction and own income, but a negative relationship with respondent-supplied reference-group income. The central result here is that the estimated relative importance of own and reference-group income in this regression is strikingly similar to that found in the hypothetical-choice experiments using the same individuals in the same dataset described above in Table 2. The satisfaction ratio of the effect of reference-group income to own income (hereafter SR) is -0.528 in column (1) (halfway between those put forward Easterlin (1995) and Stevenson and Wolfers (2008)). The analogous hypothetical-choice figure for the same individuals, estimated as γ in Eq. (1) with the social-average reference group in Table 2, is -0.458. Experienced utility (satisfaction) and decision utility (choice) then exhibit the same trade-off between income and status. This is a novel validation of subjec-

¹⁶Our preferred specifications without demographic controls are then free from the classic exclusion-restrictions problem that appears in tests of income-comparison effects, such as Clark and Oswald (1996). See de la Garza et al. (2010) on this point.

¹⁷We calculated post-stratification adjusted survey sampling weights, with the weights on gender and education adjusted so that the sum of the weights equals the control total for each stratum.

¹⁸These are all the demographic variables we have available in our survey data.

tive well-being data using information from an alternate method of estimating underlying preferences, via decision utility (Kahneman et al. 1997).

We check that similar results pertain when we use sampling weights in column (2) and when we estimate ordered probits as in Luttmer (2005) in column (3). Adding demographic variables to the regression changes the SR by around 30%, as can be seen in column (4).¹⁹ While this change from the estimate in column (1) is significant, and it is common to control for individual characteristics in happiness regressions, for the comparison of experienced utility and decision utility results, it is preferable to rely on the estimates in column (1).

The specifications in columns (5) and (6) of Table 3 appeal to the internal survey-based cell-mean measure of reference-group income. The standard errors here are clustered at the level of the reference-group cells. In column (5), where we do not use labour-force status to estimate cell average wages, we obtain a different value of SR to that in Yamada and Sato (2013). In column (6), where we do use labour-force status to calculate cell average wages, we obtain a value of $SR = -0.578$ similar to that of $\gamma = -0.458$. Finally, column (7) turns to the external cell-mean income information from the BSWS, as described above.²⁰ The sign and size of the estimated comparison-income coefficient is not sharply different from that in Column (1).

To carry out a formal test of the equality of γ from the hypothetical choices and SR from income satisfaction, we calculate the variances of γ and SR using the Delta method. Since both γ and SR take the form r/s , where r and s are stochastic variables, the variance is obtained as follows:

$$Var = \begin{pmatrix} \frac{\partial \gamma}{\partial r} & \frac{\partial \gamma}{\partial s} \end{pmatrix} \begin{pmatrix} Var(r) & Cov(r, s) \\ Cov(r, s) & Var(s) \end{pmatrix} \begin{pmatrix} \frac{\partial \gamma}{\partial r} \\ \frac{\partial \gamma}{\partial s} \end{pmatrix} = \frac{1}{\beta_s^2} Var(r) + \frac{\beta_r^2}{\beta_s^4} Var(\beta_s) - \frac{2\beta_r}{\beta_s^3} Cov(r, s),$$

where β_r and β_s are the averages and $Var(r)$ and $Var(s)$ the variances of r and s , respectively, and $Cov(r, s)$ is the covariance of r and s . The standard errors thus obtained are reported in Tables 2 and 3. The last row of Table 3 shows the t-statistics for the test

¹⁹This result is robust to the exclusion of the individual parameter of income comparisons, γ_i . The estimated coefficients on the other explanatory variables in column (4), which are available on request, are reassuringly similar to those in existing work (which has mostly used data from Western countries). Women report higher satisfaction scores than do men, while satisfaction is U-shaped in age, with a minimum at about age 46. Education in Japan is strongly positively correlated with income satisfaction, even controlling for income itself. Last, marriage is found to be positively correlated with satisfaction in our cross-section data. See Clark et al. (2013) for more details.

²⁰The number of observation drops here, as we do not have reference-group income information for some particular groups, for example housewives.

$\gamma = SR$. The equality of the trade-off between own and others' income between the hypothetical choices and the satisfaction regressions is accepted in six out of seven cases: this is our paper's main result. The only equality test that does not pass is that when the reference-group income is calculated within the survey by age, sex and education. This may seem surprising as the respondent-supplied reference-group income explicitly asks individuals their estimate of "the average personal income for people of the same age, sex, and education as you". However, we can see from column (6) that the $\gamma = SR$ test passes when we construct reference group income within the survey by age, sex and education **and** labour-force status. One interpretation is that individuals think of people who are in the same labour-force situation as them when they evaluate reference-group income.²¹

We can carry out a similar exercise using the results of the mixed logit model in Table 2. Here, the SR values estimated in Table 3 are mostly significantly different to the value of γ in Table 3. However, perhaps the main point from all of the estimations we present here is that the ratio of the effect of reference group to own income is mostly found to be between -0.4 and -0.7. On average, our results suggest that about half of the effect of a rise in one's own income will be dissipated if others experience the same rise in income as well.²²

3.2 Discussion

Our main results above suggest that the trade-offs between own and reference-group income are similar in hypothetical choice experiments and in satisfaction regressions with a variety of measures of reference-group income. The ratio of the estimated coefficients on reference-group and own income is almost always in the range -0.4 to -0.6. Income comparisons then matter, but the relativity of income in well-being is not total, so that a rise in both own and reference-group income does increase income satisfaction.

Our conclusion that trade-offs (regarding absolute and relative income) in choice and in satisfaction are similar is unlikely to be the final word. Our exercise here is not free from imperfections. First, the definitions of the reference group are different: the hypothetical-choice data is based on comparisons to the "overall average monthly income in the society," whereas in the satisfaction regression we elicited information of reference group defined

²¹In Clark and Senik (2010), most individuals chose their work colleagues when asked to whom they compared their own incomes.

²²We can reject the hypothesis that the SR is under 0.5 in absolute value in six out of the seven columns in Table 3

over age, gender, and education (which we have to do in order to have any variation in reference-group income across respondents). Second, the income variables are continuous in the choice experiments, while they are only categorical in the happiness regressions. de la Garza et al. (2010) analyse Japanese wage data and find that the results of Mincer wage equations are not that different when using continuous income data and income as the midpoint values of banded data.²³ A third point is that the analysis of choice data requires non-linear regressions, such as the conditional or mixed logit models that appear in the first two columns of Table 2. However, the satisfaction regressions are mostly estimated by OLS, so that the assumptions regarding the nature of the error terms is different. To address this issue, we show the results from the estimation of a random effects linear probability model in column (3) of Table 2. The error term here is assumed to follow a normal distribution, as in OLS.²⁴ The estimated trade-off in this linear-probability model is around -0.75.²⁵ Equally, we can recode income satisfaction into a binary variable (grouping responses of 1 and 2, and 3-5) and estimate the regression specifications in Table 3 (apart from the ordered probit in column 3) as logits. This produces similar SR figures (the SR ratios differ by 20% or less from those in Table 3, except that in column 6 which becomes one-third larger).

It may also be possible to use the information here to say something about who is in individuals' reference groups. If we have measures of the income of those with the same education, then those of the same sex, then those of the same age etc., we could estimate how self-reported reference-group income relates to these different measures. Say that we find that, conditional on the income measures by other demographics, the income of those in one's own labour-force status is not correlated with self-reported reference-group income, then we can conclude that comparisons take place across all labour-force statuses, and not within one's own labour-force status. This is a potentially useful avenue for future research.

²³See Luechinger and Raschky (2009) for how to reduce biases in estimates when income data are categorical.

²⁴The observations are grouped by question, and the robust standard errors are clustered at the individual level.

²⁵Following the same logic as in the previous-subsection, we can test whether the SR values in Table 3 are equal to this estimate of γ . This equality is rejected for five out of the seven SR estimates in Table 3.

4 Conclusion

This paper has used relatively large-scale internet survey data from Japan to lift part of the veil regarding the process of income comparisons. We showed that the results from happiness regressions addressing income comparisons can be argued to yield reliable results, as they concur with those from the hypothetical-choice experiments in Yamada and Sato (2013). This applies both to cell-mean measures of others' income, and the individual's self-reported estimate of what others earn. One implication is that worries about the endogeneity of self-reported reference-group income (I feel unhappy, so I think that others earn more) may be over-stated, as the trade-offs between own and others' income are so similar in satisfaction and hypothetical-choice regressions. The central estimate of the importance of others' income relative to own income is around one-half. As such, utility in our Japanese sample is indeed relative in income, but not completely so.

In our survey, "relevant others" were defined by age, sex and education. While we have shown that this definition of the reference group does seem to capture useful information, there is no reason why the same question cannot be asked about the neighborhood or the workplace. In general, questions about self-reported comparison income may well be a simple useful addition to current surveys, in the same way that the addition of subjective questions to many surveys has helped contribute to the current outpouring of research across the social sciences on individual and societal well-being.

References

- ABDALLAH, S., S. THOMPSON AND N. MARKS, “Estimating worldwide life satisfaction,” *Ecological Economics* 65 (March 2008), 35–47.
- AKAY, A., O. BARGAIN AND X. J. TAMAYO, “Back to Bentham: should we? Large scale comparison of decision and experience utility,” mimeo, Aix-Marseille Universite, January 2015.
- BENJAMIN, D. J., O. HEFFETZ, M. S. KIMBALL AND A. REES-JONES, “What Do You Think Would Make You Happier? What Do You Think You Would Choose?,” *American Economic Review* 102 (August 2012), 2083–2110.
- , “Can Marginal Rates of Substitution Be Inferred from Happiness Data? Evidence from Residency Choices,” *American Economic Review* 104 (November 2014), 3498–3528.
- CAPPELLI, P. AND P. D. SHERER, “Satisfaction, Market Wages, & Labor Relations: An Airline Study,” *Industrial Relations* 27 (Winter 1988), 56–73.
- CLARK, A. E., “What Really Matters in a Job? Hedonic Measurement Using Quit Data,” *Labour Economics* 8 (May 2001), 223–242.
- CLARK, A. E., P. FRIJTERS AND M. A. SHIELDS, “Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles,” *Journal of Economic Literature* 46 (March 2008), 95–144.
- CLARK, A. E., N. KRISTENSEN AND N. WESTERGARD-NIELSEN, “Economic Satisfaction and Income Rank in Small Neighbourhoods,” *Journal of the European Economic Association* 7 (04-05 2009a), 519–527.
- , “Job Satisfaction and Co-worker Wages: Status or Signal?,” *Economic Journal* 119 (March 2009b), 430–447.
- CLARK, A. E. AND A. J. OSWALD, “Satisfaction and comparison income,” *Journal of Public Economics* 61 (September 1996), 359–381.
- CLARK, A. E. AND C. SENIK, “Who compares to whom? The anatomy of income comparisons in Europe,” *Economic Journal* 120 (2010), 573–594.

- CLARK, A. E., C. SENIK AND K. YAMADA, “The Joneses in Japan: Income Comparisons and Financial Satisfaction,” ISER Discussion Paper 0866, Institute of Social and Economic Research, Osaka University, March 2013.
- DE LA GARZA, A., G. MASTROBUONI, A. SANNABE AND K. YAMADA, “The Relative Utility Hypothesis With and Without Self-reported Reference Wages,” ISER Discussion Paper 0798, Institute of Social and Economic Research, Osaka University, November 2010.
- EASTERLIN, R. A., “Will raising the incomes of all increase the happiness of all?,” *Journal of Economic Behavior & Organization* 27 (June 1995), 35–47.
- , “Income and Happiness: Towards an Unified Theory,” *Economic Journal* 111 (July 2001), 465–84.
- FERRER-I CARBONELL, A., B. M. S. VAN PRAAG AND I. THEODOSSIOU, “Vignette Equivalence and Response Consistency: The Case of Job Satisfaction,” IZA Discussion Papers 6174, Institute for the Study of Labor (IZA), November 2011.
- GOERKE, L. AND M. PANNENBERG, “Direct Evidence on Income Comparisons and Subjective Well-Being,” IAAEU Discussion Papers 201303, Institute of Labour Law and Industrial Relations in the European Union (IAAEU), March 2013.
- HUBER, J. AND K. ZWERINA, “The importance of utility balance in efficient choice designs,” *Journal of Marketing Research* 33 (1996), 307–317.
- KAHNEMAN, D., P. P. WAKKER AND R. SARIN, “Back to Bentham? Explorations of Experienced Utility,” *Quarterly Journal of Economics* 112 (May 1997), 375–405.
- KNIGHT, J., L. SONG AND R. GUNATILAKA, “Subjective well-being and its determinants in rural China,” *China Economic Review* 20 (December 2009), 635–649.
- LOUVIERE, J. J., D. A. HENSHER AND J. D. SWAIT, *Stated Choice Methods* (Cambridge University Press (Cambridge), 2000).
- LUECHINGER, S. AND P. A. RASCHKY, “Valuing flood disasters using the life satisfaction approach,” *Journal of Public Economics* 93 (April 2009), 620–633.
- LUTTMER, E. F. P., “Neighbors as Negatives: Relative Earnings and Well-Being,” *Quarterly Journal of Economics* 120 (August 2005), 963–1002.

- MANSKI, C. F., "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies* 60 (July 1993), 531–42.
- MAYRAZ, G., J. SCHUPP AND G. G. WAGNER, "Life Satisfaction and Relative Income: Perceptions and Evidence," CEP Discussion Papers dp0938, Centre for Economic Performance, LSE, July 2009.
- MCFADDEN, D., *Conditional Logit Analysis of Qualitative Choice Behavior in P. Zarembka (ed.) Frontiers in Econometrics* (Academic Press, 1974).
- MEADE, A. AND S. CRAIG, "Identifying Careless Responses in Survey Data," *Psychological Methods* 17 (September 2012), 437–55.
- OSWALD, A. AND S. WU, "Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the U.S.A.," *Science* 327 (January 2010), 576–579.
- PEREZ-TRUGLIA, R., "A Samuelsonian Validation Test for Happiness Data," SSRN WP 1658747, 2015.
- REVELT, D. AND K. TRAIN, "Mixed Logit With Repeated Choices: Households' Choices Of Appliance Efficiency Level," *Review of Economics and Statistics* 80 (November 1998), 647–657.
- SENIK, C., "Direct evidence on income comparisons and their welfare effects," *Journal of Economic Behavior & Organization* 72 (October 2009), 408–424.
- SLOANE, P. J. AND H. WILLIAMS, "Job Satisfaction, Comparison Earnings, and Gender," *Labour* 14 (September 2000), 473–502.
- STEVENSON, B. AND J. WOLFERS, "Economic Growth and Subjective Well-Being: Re-assessing the Easterlin Paradox," *Brookings Papers on Economic Activity* 39 (Spring 2008), 1–87.
- 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.

Table 1: Descriptive Statistics

	(1)		(2)	(3)	
	Our Survey (analysis sample)		NIKKEI	National Data (b)	
	Male	Female		Male	Female
Income satisfaction (1-5)	2.2	2.3			
Female		55.6	56.6		51.3
Age category					
20-29	18.1	21.5	13.7	19.5	18.7
30-39	25.4	23.2	36.4	24.5	24.1
40-49	19.2	24.8	30.1	21.8	21.7
50-59	22.5	18.9	13.9	22.1	22.6
60-65	14.8	11.7	5.9	12.1	12.8
Education (a)			N.A.		
Middle school	0.9	0.9		18.2	20.8
High school	21.2	25.9		41.6	43.4
Some college	11.1	32.1		11.4	24.5
College	66.7	41.1		28.3	10.7
Marital Status					
Single	33.2	25.8	30.0	32.0	23.4
Married	63.3	67.9	60.7	61.8	57.6
Divorced/Separated/Widowed	3.5	6.3	9.3	6.2	19.0
Region					
Hokkaido		4.7	4.0		4.3
Tohoku		4.4	4.1		7.4
Kanto		44.2	46.9		32.9
Koshinetsu		4.2	3.7		6.7
Chubu		10.0	9.5		11.9
Kansai		20.6	19.7		16.3
Chugoku		3.9	3.8		6.0
Shikoku		1.9	1.9		3.1
Kyushu		6.3	6.4		11.4
Student		3.4	N.A.		7.6
Unemployment rate		4.0	N.A.		4.9
Annual income (Thousand Yen)					
Own income in survey	4,902	2,709			
Reference income in survey	5,993	4,458			
Income of Cell average (c)	4,902	2,709			
Income of cell average from External source (d)	5,555	3,237			

All figures except for annual income (in million JPY) are percentages for each category.

(a) Those who are currently students are excluded from this figure.

(b) The demographic characteristics come from the Population Estimates by the Statistics Bureau (Sep. 2009); education data are from the Employment Status Survey (Table 3; 2007) from the Statistics Bureau; marital-status data are obtained from the Population Statistics of Japan (Table 6.21; 2008) from the National Institute of Population and Social Security Research; region data come from the Population Statistics of Japan (Table 9.5; 2008); income information is from the Employment Status Survey (2008); and unemployment data come from the Labour Force Survey (Feb. 2010) from the Statistics Bureau.

(c) Cells defined by age, education, and gender.

(d) Cells defined by age, education, gender, and labour-force status.

Table 2: Regression Results from the Hypothetical Choice (Decision Utility) Experiments

Model	(1)	(2)	(3)
Dep. Var: Utility	Conditional logit	Mixed logit	Linear Probability
Own income y	0.048*** (0.001)	0.053*** (0.002)	0.043*** (0.004)
Reference income \bar{y}	-0.022*** (0.001)	-0.036*** (0.001)	-0.033*** (0.005)
Estimated γ	-0.458	-0.676	-0.766
S.E. of γ from the delta method	(0.012)	(0.018)	(0.014)
Observations	48172	48172	48172
Pseudo R-squared	0.249		0.082

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes. Robust standard errors clustered by subject in parentheses. The results in columns (1) and (2) are reproduced from Yamada and Sato (2013).

Table 3: Income Satisfaction (Experienced Utility) Regression Results with various reference-income measures

Type of reference income	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Respondent-supplied			Within-survey 1		Cell mean	
Dep. Var	Income			Within-survey 1		Within-survey 2	
External:				Within-survey 1		Within-survey 2	
				Within-survey 1		Within-survey 2	
Own income y	0.290*** (0.016)	0.256*** (0.019)	0.362*** (0.020)	0.354*** (0.020)	0.330*** (0.054)	0.320*** (0.050)	0.466*** (0.040)
Reference income \bar{y}	-0.153*** (0.020)	-0.131*** (0.025)	-0.193*** (0.024)	-0.137*** (0.021)	-0.335*** (0.077)	-0.185*** (0.064)	-0.195*** (0.034)
Constant	0.272 (0.273)	0.502 (0.318)		-0.290 (0.361)	2.371*** (0.749)	0.257 (0.592)	-1.819*** (0.452)
Other Controls	No	No	No	Yes	No	No	No
Observations	10203	10203	10203	10203	10203	10203	7880
R-squared	0.032	0.026	0.013	0.078	0.037	0.030	0.061
Ratio of effect of \bar{y} to that of y (SR)	-0.528 (0.059)	-0.512 (0.081)	-0.532 (0.058)	-0.387 (0.063)	-1.015 (0.131)	-0.578 (0.140)	-0.418 (0.055)
S.E. of SR from Delta method							
Test: SR = the value of γ from Yamada and Sato (2013)	1.11	0.63	1.18	1.18	4.21	0.84	0.76
T-statistics							

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: These income satisfaction regressions are estimated by OLS, except for column (3) which shows the ordered probit estimates using the same specification as in column (1). The model in column (2) is identical to that in column (1) but applies post-stratification sampling weights. Cell-mean averages within the survey are defined over (age, sex, education) in column (5), and (age, sex, labour-force status and education) in columns (6) and (7). The standard errors are clustered by (age, sex, education) in column (5), and (age, sex, labour-force status and education) in columns (6) and (7). The other controls in column (4) are gender, age, age-squared, education dummies (5), marital-status dummies (2), region dummies (8), and individual values of γ_i (the individual values of preferences over income comparisons) from Yamada and Sato (2013).^a The number of observations in columns (7) is lower than the other columns as we have no suitable external information for the income of housewives.

^aWe introduce this latter to reflect some kind of individual-specific fixed effect

Appendix: Questionnaire

Satisfaction and Income-Comparison Questions:

1. *How satisfied are you with the amount of your current income?*

1. Extremely dissatisfied
2. Dissatisfied
3. Neither dissatisfied nor satisfied
4. Satisfied
5. Extremely satisfied

2. *How much are you concerned, anxious or envious about other people's income?*

1. Not at all
2. Not much
3. Neither yes nor no
4. Fairly much
5. Very much

3. *When you compare the amount of your income with that of someone else, which of the following best describes that person?*

1. Family
2. Neighbour
3. Friend
4. Colleague
5. Don't compare
6. Others

Demographic Variable Questions:

1. What is your age in years?
2. What is your sex?
3. What is your highest educational qualification?
4. What is your occupation?
5. What is your marital status?
6. What best describes your current residence? (Type of dwelling, for example a house that you own)
7. Where do you live? (Which prefecture in Japan)
8. What was your approximate personal income (before taxes) in 2009? Please also include pension income, remittances, and so on.
9. About how much do you suppose was the average personal income (before taxes) in 2009, for people of the same age, sex, and education level as you?