Do People Seek to Maximize Happiness? Evidence from New Surveys*

Daniel J. Benjamin
Cornell University and NBER

Ori Heffetz
Cornell University

Miles S. Kimball
University of Michigan and NBER

Alex Rees-Jones
Cornell University

First Draft: July 26, 2010
This Draft: April 2, 2011

Abstract

Are subjective well-being (SWB) measures a good empirical proxy for utility? We evaluate one necessary assumption: that people's preferences coincide with what they predict will maximize their SWB. Our method is to present survey respondents with hypothetical scenarios and elicit both choice and predicted SWB rankings of two alternatives. While choice and predicted SWB rankings usually coincide, we find systematic reversals. Furthermore, we identify factors—such as predicted sense of purpose, control over one's life, family happiness, and social status—that help explain choice controlling for predicted SWB. We explore how our findings vary with the SWB measure and the choice situation.

JEL Classification: D03, D60

Keywords: happiness, life satisfaction, subjective well-being, utility

* We are extremely grateful to Dr. Robert Rees-Jones and his office staff for generously allowing us to survey their patients and to Cornell's Survey Research Institute for allowing us to put questions in the 2009 Cornell National Social Survey. We thank Gregory Besharov, John Ham, Benjamin Ho, Erzo F. P. Luttmer, Michael McBride, Ted O'Donoghue, Matthew Rabin, Antonio Rangel, and Robert J. Willis for especially valuable early comments and suggestions. We are grateful to participants at the CSIP Workshop on Happiness and the Economy held at the Federal Reserve Bank of San Francisco, participants at the Cornell Behavioral/Experimental Lab Meetings, and seminar audiences at Cornell University, Deakin University, Syracuse University, NBER, the Stanford Institute for Theoretical Economics (SITE), and Florida State University for helpful comments. We thank Eric Bastine, J.R. Cho, Kristen Cooper, Isabel Fay, John Farragut, Geoffrey Fisher, Sean Garborg, Arjun Gokhale, Jesse Gould, Kailash Gupta, Han Jiang, Justin Kang, June Kim, Nathan McMahon, Elliot Mandell, Cameron McConkey, Greg Muenzen, Desmond Ong, Mihir Patel, John Schemitsch, Brian Scott, Abhishek Shah, James Sherman, Dennis Shiraev, Elizabeth Traux, Charles Whittaker, Brehnen Wong, Meng Xue, and Muxin Yu for their research assistance. We thank the National Institute on Aging (grant P01-AG026571/01) for financial support.
E-mail: db468@cornell.edu, oh33@cornell.edu, mkimball@umich.edu, arr34@cornell.edu.

All things considered, how satisfied are you with your life as a whole these days?

Taken all together, how would you say things are these days—would you say that you are very happy, pretty happy, or not too happy?

Much of the time during the past week, you felt you were happy. Would you say yes or no?[1]

Economists increasingly use survey-based measures of subjective well-being (SWB) as an empirical proxy for utility. In some applications, SWB data are used for conducting welfare evaluations that are difficult to conduct with standard revealed preference methods. Examples include measuring the negative externality from neighbors' higher earnings (Luttmer, 2005), individuals' tradeoff between inflation and unemployment (DiTella, MacCulloch, and Oswald, 2003), and the effect of health status on the marginal utility of consumption (Finkelstein, Luttmer, and Notowidigdo, 2008). In order for such SWB-based evaluations to assess welfare as conceived in the standard economic approach, a necessary assumption is that an agent's preferences over alternatives and her measured SWB under each alternative yield identical rankings of the alternatives. This paper provides evidence for evaluating that assumption. Put differently but equivalently, the goal of this paper is to assess whether individuals seek to maximize SWB (as commonly measured).

In pursuing that goal, this paper also empirically addresses two opposing theoretical views regarding the relationship between SWB and utility. The first, reflected at least implicitly in a large number of papers (e.g., Gruber and Mullainathan, 2005; Frey, Luechinger, and Stutzer, 2008; Welsch, 2009), is that SWB data represent idealized revealed-preference utility in the sense of what individuals would choose if they were well-informed about the consequences of their choices for SWB. The second view, explicitly laid out in Kimball and Willis (2006) and Becker and Rayo (2008), is that even well-informed agents will be willing to trade off SWB for other things they care about, making SWB and utility distinct.[2] This paper assesses these two

[1] The first of these three questions is from the World Values Survey (WVS); similar questions appear in the Euro-Barometer Survey, the European Social Survey, the German Socioeconomic Panel (G-SOEP), and the Japanese Life in Nation survey. The second question is from the U.S. General Social Survey (GSS); similar questions appear in the Euro-Barometer survey, the National Survey of Families and Households (NSFH), and the WVS. The third question is from the University of Michigan's Survey of Consumers; similar questions appear in the Center of Epidemiologic Studies Depression Scale (CES-D), the Health and Retirement Study (HRS), and the Gallup-Healthways Well-Being Index (GHWBI).

[2] For other elaborations of this view, see also Kimball, Levy, Ohtake and Tsutsui (2006), Tsutsui, Kimball, and Ohtake (2010), and Kimball, Nunn and Silverman (2010).

views empirically by measuring the concordance between choices and what individuals think will maximize their SWB.[3]

We pose a variety of hypothetical decision scenarios to three respondent populations: a convenience sample of 1,066 adults, a representative sample of 1,000 adult Americans, and 633 students. Each scenario has two alternatives. For example, one scenario describes a choice between a job that pays less but allows more sleep versus a job with higher pay and less sleep. We ask respondents which alternative they would choose. We also ask them under which alternative they expect greater SWB; we assess this "predicted SWB" using measures based on each of the three commonly-used SWB questions posed in the epigraph above. Importantly, since people often mispredict the happiness consequences of their choices (e.g., Gilbert, 2006), we compare an individual's choices with her *ex ante beliefs* regarding her choices' effects on SWB.[4]

We have two main results. First, we find that overall, respondents' SWB predictions are a powerful predictor of their choices. On average, predicted SWB and choice coincide in our data 83 percent of the time. We find that the strength of this relationship varies across choice situations, subject populations, survey methods, questionnaire structure variations, and measures of SWB, with coincidence ranging from well below 50 percent to above 95 percent.

Our second main result is that discrepancies between choice and predicted SWB rankings are systematic. Moreover, we can identify other, non-SWB factors that help explain respondents' choices. In some surveys, in addition to measuring participants' choices and predicted SWB, we also ask participants to predict how each alternative would affect particular aspects of life other than SWB. The aspects that systematically contribute most to explaining choice, *controlling for SWB*, are sense of purpose, control over life, family happiness, and social status. At the same time, and in line with our first main result above, when we examine how well SWB compares with the other factors we measure, we find that across our scenarios, populations, and methods,

[3] Some researchers adopt yet a third view, the philosophical position that happiness *per se* is the appropriate maximand for a policymaker (e.g., Bentham, 1789; Layard, 2005; Bronsteen, Buccafusco, and Masur, forthcoming). See Bernheim (2009) for a categorization and discussion of alternative philosophical approaches to welfare economics. Throughout this paper, we adopt the standard economic perspective (common to the two views mentioned in the text) that the appropriate maximand for welfare analysis is an individual's preferences. However, the value of our results does not hinge on this perspective. Even if one believes that happiness is the appropriate maximand, one would like to know under what circumstances people choose to maximize it on their own, and under what circumstances a policymaker would have to intervene in order to maximize it.

[4] In the terminology of Kahneman, Wakker, and Sarin (1997), we compare "decision utility" (what people choose) with "predicted utility" (what people predict will make them happier).

SWB is by far the single best predictor of choice.

These results provide a nuanced answer to our motivating question. People do not exclusively seek to maximize SWB (as operationalized by common SWB measures), yet SWB does seem to be a uniquely important argument in the utility function. Furthermore, as discussed shortly, the choice-SWB relationship varies systematically with scenario and SWB measure.

We use a variety of survey versions and empirical approaches in order to test the robustness of our main results to alternative interpretations. For example, while most of our data are gathered by eliciting both choice and predicted SWB rankings from each respondent, in some of our survey variations we elicit the two rankings far apart in the survey, or we elicit only choice ranking from some participants and only SWB ranking from others. To take another example, we assess the impact of measurement error by administering the same survey twice (at two separate times) to some of our respondents. While these different approaches affect our point estimates and hence the relative importance of our two main results, both results appear to be robust.

As steps toward providing practical, measure-specific and situation-specific guidance to empirical researchers as to *when* the assumption that people try to maximize predicted SWB is a better or worse approximation, we analyze how our results differ across SWB measures and across scenarios. Comparing SWB measures, we find that a "life satisfaction" measure (modeled after the first question in the epigraph) is a better predictor of choice than either of two "happiness" measures (modeled after the second and third questions in the epigraph) which perform similarly to each other. Comparing scenarios, we find that in scenarios constructed to resemble what our student respondents judge as representative of important decisions in their lives, predicted SWB coincides *least* often with choice, and other factors add relatively more explanatory power. This suggests that caution is especially warranted in drawing welfare conclusions from SWB data in such situations. Importantly, we also find that in scenarios where one alternative offers more money, respondents are systematically more likely to choose the money alternative than they are likely to predict it will yield higher SWB. This suggests that survey-based SWB measures may understate the utility benefits of higher income.

Our work is related to a literature in philosophy that poses thought experiments in hypothetical scenarios in order to demonstrate that people's preferences encompass more than their own happiness. That literature focuses on extreme situations such as being hooked up to a

machine that guarantees happiness.[5] In contrast, by focusing on realistic choice situations rather than on extreme ones, and by focusing on commonly-used SWB measures rather than on the concept of happiness, we seek to assess to what extent measured happiness is a good proxy for preferences in empirically-relevant situations.[6] As far as we are aware, the work whose methodology is most similar to ours is one of the experiments in Tversky and Griffin (1991), but their research question is different.[7]

The paper is organized as follows. Section I discusses the survey design and subject populations. Section II asks whether participants choose the alternative in our decision scenarios that they predict will generate greater SWB. Section III asks whether aspects of life other than SWB help predict choice, controlling for SWB, and compares the relative predictive power of the factors that matter for choice. Section IV presents robustness analyses. Section V discusses additional results. Section VI concludes, and discusses other possible applications of our methodology. Finally, the Appendix lists our decision scenarios, and the Web Appendix provides a chronology of all pilots; details information on all survey instruments; describes all of our robustness analyses and reports all the resulting tables; and reports an extensive set of additional analyses and results.

---

[5] In what is perhaps the most famous thought experiment, a reader is asked to introspect whether he or she would choose to be hooked up to an "experience machine" that would guarantee complete happiness for the rest of life, without awareness that the experience is not real (Nozick, 1974, pp.42-45). It is argued that if a reader would refuse the machine, he or she must not care exclusively about happiness. Less fancifully, extreme examples such as sacrificing one's life for a cause one believes in may also suggest that happiness is not the only maximand.

[6] Like us, Becker and Rayo (2008) propose empirical tests of whether things other than happiness matter for preferences in empirically-relevant choice situations. Relatedly, Truglia (2010) tests empirically whether the utility function inferred from consumption choices is distinguishable from the estimated happiness function over consumption. In contrast to our approach, where we measure individuals' predictions about SWB, these alternative tests rely on *ex post* measured happiness. Hence these tests assume that people correctly predict the happiness consequences of their choices, while our tests do not. However, we still interpret individuals' choices as revealing their "true preferences"—i.e. their assessment of their welfare—given their (potentially mistaken) beliefs about the consequences of their choices. This interpretation may be unjustified in situations where, e.g., due to social or moral pressure, choices do not reveal "true preferences" (see Koszegi and Rabin, 2008). To the extent that the decision scenarios we study involve such situations, our finding that people sometimes make choices that do not maximize predicted SWB may overstate the discrepancy between "true preferences" (or *welfare*) and SWB. However, our finding that people often choose an option that generates higher income despite predicting lower SWB for that option leans against this view; social or moral pressure would probably dictate choosing the *lower*-income option.

[7] In a between-subjects experiment with 66 undergraduates, Tversky and Griffin's participants were presented with a hypothetical choice between a job with a higher salary where co-workers earn even more versus a job with a lower salary where co-workers earn even less. Consistent with what we find in a similar choice problem, Tversky and Griffin find that participants are more likely to choose the higher absolute salary but say they would be happier with the higher relative salary. Tversky and Griffin interpret the result as supporting their theory that payoff levels are weighted more heavily in choice, while contrasts between payoffs and a reference point are weighted more heavily in SWB judgments.

I. Survey Design

While our main evidence is based on 29 different survey-versions, all versions share a similar underlying structure. Respondents are presented with a sequence of hypothetical scenarios, and in each they face a choice between two options. To illustrate, our 'Scenario 1' highlights a tradeoff between sleep and income. Followed by its SWB and choice questions, it appears on one of our questionnaires as follows:

> Say you have to decide between two new jobs. The jobs are exactly the same in almost every way, but have different work hours and pay different amounts.
>
> **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.
>
> Between these two options, taking all things together, which do you think would give you a happier life as a whole?

Option 1: Sleep more but earn less			**Option 2: Sleep less but earn more**		
definitely happier	*probably happier*	*possibly happier*	*possibly happier*	*probably happier*	*definitely happier*
X	X	X	X	X	X
Please circle one X in the line above					

> If you were limited to these two options, which **do you think you would choose**?

Option 1: Sleep more but earn less			**Option 2: Sleep less but earn more**		
definitely choose	*probably choose*	*possibly choose*	*possibly choose*	*probably choose*	*definitely choose*
X	X	X	X	X	X
Please circle one X in the line above					

In within-subject questionnaires, respondents are asked both the SWB question and the choice question above. In between-subjects questionnaires, respondents are initially or exclusively asked only one of the two questions.

I.A. Populations and Studies

We conducted surveys among 2,699 respondents from three populations: 1,066 patients at a doctor's waiting room in Denver who voluntarily filled out our questionnaires while waiting for their appointment; 1,000 adults who participated by telephone in the 2009 Cornell National Social Survey (CNSS) and form a nationally representative sample;[8] and 633 Cornell students who were recruited on campus and participated for pay or for course credit.[9] The Denver and Cornell studies include both within-subject and between-subjects survey variants, while the CNSS study is exclusively within-subject.[10]

Table 1 summarizes the design details of these studies. It lists each study's respondent population, sample size, scenarios used (see I.B below), types of questions asked (see I.C below), and other details such as response scales, scenario order, and question order.[11] The rest of this section explains the details summarized in the table.

I.B. Scenarios

Our full set of 13 scenarios is given in the Appendix, starting with Scenario 1 ("sleep versus income") from the example above. Referring to scenarios by their scenario number (1 through 13), table 1 reports which scenarios are used in which studies, and in what order they appear on different questionnaires. As detailed in the Appendix, some scenarios are asked in different variations (e.g. different wording, different quantities of money/sleep/commute time, etc.) and some scenarios are tailored to different respondent populations (e.g., while we ask students about school, we ask older respondents about work). In constructing the scenarios, we were guided by four considerations.

First, we chose scenarios that highlight tradeoffs between options that the literature suggests might be important determinants of SWB. Hence, respondents face choices between jobs and housing options that are more attractive financially versus ones that allow for: in

[8] The CNSS is an annual survey conducted by Cornell University's Survey Research Institute. For details: https://sri.cornell.edu/SRI/cnss.cfm.

[9] We conducted additional, smaller, studies among another 924 participants. Briefly, they included a pilot of the Cornell surveys with University of Chicago Students (102 respondents), a survey of Cornell students eliciting important decisions they face (171 respondents), as well as two surveys among the Denver sample (429 respondents) and a survey among Cornell students (222 respondents) that measured willingness-to-pay for greater SWB, among other things not related to this paper. For additional details, see the Web Appendix.

[10] In our Cornell between-subjects surveys, we use a design that allows us to also elicit within-subject data from our respondents, as described in section III.

[11] The median age in our Denver, CNSS, and Cornell samples is, respectively, 47, 49, and 21; the share of female respondents is 76, 53, and 60 percent. For summary statistics, see Web Appendix table A3.

Scenario 1, more sleep (Kahneman et al., 2004; Kelly, 2004); in Scenario 12, a shorter commute (Stutzer and Frey, 2008); in 13, being around friends (Kahneman et al., 2004); and in 3, making more money *relative* to others (Luttmer, 2005; see Heffetz and Frank, forthcoming, for a survey).

Second, since some of us were initially unsure we would find any divergences between predicted choice and SWB, in our earlier surveys we focused on choice situations where one's SWB may not be the only consideration. Hence, in Scenario 4 respondents choose between a career path that promises an "easier" life with fewer sacrifices versus one that promises posthumous impact and fame, and in Scenarios 2 and 11 they choose between a more convenient or "fun" option versus an option that might be considered "the right thing to do."

Third, once we found divergences between predicted SWB and choice, in our later surveys (the Cornell studies) we wanted to assess the magnitude of these divergences in scenarios that are representative of important decisions faced by our respondent population. For this purpose we asked a sample of students to list the three top decisions they made in the last day, the last month, the last two years, and in their whole lives.[12] Naturally, decisions that were frequently mentioned by respondents revolved around studying, working, socializing and sleeping. Hence, in the resulting Scenarios 7-10, individuals have to choose between socializing and fun versus sleep and schoolwork; traveling home for Thanksgiving versus saving the airfare money; attending a more fun and social college versus a highly selective one; and following one's passion versus pursuing a more practical career path. To these scenarios we added Scenario 6, where respondents choose between giving up two hours a week versus paying a weekly sum of money. With this relatively pedestrian scenario we also hoped to get an insight into our subjects' time-versus-money tradeoff.

Fourth, as an informal check on our methods, we wanted to have one falsification-test scenario where we expected a respondent's choice and SWB ratings to coincide. For this purpose, we added Scenario 5, in which respondents face a choice between two food items (apple versus orange) that are offered for free and for immediate consumption. Since we carefully attempted to avoid any non-SWB differences between the options, we hypothesized that in this scenario, predicted SWB would most strongly predict choice. (Notice that this scenario has the additional attraction of being similar to prevalent decisions in almost everyone's

[12] The sample included 102 University of Chicago students; results were subsequently supported by surveying another 171 Cornell students. See the Web Appendix for full details (including the classification of responses by "decision type").

life, which is our third consideration above.)

I.C. Main Questions

(a) Choice question: In all studies, for each scenario, the choice question is presented as in our example above. In our analysis, we convert the horizontal six-point response scale into an intensity-of-choice variable, ranging from 1 to 6, or into a binary choice variable ("choose Option 1" versus "choose Option 2").[13] We view answers to this question as measures of preference, and the binary variable as a measure of (hypothetical) pairwise-choice.

(b) SWB question: While the language in the choice question is always kept the same, we vary the SWB question across questionnaire versions in order to examine how choice relates to several different, common SWB measures. In our Denver within-subject study we ask three different versions of the SWB question, modeled after what we view as three "families" of SWB questions that are commonly-used in the literature (see examples in the epigraph):

(i) life satisfaction: "Between these two options, which do you think would make you more satisfied with life, all things considered?";

(ii) happiness with life as a whole: "Between these two options, taking all things together, which do you think would give you a happier life as a whole?"; and

(iii) felt happiness: "Between these two options, during a typical week, which do you think would make you feel happier?"[14]

Since our between-subject tests have less statistical power than our within-subject tests, we ask only version (*i*) in our Denver between-subject surveys. In the CNSS study, where we were limited to one version of the SWB question due to design constraints, we ask only version (*ii*). As with the choice question, in the Denver study there are six possible answers (see the example

[13] Responses in the CNSS study are elicited as a binary choice rather than as a six-point scale, because in telephone interviews the binary format is both briefer for interviewers to convey and easier for respondents to understand.

[14] A fourth family of SWB questions, notably seen in the Gallup World Poll and Pew Global Attitudes surveys, and more recently on the HRS and its sister studies in Europe, asks respondents to rank their lives on an objective scale without using happiness or satisfaction language. For example, the Gallup World Poll asks: "Please imagine a ladder/mountain with steps numbered from zero at the bottom to ten at the top. Suppose we say that the top of the ladder/mountain represents the best possible life for you and the bottom of the ladder/mountain represents the worst possible life for you. If the top step is 10 and the bottom step is 0, on which step of the ladder/mountain do you feel you personally stand at the present time?" This family of SWB questions is used when difficulties in translation make standard SWB questions unreliable. We did not include this type of measure in our experimental surveys, but consider studying ladder/mountain measures with the tools developed here a promising topic for further research.

response-scales above) which we convert into either a six-point variable or a binary pairwise-SWB-ranking variable; in the CNSS study, response is binary.

Finally, as described shortly, in our Cornell studies we ask respondents about twelve different aspects of life, of which (one's own) happiness is only one. In that study we use versions of (*ii*) and (*iii*) that are modified to remain meaningful, with fixed wording, across aspects. The modified (*ii*) and (*iii*) result in these two new versions:

(iv) own happiness with life as a whole: "Between these two options, taking all things together, which option do you think would make your life as a whole better in terms of … [your own happiness]"; and

(v) immediately-felt own happiness: "Between these two options, in the few minutes immediately after making the choice, which option do you think would make you feel better in terms of … [your own happiness]."

The modified response scale now includes a middle "no difference" response, and has seven possible answers (*Option 1 definitely better; Option 1 probably better; Option 1 possibly better; no difference; Option 2 possibly better*, etc.). Respondents are allowed to indicate "no difference" for aspects of life that, in a given scenario, they judge irrelevant. We ask both (*iv*) and (*v*) in the Cornell within-subject surveys, and only (*iv*) in the Cornell between-subjects surveys.

(c) Questions about meta-choice and about other aspects of life: For completeness, we briefly mention here that some of our surveys contain the following additional questions. First, in all questionnaires of the Denver and Cornell within-subject studies, the choice question is followed by what we refer to as a meta-choice question: "If you were limited to these two options, which **would you want yourself to choose**?" Second, as mentioned above, the SWB question in all Cornell studies is modified to elicit ratings of the two scenario options in terms of "own happiness" as well as eleven additional aspects of life. For example, in versions (*iv*) and (*v*) of the SWB question (see language above), [your own happiness] may be followed by [your family's happiness], [your health], [your romantic life], etc.[15] We discuss these additional

[15] In some questionnaire versions, we separate "own happiness" from the other eleven aspects, and ask respondents first only about own happiness in each scenario, and then, re-presenting the scenarios, we ask about the other aspects. In these versions, we refer to the question on own happiness as an "isolated" measure of SWB (see table 1).

questions in sections III and IV, where we analyze the data they yield.

II. Do People Respond to the Choice and SWB Questions in the Same Way?

In this section we look at respondents' ranking of Option 1 versus Option 2 in terms of hypothetical choice compared with their ranking of the options in terms of predicted SWB. Viewing responses to the choice and the SWB questions as binary variables, we ask: *Do respondents answer the two questions as if they were answering the same question, or do they answer them systematically differently?* (We postpone analyzing responses as multi-point variables to section III.) Our analysis provides a first test of the null hypothesis that individuals seek to maximize SWB.

To answer our question, table 2 reports the distribution of binary responses to our surveys' choice and SWB questions by study and scenario, along with *p*-value statistics from equality-of-proportions tests. The table pools responses across SWB question variants (see I.C(b) and table 1 above); we discuss results by specific SWB measure below. The rest of this section interprets the figures reported in the table.[16]

II.A. Within-Subject Analysis

We start with table 2's leftmost column, which reports Scenario 1 figures from the Denver within-subject questionnaires (our "sleep versus income" scenario from the example in section I). The column's top four cells report a vertically-stacked 2×2 contingency matrix, consisting of the joint binary distribution of subjects who favor Option 1 in the choice question, and those who favor it in the SWB question. Looking at these four cells, we point out two facts that illustrate the present section's two main findings. First, the top two cells reveal that SWB response is highly predictive of choice response: between the two cells, 87 percent of respondents rank Option 1 versus Option 2 in the choice question the same as they do in the

---

In other versions, where the twelve aspects appear together either in the order specified above or in reverse order, we refer to the own happiness question as a "first/last in a series" measure.

[16] Non-response in our surveys was generally low. In the Cornell studies, virtually all questions had a non-response rate below 2 percent (one Cornell respondent was excluded due to obvious confusion with instructions). In the CNSS, less than 5 percent of respondents answered "Do not know" or refused to answer in any of the questions. Due to the less-structured recruiting method used in our Denver doctor's office studies, some questions from those studies had non-response rates as high as 20 percent. However, the majority of this non-response is driven by respondents being called in for their appointments, alleviating concerns of selection bias. Comparing the completed responses of subjects who did not finish the complete survey to the responses of those who finished the entire survey, we find no evidence of a difference in average responses.

SWB question. Second, the next two cells reveal systematic differences across the two questions among the remaining 13 percent of respondents: while 12 percent rank Option 1 (sleep) above Option 2 (income) in the SWB question and reverse this ranking in the choice question, only 1 percent do the opposite. This asymmetry suggests that on average, respondents react to the two questions systematically differently. The fifth cell reports the *p*-value from a paired Liddell exact test, a (nonparametric) equality-of-proportions test (Liddell, 1983).[17] The null hypothesis—namely, that the proportion of respondents who rank Option 2 above Option 1 does not depend on whether they answer the choice or the SWB question—is easily rejected.

Examining the top five rows in table 2 for the rest of the Denver within-subject columns verifies that the two main findings above are not unique to Scenario 1: in the remaining five scenarios, 81 to 90 percent of respondents rank the two options identically across the choice and SWB questions; yet in four out of five cases, choice-SWB reversals among the remaining 10 to 19 percent of respondents are asymmetric, and the equality-of-proportions null hypothesis across the two questions is easily rejected. In these cases, respondents rank income above legacy, concert above duty, low rent above short commute, and income above friends in higher proportions in the choice question than in the SWB question. Interestingly, there appears to be a systematic tendency among respondents to favor money in the choice question more than in the SWB question.

Similarly, the CNSS column suggests that qualitatively, Scenario 1's findings carry over from our Denver study—a pencil-and-paper survey with six-point response scales administered to a convenience sample—to the CNSS study—a telephone survey with binary response scales administered to a nationally representative sample. While the proportion of participants with no choice-SWB reversals increases to 92 percent, almost all of the rest—7 out of the remaining 8 percent—favor Option 1 (sleep) in the SWB question and Option 2 (income) in the choice question. The direction of this asymmetry is hence the same as in the Denver sample, and equality of proportions is again easily rejected.

Last among our within-subject data, results from the Cornell surveys are reported on the

[17] Conceptually, the Liddell exact test works as follows. For each respondent who answered both the SWB and the choice questions, one of his or her two responses is randomly (and hence, possibly falsely) reassigned to be treated as choice-response and the other as SWB-response. The difference between the proportions who favored Option 1 over Option 2 in the (reassigned) choice versus SWB questions is recorded. This procedure is then repeated for every possible permutation of assignments. The Liddell *p*-value reports the fraction of repetitions that resulted in a difference in proportions equal to or greater than the actual difference in proportions in our data. Computationally, *p*-values are calculated using an exact formula.

second page of table 2. The structure of this part of the table is similar to the corresponding Denver and CNSS parts, with the following three differences that result from the fact that the Cornell questionnaires allow for an additional "no difference" response in the SWB question: (a) an additional row below the top four rows reports the proportion of respondents who choose the "no difference" response; (b) the top four rows report vertically-stacked contingency matrices as before, only here they exclude these "no difference" responses (their sum is normalized to 100 percent after dropping these responses); and (c) the "no difference" responses are excluded from the Liddell tests.[18]

Starting again with Scenario 1 in the leftmost column, choice-SWB reversals (in the third and fourth rows, 24 percent together) are still a minority, although they are almost twice to three times more common in the Cornell sample than in the Denver and CNSS samples. Importantly, consistent with the Denver and CNSS data, in virtually all of these reversals—23 of the 24 percent—Option 1 (sleep) is ranked above Option 2 (income) in the SWB question and below it in the choice question. Equality of proportions is, again, strongly rejected for this scenario.[19]

Moving to the rest of the Cornell within-subject columns reveals a similar story. Equality of proportions is strongly rejected for all the remaining nine scenarios (2-10) as well, with the exception of Scenario 5. Recall that we constructed Scenario 5 ("apple versus orange") as a falsification test, to alert us if our methods produced choice-SWB reversals they should not. We predicted that, barring problems with our methods, choice and SWB should largely coincide for this scenario. Reassuringly, the results support this prediction. Indeed, only 5 percent of responses exhibit reversals in this scenario, by far the lowest fraction among the ten scenarios.[20]

---

[18] Since respondents must choose between Option 1 and 2 in the choice question, we cannot observe whether individuals indicating "no difference" for SWB support or weaken the equality-of-proportions null hypothesis. These individuals' responses in the choice question mirror the distribution of choice-responses among the rest of the respondents reasonably closely (Web Appendix table A5), and, reassuringly, the choice proportions in table 2 are virtually unaffected by excluding these individuals. Moreover, under the null hypothesis that choice is determined solely by predicted SWB, the distribution of choice-responses should be closer to 50-50 for individuals indicating SWB "no difference." Hence, the responses of these respondents actually provide additional suggestive evidence against the null hypothesis.

[19] Comparing each of the top four cells in the scenario 1 column across the three within-subject samples reveals that the reported proportions differ dramatically between the samples. Given the very different populations and, in the CNSS study, the very different survey methods, this finding in itself is not surprising. (For example, we speculate that since a telephone survey is harder to understand, more respondents answered the two questions in the same way, taking the "artificial consistency" mental shortcut discussed in II.B below.)

[20] Furthermore, we find no evidence that these reversals are in one systematic direction. At the same time, a sizeable 37 percent of respondents indicate "no difference" in the SWB question in scenario 5—by far the highest. This may suggest that scenario 5 is "cleaner" than we intended it to be: not only non-SWB aspects of life, but even own

As to the two other scenarios that are used in both the Denver and Cornell studies—Scenarios 3 and 4—choice-SWB reversals maintain their direction: in both studies, (absolute) income is ranked above relative income (Scenario 3) and above legacy (Scenario 4) in the choice questions more often than in the SWB questions. (While equality of proportions is rejected in the Cornell data but not in the Denver data in Scenario 3, it is rejected in both studies in Scenario 4.)

Finally, in Scenarios 6 and 8, which are used only in the Cornell studies and include a "money" option, we once again find that respondents favor money in the choice question more than in the SWB question. That this tendency holds in all seven scenarios that trade off more money/income for something else—be it more sleep, higher relative income, a legacy, a shorter commute, being around friends, having more time, or visiting family—suggests that survey-based SWB measures understate the utility benefits of higher income.

Our two main findings—that the ranking of the two options is identical across the choice and SWB questions for most respondents and in most scenarios, but that respondents react to the two questions systematically differently—hold not only in the pooled data, but also for each SWB question variant (*i*)-(*v*) separately. We show this in Web Appendix table A4, which reports versions of table 2 by SWB measure. Interestingly, we find some differences across the measures in the prevalence of choice-SWB reversals. In the Denver sample, the life satisfaction question variation (*i*) comes closest to matching choice, with only 11 percent of reversals, averaged across all scenarios. In comparison, happiness with life as a whole (*ii*) and felt happiness (*iii*) yield more reversals—17 percent each. In the Cornell sample, own happiness with life as a whole (*iv*) and immediately felt own happiness (*v*) both yield 22 percent reversals. We return to the comparison between different SWB measures in section V.A below.

II.B. Between-Subjects Analysis

Our within-subject analysis above is based on both choice and SWB responses elicited from each individual. However, empirical work that uses SWB data relies on surveys that measure SWB alone, not together with choice. Thus, two potential biases could compromise the relevance of our findings to existing SWB survey data and their applications. On the one hand,

---

happiness is deemed by many respondents irrelevant in what they may perceive as a context of *de gustibus non est disputandum*.

asking a respondent both questions might generate an "artificial consistency" between the two responses. For example, respondents might think they ought to give consistent answers, or might give consistent answers as an effort-saving mental shortcut. On the other hand, an "artificial inconsistency" bias is also possible if respondents infer from being asked more than one question that they ought to give different answers, or if the presence of the other question focuses participants' attention on the contrast between the wordings.

To address these concerns, we compare the above results from the Denver and Cornell within-subject studies with their counterpart between-subjects studies, in which respondents are asked either just the choice or just the SWB question. For ease of comparison, the between-subjects results are reported below the within-subject results on the two pages of table 2.

We start with the Denver sample. While the Denver within-subject questionnaires use the three SWB variants (*i*)-(*iii*), their between-subjects counterpart uses only variant (*i*), the life satisfaction measure. Looking again at Scenario 1, of the between-subjects respondents who are presented with this SWB question, 34 percent rank Option 2 (income) above Option 1 (sleep). Of the respondents who are presented with the choice question, the proportion is 44 percent. The direction of the discrepancy is the same as in the within-subject data and, as reported by the Fisher exact test *p*-value, the difference is statistically distinguishable from zero.[21] In contrast, the other columns show that there is essentially no difference between the responses to the choice and the SWB questions in the rest of the scenarios in the Denver between-subjects study.

Before comparing these proportion point estimates across the within- and between-subjects data, we first compare statistical results. Since the unpaired Fisher test has less power than the paired Liddell test, table 2 also reports Fisher test results for the within-subject data. They are calculated by randomly assigning our subjects into two equal groups, and examining only either their choice or their SWB response (hence "unpairing" these data). The reported average *p*-values (and average *n*) from 500 such repetitions show that a Fisher test rejects the no-reversals null in (on average) one or two of our six within-subject scenarios (compared with one of the five between-subjects scenarios). In other words, when using a comparable low-power

[21] We use the Fisher exact test as a non-parametric unpaired equality-of-proportions test. Conceptually, it works as follows. Remember that each respondent answers either the choice or the SWB question. These responses are pooled and reassigned (possibly falsely) into the two groups, and the proportion of those who rank Option 1 above Option 2 in each group is recorded. This procedure is repeated for all possible permutations of assignments. The reported p-value is the fraction of repetitions that resulted in difference in proportions at least as great as that between the two actual groups. Computationally, p-values are calculated using an exact formula.

test, overall statistical results seem comparable.

In order to compare the *magnitudes* of the choice-SWB reversals in the between- and within-subject data, we focus on only the three scenarios that are identical across the two studies[22] and on only those of the Denver within-subject respondents who answer the life satisfaction question (*i*) (reported in Web Appendix table A4). In this comparison, we find that while in Scenario 1 (income vs. sleep) discrepancies increase from 3 percentage points in the within- to 10 in the between-subjects data, in Scenario 13 (friends vs. income) they decrease from 7 to virtually nil, and in Scenario 3 (absolute income vs. relative income) they are essentially nil in both studies. Our overall interpretation of our Denver data with the life satisfaction measure—the measure with the fewest choice-SWB reversals—is that while there are clear differences across the within- and between-subjects studies, the evidence in either supports the same two main findings, albeit with differences in relative emphasis.

Finally, our Cornell between-subjects questionnaires ask the "own happiness with life as a whole" SWB question variant (*iv*). Starting again with the first scenario, results are strikingly similar to those in the Cornell within-subject data. On magnitudes, 66 percent favor Option 2 (income) in the choice question while 46 percent favor Option 2 in the SWB question, similar to the corresponding proportions—68 and 45 percent—among the Cornell within-subject respondents who are asked SWB question variant (*iv*) (see Web Appendix table A4).[23] On statistical results, in spite of the reduced power of the unpaired Fisher test in the bottom row compared with the paired Liddell test directly above it, equality of proportions between the choice and SWB questions is still easily rejected in the between-subjects study.

Looking at the rest of the scenarios (in both tables) suggests that the finding in Scenario 1—that the choice-SWB difference in proportions in the between-subjects data is similar to but slightly smaller than that in the within-subject data—is typical of only some scenarios. Of the ten scenarios in the Cornell studies, the difference maintains its sign but increases in one scenario (from 7 to 12 percent); decreases in four (from 23 to 20; 10 to 4; 11 to 9; and 11 to 8 percent); disappears in two (10 to 0; and 12 to 0 percent); and changes sign in three, including two very

[22] As reported in table 1 and the Appendix, only scenarios 1, 3, and 13 are comparable across the studies: scenario 2 is only included in the between-subjects study, and different versions of scenarios 4 and 12 are presented across the within- and between-subjects studies. For further explanation see the Web Appendix.

[23] Remember that in both studies, these figures exclude "no difference" responses in the SWB question. We note that such "no difference" responses seem more prevalent in the within- than in the between-subjects data. See Web Appendix table A4, where we report their prevalence by design variations.

large changes (1 to -5; 20 to -13; and 24 to -11 percent). The Fisher test at the bottom row rejects the null hypothesis in three of the ten scenarios in the Cornell between-subjects data—strikingly similar to the Fisher test reported in the "Average *p*-value from 500 repetitions" row for the within-subject data.

As with the Denver data, our overall interpretation is that while there are differences across the between- and the within-subject studies—in particular, choice-SWB reversals are on average less pronounced in the between-subjects study—either set of studies supports our two main findings.

II.C. Measurement Error

Our analysis above suggests that in many scenarios, individuals do not respond to the choice and SWB questions as if they were responding to the same question. However, in a given scenario, such rejection of the null hypothesis could be explained by differences in measurement error across the two questions—for example, because it is easier to introspect about choice than about SWB, or vice versa. An individual whose "true" ranking of the options is identical across the questions is more likely to mistakenly rank the "wrong" option higher in a question with greater measurement error, leading to ranking proportions closer to 50-50 for that question.

Looking across table 2's columns reveals that cross-question differences in the measurement error variances for choice and SWB in the same direction in all scenarios in a study cannot explain our data. For example, in the Denver within-subject data, choice proportions are closer to 50-50 in Scenarios 1, 11, and 13, but SWB proportions are closer in Scenarios 4 and 12. We postpone further discussion of measurement error to the end of section III.

To summarize, the two main findings in this section are (a) that most respondents in most scenarios do not exhibit choice- versus SWB-ranking reversals, and (b) that when they do, their pattern of reversals is systematic. Overall, the two findings hold up well—although with differences in relative strength—across scenarios, populations, and designs. Furthermore, these findings cannot be explained by a measurement error structure that is stable across scenarios.

III. Do Non-SWB Factors Help Predict Choice?

In this section we take a second approach to assessing whether individuals seek to

maximize SWB. We ask: *Can we identify non-SWB factors that help explain choices, controlling for predicted SWB?* We also analyze *by how much* respondents' choices can be explained by their predicted SWB and non-SWB aspects of life together, compared with their predicted SWB alone.

We address these questions using data from the Cornell sample, where we ask respondents to rank the options on a set of eleven non-SWB aspects of life, in addition to ranking them on choice and SWB (see section I.C). Specifically, in addition to being asked about "your own happiness," respondents are also asked about: your family's happiness, your health, your romantic life, your social life, your control over your life, your life's level of spirituality, your life's level of fun, your social status, your life's non-boringness, your physical comfort, and your sense of purpose. While still a limited list, we chose these aspects of life in an attempt to capture what else, besides one's own happiness, might enter one's preferences and hence matter for choice. Our list is guided by economists' and philosophers' enumeration of "capabilities" (Sen, 1985; Nussbaum, 2000), non-hedonic components of SWB proposed by psychologists (White and Dolan, 2009), and our own introspection (for example, we tried to include all aspects that money can buy, such as physical comfort).

As mentioned above, in our Cornell between-subjects surveys we use a design that allows us also to elicit within-subject data from our 201 participants. This is done by presenting subjects with the between-subjects part of the survey, followed by an additional, within-subject part.[24] In the between-subjects analysis in section II we used only data from the first, between-subjects part. In contrast, in this section we pool data from both parts, treating them as within-subject data. Further pooling these data with the original Cornell within-subject data (432 respondents) yields an augmented sample of 633 Cornell within-subject respondents. In this section we analyze this pooled data set. In section IV.E and in the Web Appendix we show that our main results hold in the constituent subsamples.

[24] Specifically, we present the entire sequence of ten scenarios three times. First, each scenario is presented and is followed by only a choice question (for half the respondents) or only a SWB question (for the other half). Second, after respondents finish answering their question for each of the ten scenarios, the ten scenarios are presented again, each followed by only the question (SWB or choice) respondents have not seen yet. Finally, the ten scenarios are presented for a third time, with each scenario followed by the non-SWB aspect questions. Respondents are specifically instructed to answer the surveys in exactly the order questions are presented, and the experimenters verify that they do (in the rare cases where a respondent was observed to flip through the pages, she/he was promptly reminded of this instruction). With this design, excluding data collected after the first round of scenario-presentation results in between-subjects data.

III.A. Response distributions

Figure 1 displays, by scenario, the histograms of raw, multi-point responses to the choice, SWB, and eleven non-SWB-aspect questions. Note first that the choice responses—and also the SWB responses, although to a lesser extent—tend to be bimodal with most of the mass on "definitely" or "probably," suggesting that the choice-SWB reversals discussed in section II are *not* the result of widespread near-indifferences. Second, notice that we were rather successful in constructing Scenario 5 (apple versus orange): almost everyone indicates "no difference" on non-SWB aspects of the options. While 37 percent also indicate "no difference" on SWB, the low count of reversals in Scenario 5 suggests that for the other respondents, variation in choice is strongly related to variation in SWB. Finally, note that in many other scenarios, there is substantial variation in the non-SWB aspect rankings, and that the histogram of choice responses sometimes looks rather different from the histogram of SWB responses.

III.B. Predicting choice

Table 3 presents a variety of specifications in which we regress choice on SWB and non-SWB aspects of life, aggregating data across the ten scenarios (we discuss regressions by scenario in section V.B below). We want to estimate the relationship from the within-scenario—rather than the between-scenario—variation in responses. For this purpose, in the probit and ordered probit specifications, we include scenario fixed effects.[25] In the OLS specifications, we demean all variables at the scenario level. Doing so yields identical coefficients to a fixed-effects OLS specification but has the advantage that the R^2's reflect only the within-scenario explanatory power of the regressors. Interpreting the first three columns in table 3 as an attempt to predict choice responses in our data, we first focus only on R^2's. We return to interpreting coefficients in the next subsection.

The first column of table 3 reports an OLS regression of six-point choice on seven-point SWB. As reported by the R^2, 0.38 of the variation in choice is explained by SWB (own happiness) alone. In comparison, a regression of the same choice measure on our eleven non-SWB aspects (each as a seven-point variable) yields an R^2 of 0.21 (second column of table 3).

[25] Including fixed effects in non-linear models is problematic in cases where the number of fixed effects is asymptotically infinite. Since our asymptotics rely on an infinite number of respondents, not of scenarios, the inclusion of scenario fixed effects does not harm the consistency of our estimates.

Hence, we find that SWB predicts choice substantially better than all of the other aspects combined. In the third column we regress choice on both SWB and the eleven non-SWB aspects. The R^2 of 0.41 is again substantially higher than that in the second column but is only slightly higher than that in the first column.[26] The pattern in these three columns is similar when we relax the linear functional form, replacing each regressor with a set of six dummy variables (not reported): the respective R^2's are 0.46, 0.33, and 0.50. In summary, when we pool data across scenarios we find that adding the non-SWB aspects to the regression of choice on SWB increases explanatory power, but the increase is quite modest. (However, the increase is substantial in some of the individual scenarios, as we report in section V.B.)

III.C. Comparing the coefficients

In order to interpret—and compare—the regression coefficients in table 3, we need to provide more structure. We hence assume that choices result from maximizing a utility function,

$$U(H(\boldsymbol{X}), \boldsymbol{X}), \tag{1}$$

where H is SWB and $\boldsymbol{X}$ is a vector of non-SWB factors that might affect utility both directly, and indirectly through H. If people seek to maximize SWB alone (as opposed to trading off SWB for other, non-SWB factors), then the (vector) partial derivative $\partial U/\partial \boldsymbol{X}$ will be identically zero. To a first-order approximation, the difference in utility between two options in a decision problem is

$$\Delta U \approx (\partial U/\partial H)\, \Delta H + (\partial U/\partial \boldsymbol{X})\, \Delta \boldsymbol{X}, \tag{2}$$

where ΔH and $\Delta \boldsymbol{X}$ are the differences between the options in predicted SWB and predicted non-SWB aspects, respectively. Naturally, the smaller are these differences, the more justified is the approximation—a point we return to shortly.

We can now interpret the OLS regression from the third column of table 3 as estimating the implied regression equation,

[26] Bootstrapped standard errors yield the following 95-percent confidence intervals around the three respective R^2's: [0.36, 0.40], [0.19, 0.23], and [0.39, 0.43].

(3) $$\Delta U_{is} = \beta_H \Delta H_{is} + \boldsymbol{\beta}_X \Delta \boldsymbol{X}_{is} + \varepsilon_{is},$$

where ΔU_{is}, ΔH_{is}, and $\Delta \boldsymbol{X}_{is}$ are, respectively, respondent i's choice, SWB, and non-SWB responses in Scenario s (raw responses demeaned at the scenario level); and ε_{is} is an error term, which captures the direct effects of unmeasured differences between the options as well as choice mistakes. Under this interpretation, the twelve regression coefficients estimate marginal utilities.

Under this approximation, the null hypothesis $\partial U/\partial \boldsymbol{X} \equiv \mathbf{0}$ can be rewritten as the hypothesis that the vector $\boldsymbol{\beta}_X = \mathbf{0}$, or that all eleven non-SWB coefficients in the third column are zero—a hypothesis we can easily reject (F-test $p = 0.000$). This result is robust to treating the choice measure as ordinal rather than cardinal, or as binary (the ordered probit and probit specifications in the fifth and sixth columns); to relaxing the linearity of our SWB measure by replacing it with a set of six dummy variables; and to combinations of these specifications (see tables A7-A9 in the Web Appendix for these and other specifications). Furthermore, with the exception of Scenario 8 (where F-test $p = 0.086$), the result holds in each individual scenario.[27] This and the robustness to relaxing the linearity of ΔH suggest that not all the marginal utilities $\partial U/\partial \boldsymbol{X}$ are zero, regardless of how good an approximation we use in equation (2).

To move from testing the null hypothesis to interpreting the magnitudes of coefficients requires additional assumptions. Econometrically, the coefficients in regression equation (3) are consistently estimated if $E(\varepsilon_{is} \mid \Delta H_{is}, \Delta \boldsymbol{X}_{is}) = 0$. We measure a long list of non-SWB aspects as an attempt to make this assumption more plausible; an unmeasured factor will bias the coefficient on the regressors it is correlated with. Psychologically, the coefficients are comparable if respondents respond to the seven-point scales similarly across the twelve aspects.

Comparing the coefficients in the third column of table 3, the coefficient on SWB is by far the largest. A one-point increase in our seven-point measure of predicted SWB is associated with a highly significant 0.46-point increase in our six-point choice measure. The largest non-SWB coefficients are on sense of purpose (0.12), control over one's life (0.08), family happiness (0.08), and social status (0.06). The relative sizes of the coefficients are similar in alternative

[27] We show in table A10 in the Web Appendix that this result holds even when the regressions include only aspects for which more than a trivial fraction of respondents (e.g. 15 percent) indicate answers other than "no difference." In other words, it holds even when we include only the most reliably-estimated coefficients. Interestingly, table A10 shows that the only large and robust non-SWB coefficient in the "apple versus orange" scenario is that on "physical comfort"; this seems consistent with the *de gustibus* (or taste-related) interpretation of this scenario.

specifications (e.g., the ordered probit column), but we remind the reader that the data are pooled across surveys that use two opposite orders in which aspects are presented, and order matters for the coefficient estimates (see section IV.E). While the rejection of $\boldsymbol{\beta}_X = \mathbf{0}$ suggests that SWB is not the only argument in the utility function, a comparison of the coefficients suggests that the marginal utility of SWB is several times larger than the marginal utilities of even the most significant non-SWB aspects we measure.[28]

Finally, related to the question to what extent people seek to maximize SWB, we can ask to what extent people prefer a higher level of some non-SWB aspect *because* it increases SWB. Rewriting equation (2) as $\Delta U \approx [(\partial U/\partial H)\,(\partial H/\partial \boldsymbol{X}) + (\partial U/\partial \boldsymbol{X})]\,\Delta \boldsymbol{X}$ decomposes the net effect on utility of a change in $\boldsymbol{X}$ into an indirect effect that operates through ΔH and a direct effect. The coefficients in the second column of table 3 estimate the sum of the indirect and direct marginal effects, $(\partial U/\partial H)\,(\partial H/\partial \boldsymbol{X}) + (\partial U/\partial \boldsymbol{X})$. In the third column, the coefficient on SWB estimates $\partial U/\partial H$, while the coefficients on the non-SWB aspects estimate the direct marginal effects, $\partial U/\partial \boldsymbol{X}$. Under this causal interpretation of the coefficients, comparing the second and third columns suggests that, for example, while the effects of health and life's non-boringness on utility operate only through their effects on SWB, the entire effect of social status operates directly (without affecting SWB). In between these extremes, aspects such as family happiness, control over one's life, life's level of fun, physical comfort, and sense of purpose seem to have roughly equal direct and indirect effects on utility.

III.D. Measurement error

Measurement error in ΔH_{is} or $\Delta \boldsymbol{X}_{is}$ will cause the assumption $E(\varepsilon_{is} \mid \Delta H_{is}, \Delta \boldsymbol{X}_{is}) = 0$ to fail, therefore biasing the coefficient estimates. Moreover, measurement error in ΔH_{is} could also invalidate our tests of the null hypothesis that $\boldsymbol{\beta}_X \equiv \mathbf{0}$; even if predicted non-SWB aspects do not enter utility directly, if they affect predicted SWB, and if predicted SWB is measured with error, we could find—as we do—that the predicted non-SWB aspects are statistically significant predictors of choice. In order to address these concerns, we collected repeated observations on a sub-sample (of 230) of our Cornell respondents. This enables us to estimate a measurement-error-corrected equation (3). Specifically, we use Simulation-Extrapolation (SIMEX) (Cook and

[28] However, we believe that the most plausible bias from unmeasured factors exaggerates the coefficient on SWB. Specifically, an unmeasured factor whose effect on ΔH has the same sign as its direct effect (i.e., not through ΔH) on ΔU will bias upward the coefficient on ΔH.

Stefanski, 1994), a semi-parametric method that assumes homoskedastic additive measurement error but does not make assumptions about the distribution of the regressors.[29] Intuitively, the SIMEX method proceeds in two steps. First, it simulates datasets with additional measurement error and uses them to estimate the function describing how the regression coefficients change with the amount of measurement error. Then the algorithm extrapolates in order to estimate what the coefficients would be if there were no measurement error in the original data. Table 3 displays the results of applying this method to our regressions. As expected, relative to the OLS results, the coefficient on own happiness increases, and remains by far the most predictive regressor. However, the non-SWB aspects with largest coefficients and statistical significance in the OLS regressions remain statistically significant and also increase.

These results suggest that SWB is not the only argument in the utility function and that non-SWB aspects of life enter preferences as well. At the same time, among the aspects we measure, SWB has by far the highest marginal utility.

IV. Robustness

To examine the robustness of our results from sections II and III, we conduct a long list of additional analyses. Full details, including all tables, charts, and statistics, are available in the Web Appendix. In this section we briefly summarize our findings. Unless stated otherwise, they are based on our within-subject data from either the Denver or the (augmented) Cornell samples.

IV.A. Are results driven by only a few individuals?

A natural interpretation of our evidence from sections II and III is that most people's preferences assign high but not exclusive weight to SWB, and therefore people usually but not always choose the option they predict would maximize SWB. However, an alternative interpretation is that most people exclusively seek to maximize SWB, and our results are driven by a small minority of participants who do not. In contrast with this latter interpretation, we find

[29] We choose to use the SIMEX method over several more common measurement error correction methods (such as IV or regression disattenuation) for several reasons. Primarily, the other methods are highly inefficient. Moreover, the SIMEX method is flexible in its treatment of the measurement error structure, it accommodates misclassified categorical data, and it easily accommodates non-linear models such as probit or ordered probit regressions. For additional discussion of SIMEX see the Web appendix, and for IV results see table A12 there. For example, in our IV estimates, standard errors are so large that we cannot reject that any of the OLS or SIMEX coefficients reported in table 3 are equal to the IV coefficients, with one exception: the 0.46 coefficient on own happiness in the OLS regression (third column) is significantly lower than the corresponding IV coefficient, which is 0.75.

that most respondents (both Denver and Cornell) exhibit at least one reversal and that very few exhibit reversals in half or more of the scenarios.

Moreover, we explore whether some of the respondents who do not exhibit a choice-SWB reversal in a given scenario *would have* done so if that scenario's tradeoff between SWB and non-SWB factors had assigned a different "price" to SWB. We do this by presenting some of our Denver respondents with four different versions of Scenario 4, a choice between a career as an artist who earns $40,000 annually and will have a lasting legacy versus a higher-paying career as a commercial artist who will leave no legacy. The versions differ from each other in the commercial artist's annual income: it is $42,000, $60,000, $80,000, or $100,000. We find that the fraction of participants who exhibit a choice-SWB reversal varies from as low as 3 percent in the $42k version to as high as 17 percent in the $80k version. Overall, 24 percent exhibit a reversal in at least one income-version, suggesting that the fraction of individuals for whom we observe a reversal in a given scenario is a lower bound on the total number who would exhibit a reversal in the same scenario with *some* "price of SWB."

IV.B. Scenario-order effects and participant fatigue

We investigate the effects of scenario order on responses with our Denver sample, where respondents face the six scenarios in one of two opposite orders (see table 1). Scenario-order effects could arise, for example, due to increasing fatigue or boredom among respondents. They could either increase or decrease the number of choice-SWB reversals through increasing response variance, decreasing respondents' effort, and in general changing artificial consistency or inconsistency patterns. While we indeed find evidence of scenario-order effects on responses in individual scenarios, they do not seem to influence our main findings. Averaged across the six scenarios, 15 percent of respondents exhibit a choice-SWB reversal in a scenario, regardless of whether it appears in the first or second half of the survey (Fisher exact test $p = 0.87$).

IV.C. Self-reported artificial consistency/inconsistency and mistakes

In section II we address artificial consistency and inconsistency concerns, as well as measurement error that in all scenarios is either larger for choice than for SWB, or vice versa. In section III, our measurement-error-corrected regressions address the possibility of response mistakes that are uncorrelated across repetitions of the survey. As an additional check on these

and other concerns, after our Cornell respondents finish responding to all the decision scenarios, we directly ask them whether their choice-SWB reversals were a mistake; whether they think they would regret them; and whether they were trying to make their choice and SWB responses consistent. Interestingly, only 7 percent report their reversals as a mistake, 23 percent report that they would regret them, and 20 percent report consciously trying to make their choice and SWB consistent. We repeat section III's analysis excluding groups of respondents based on their responses. We find, in line with our findings from our earlier approaches, that while mistakes and artificial consistency/inconsistency are likely present in our data, they alone cannot explain our results.

IV.D. Self-control

A related concern is that choice-SWB reversals may reflect a self-control problem (as in Laibson, 1997 and O'Donoghue and Rabin, 1999), rather than a preference for non-SWB aspects of life.[30] To address this concern, in some versions of the survey, in addition to asking participants what they would choose, we also ask them what they would want themselves to choose (the meta-choice question mentioned in Section I.C(c)). We reasoned that if, for example, a participant preferred Option 1 but would choose Option 2 due to a self-control problem, the participant would indicate favoring Option 2 in the choice question but Option 1 in the meta-choice question. Aggregating across all surveys that include the meta-choice question (see table 1), we find reversals between choice and meta-choice in 28 percent of the cases. However, while self-control problems may be relevant in these cases, our main conclusions from section III appear to be robust to excluding these observations. For example, in our benchmark specification from the third column of table 3, while excluding observations with a choice/meta-choice inconsistency increases the coefficient on own happiness, it also increases the coefficient on sense of purpose.

[30] In most of our decision scenarios, it is not obvious how a self-control problem would be implicated, but there are a few where it could be. For example, in scenario 7, one of the alternatives is staying out later with friends, while the other is going to bed earlier to feel better and be more productive the next day. A respondent who correctly anticipates having a self-control problem might respond that she would choose to stay out late, even though her welfare would be maximized by going to bed earlier.

IV.E. Context of choice, SWB, and non-SWB aspect questions

Closely related to concerns of artificial consistency and inconsistency, it is possible that our results are affected by the organization of our choice, SWB, and non-SWB aspect questions. Specifically, respondents' interpretations of the questions or their understanding of the meaning of the related concepts may be context-dependent.[31] For example, when the choice and SWB questions appear close together (as in our example scenario in section I), respondents may interpret each of the questions differently from how they interpret each when presented separately. Additionally, interpretations of the questions may depend on whether the choice question appears before or after the SWB question in each scenario.

It is similarly possible that respondents' interpretation of "own happiness" depends on whether it appears as a stand-alone question or as merely one aspect of life in a list of twelve. For example, when "own happiness" appears as an aspect in a list, respondents may interpret it to refer to only the part of own happiness that is not correlated with—or affected by—the other aspects on the list ("happiness controlled for other aspects of life"). Finally, the order in which the twelve aspects are presented may affect responses, for example because order affects the aspects' interpretation, or because respondents pay less attention to aspects at the bottom of the list.

As mentioned in sections I (see table 1) and III, different versions of our surveys vary in whether the choice and SWB questions are asked close together or far apart, and in the order the questions are asked; they also vary in the distance between the SWB and non-SWB aspects, and in the order of aspects. Repeating our analysis in section III by questionnaire organization verifies that such order and context effects do indeed matter. We briefly summarize the main findings. First, aspects listed earlier have larger coefficients. Second, own happiness as part of a twelve-aspects list has a smaller coefficient than as a stand-alone question. Third, a design where we present each scenario to solicit only choice for that scenario, then present each scenario again to solicit only happiness, and then present each scenario yet a third time to solicit only non-SWB aspect ratings yields the lowest coefficient on own happiness (0.31). That design also yields, in one of the aspect orderings, the highest coefficient on sense of purpose (0.27), making it the highest coefficient on any of our non-SWB aspects in any of the designs. Yet, in all designs, non-

[31] Notice the important difference between this possibility and the possibility of cross-respondent differences in the interpretations or understanding of the *scenarios*. The latter possibility is a lesser concern as long as a respondent's interpretation or understanding of a scenario remains the same across the choice and SWB questions.

SWB aspects are statistically significant, and the coefficient on own happiness has a higher point estimate than any of the non-SWB aspects. While the context in which we ask our key survey questions definitely matters, our basic results appear to be robust.

IV.F. An alternative approach: willingness-to-pay questions

Our empirical methodology in section III was to assess the marginal utility of SWB and other factors by confronting participants with a series of scenarios, and estimating the coefficients from a regression of choice on predicted SWB and non-SWB aspect ratings of the choice options. As another source of evidence on marginal utilities, which we can use to cross-validate our regression-based findings, we also ask questions aimed to directly elicit participants' willingness-to-pay (WTP) for improving SWB and other aspects of their lives.

Near the end of the Cornell surveys, after participants had faced all of the scenarios, we ask them, for the list of twelve aspects, the maximum amount of minutes per week they would be willing to invest in improving their rating of their life on that aspect by one point on a ten-point scale. Consistent with our regression-based findings, own happiness has the largest WTP.[32] Moreover, the correlation between mean WTP and the coefficient from a univariate OLS regression of choice on each aspect individually is 0.67 ($p < 0.02$).[33] SWB has a large impact on this correlation coefficient because it has both a high WTP and a high regression coefficient. To assess the robustness of the correlation, we omit SWB from the set of aspects; in that case, the correlation is 0.49, economically sizeable but not statistically significant ($p = 0.12$).[34] We interpret these WTP-based findings as broadly supportive of our scenario-based findings.

[32] We measure WTP in units of time because we expect time to be more comparable across respondents than money for our student sample. Response categories are 0, 5, 10, 15, 20, 30, 45, 60, 75, and 120 minutes. Both mean WTP for own happiness (69 minutes) and its median (60 minutes) are well below the highest response option of 120 minutes (34 percent chose the latter). Across all of our survey populations, we similarly find that WTP for SWB is typically below the highest response option in a variety of questions including, for example, asking respondents what fraction of their income or what amount of meditation time they would be willing to devote to raise happiness. Although there are well-known reasons to be skeptical about contingent valuation questions, especially those far removed from everyday experience, this response pattern may be interpreted as suggesting that respondents think about a tradeoff between SWB and other aspects of their life when answering these questions.

[33] The univariate coefficient, as opposed to the coefficient from the multivariate regression reported in table 3, is the appropriate one to use for comparability with the WTP question because the WTP question does not partial out spillover effects on the other characteristics when eliciting WTP to improve one characteristic.

[34] These correlations might be inflated if aspects listed earlier have both larger coefficients (see section IV.E) and higher WTP; however, we find high correlations even between coefficients measured in one aspect order and WTPs measured in the opposite order.

V. Additional Results

To provide additional results that may be useful for happiness researchers and policy makers, we conduct the analysis from section III separately across SWB measures, scenarios, and respondent characteristics. This section is a brief summary of a more thorough treatment in the Web Appendix.

V.A. Comparing SWB measures

Across our surveys, we ask five different SWB question variants, based on three families of SWB questions asked in large-scale surveys used for empirical work: life satisfaction, happiness with life as a whole, and felt happiness (see section I.C(b) and table 1). In section II.A we compare the frequency of choice-SWB reversals across our different SWB measures. We now compare how well the measures predict choice using R^2's from univariate OLS regressions of our multiple-point choice variable on our multiple-point SWB measures.

As in section III, we demean our variables at the scenario level. In the Denver sample, the life satisfaction question variant (*i*) is the best predictor of the choice question, with $R^2 = 0.65$. Happiness with life as a whole (*ii*) and felt happiness (*iii*) come second and third, respectively, with $R^2 = 0.59$ and 0.55. The felt happiness R^2 is statistically significantly lower than the life satisfaction R^2 ($p = 0.02$ calculated using bootstrapped standard errors), and the R^2 for happiness with life as a whole is not statistically distinguishable from the other two. In the Cornell sample, own happiness with life as a whole (*iv*) and immediately felt own happiness (*v*) have $R^2 = 0.39$ and 0.37, not statistically distinguishable from each other.

These R^2's and our findings in II.A paint a consistent picture. While in the Denver data the life-satisfaction-type SWB question is more predictive of choice than the happiness-type SWB questions, in both Denver and Cornell the felt happiness and the happiness with life as a whole questions predict choice similarly. One possible hypothesis as to why some SWB measures predict choice better is that they encourage participants to report the present value of SWB flows over time. However, our finding that variant (*v*)—about happiness "in the few minutes immediately after making the choice"—is as predictive of choice as variant (*iv*)— about happiness in "life as a whole"—is inconsistent with this view.

V.B. Heterogeneity across decision scenarios

For applied work, it is useful to know in *which* situations the assumption that people try to maximize only SWB is a better or worse approximation. Table 4 shows the benchmark OLS specification from table 3, conducted separately for each of the ten scenarios in the Cornell data. The "Incremental R^2" row reports the difference between the reported R^2's (from the reported multivariate regressions) and R^2's from *univariate* regressions of choice on only SWB (which are not reported).

Mechanically, if utility is increasing in SWB, then a ranking of decision alternatives resulting from utility-maximization and a ranking resulting from SWB-maximization will coincide in choice settings where predicted SWB happens to be the only aspect that varies across the choice alternatives. As discussed above, Scenario 5 (apple versus orange)—which was designed to be as close as possible to such a case—has very little variance in non-SWB aspects and the fewest choice-SWB reversals (see figure 1 and table 2). As expected, the R^2 in a univariate regression of choice on predicted SWB is the highest (at 0.56) in Scenario 5, and the incremental R^2 from adding all other aspects is the lowest (at 0.02). If this type of minor decision—which possibly comprises most decisions in life—generally features low variance in non-SWB aspects, then the assumption that people try to maximize predicted SWB is a good approximation in such settings. The next highest univariate R^2 is for Scenario 3 (absolute income vs. relative income, at 0.49). This result is a surprise to us. We did not expect predicted SWB to capture the issues involved in the choice between absolute and relative income as well as it does.

Interestingly, at the other extreme, the four scenarios we designed to be representative of typical important decisions facing our college-age Cornell sample—Scenarios 7-10 (socialize versus sleep, family versus money, education versus social life, and interest versus career)—are among the scenarios with the lowest univariate R^2 and, correspondingly, the highest incremental R^2 from adding non-SWB aspects as regressors. Indeed, in Scenarios 7 and 10, where univariate R^2 is the lowest—at 0.25 and 0.24, respectively—incremental R^2 is 0.07 and 0.13. Here, non-SWB aspects increase predictive power (as measured by R^2) substantially, by 28 and 54 percent respectively. This in turn suggests that one should be especially cautious in interpreting measured SWB as utility in empirical applications that focus on important life decisions.

The rest of the scenarios fall somewhere in between. They include the scenarios that were designed to explore common themes from the happiness literature and, surprisingly, those designed as situations where we most expected to find tensions between SWB and non-SWB

factors (see section I.B).

Finally, table 4 reveals a heterogeneity in coefficients across scenarios (Chow test p = 0.000). Moreover, this heterogeneity appears to be systematic, with an increased coefficient on SWB and decreased coefficients on non-SWB factors in scenarios where there is less variance across the choice alternatives in non-SWB factors. In contrast, our simple theoretical framework from section III predicts constant marginal utilities and hence constant coefficients across scenarios. We speculate that this empirical heterogeneity results from the fact that in scenarios where some aspect varies less, that aspect becomes less salient as an input into utility, and respondents tend to take it into account less when making a choice.[35] Regardless of the explanation for the heterogeneity, it reinforces the assessment that predicted SWB may be a better approximation of utility in minor than in important decisions.

V.C. Preference heterogeneity across respondents

For a policymaker deciding whether to use SWB data as a proxy for utility for a particular demographic group, it may be useful to know how the propensity to maximize predicted SWB varies with demographic characteristics, such as gender, age, race, education, and income. In the Web Appendix, we report regressions that predict choice-SWB reversals as a function of respondent characteristics. In the Cornell sample, black respondents are 24 percentage points more likely than white ones to exhibit a choice-SWB reversal (with a standard error of 10 percentage points). However, on the whole across the Denver, CNSS, and Cornell samples, there is relatively little evidence for heterogeneity across demographic groups.

To see how heterogeneity in maximizing predicted SWB relates to well-understood dimensions of individual heterogeneity, we also predict the probability of a choice-SWB reversal

[35] In the Web Appendix we examine and rule out several alternative explanations. First, since our regression equation (3) relies on a first-order approximation of the utility function (1), one possible explanation of the above empirical heterogeneity is that the approximation breaks down in scenarios where the choice alternatives differ greatly in some aspects. Second, coefficients may be different in scenarios where both choice alternatives are far from the respondent's current situation, either because the relevant marginal utilities are not the same or because the respondent has difficulty predicting choice in fanciful hypothetical settings. We test these two explanations by running separate regressions for scenarios where we expect the local approximation to hold well or not; we find little evidence for differences in coefficients across these categories. Third, the coefficients may be identified off of different respondents in different scenarios. However, Cook's influence statistic (calculated at the respondent level) is highly correlated across most scenarios, and influence-robust iteratively reweighted least squares regressions still yield coefficients that vary across scenarios. This suggests that identification off of different respondents is not driving the variation in our coefficients across scenarios. We cannot rule out other explanations, such as respondents scaling their responses differently in different scenarios.

as a function of scores on the "Big 5" personality traits, which we measure in the Cornell within-subject sample using John and Srivastava's (1999) BFI scale. We find that a one standard deviation increase in conscientiousness is associated with an 2 percent lower likelihood of a choice-SWB reversal, while a one standard deviation increase in neuroticism (i.e., moody, tense) is associated with a 2 percent higher likelihood.

VI. Discussion

When aggregating our results across scenarios, we find that choices maximize predicted SWB (especially "life satisfaction") for most people most of the time. This may partially allay one source of concern about using SWB data as a welfare measure. However, care must be taken in interpreting our findings. The amount of choice-SWB concordance we find represents an upper bound on the degree to which the use of SWB measures is justified as a welfare measure in empirical applications. Applications always require additional assumptions that we do not test. For example, a typical assumption is that SWB measures are comparable and can be aggregated across individuals.[36]

When comparing scenarios, our results suggest that, first, researchers should be especially cautious about using SWB data as a proxy for utility in settings that are perceived to involve personally-important decisions. Second, in settings where one alternative involves higher income or more money, predicted SWB systematically understates the utility benefits of the money alternative. This in turn suggests that the increasingly-common practice of estimating willingness to pay for more sex, less pollution, etc., by comparing the coefficient on income with that on another variable in multivariate SWB regressions biases these estimates upwards.

Our scenario-based methodology could be usefully applied in several new directions. First, the method of assessing the correspondence between choice and predicted SWB could be used to assess new SWB measures. In the Web Appendix, we describe pilot data we collected on two such measures, neither of which appears to predict choice any better than existing measures. Measures that explicitly encourage integration of SWB over time, as well as ladder/mountain-

[36] Our results may also overstate the extent to which standard SWB questions provide a good measure of preferences because standard questions are asked absolutely ("How satisfied are with your life?"), while our SWB questions are asked relatively ("Between these two options, which do you think would make you more satisfied with life?"). Different individuals may apply different scales to a greater extent for an absolute measure, making it more difficult to translate an absolute SWB measure into a meaningful utility number than might be suggested by our results.

type measures, seem especially promising areas for further research.[37]

Second, our method could be used to provide more tailored guidance for applied work by asking about scenarios that are intended to address specific issues of interest. For illustrative purposes, we pilot four such scenarios at the end of our Cornell repeat-survey. For example, to reconcile the intuition that Americans today are better off than in the past with the finding that average SWB has remained flat in the U.S. over the past decades (Easterlin, 1974, 1995; see Stevenson and Wolfers, 2008, for a recent assessment), we ask respondents to rank being born in 1950 versus being born in 1990 in both choice and SWB questions. Although our 209 respondents overwhelmingly favor being born in 1990 in both questions (87 and 79 percent, respectively), the 9 percent who choose 1990 despite believing that they would be happier in 1950 is statistically significantly larger than the 1 percent exhibiting the reverse response pattern (Liddell test p-value < 0.001). This result indeed suggests that some people prefer being born later even if it does not make them happier. For another example, to reconcile the intuition that expanding political and economic freedoms for women have made women better off with the finding that average SWB among women has declined in the U.S. since the 1970s, both absolutely and relative to men (Stevenson and Wolfers, 2009), we ask respondents to rank living in a world with or without these expanded freedoms for women. Again, significantly more respondents choose a world with these expanded freedoms for women in spite of believing that a world without them would make them happier than the reverse (Liddell $p = 0.007$). For further examples and full details, see the Web Appendix.

Finally, some researchers have attempted to identify the key non-SWB aspects of life that are associated with greater welfare (e.g., Sen, 1985). Others have called for an SWB-based "national well-being index" to provide a measure of welfare that captures factors not represented in economic indicators such as GDP (e.g., Diener et al., 2009). To the best of our knowledge, our paper is the first attempt to empirically estimate the appropriate weights on SWB and other factors for combining them into an overall index of welfare. Our method could be applied more systematically for this purpose.

[37] Since different SWB questions seem to capture distinct dimensions of well-being that correlate differently with income and other variables (e.g. Kahneman and Deaton, 2010), future research could also explore whether a combination of SWB questions predicts choice better than any individual SWB question alone.

References

Becker, Gary S., and Luis Rayo. 2008. "Comment on 'Economic growth and subjective well-being: Reassessing the Easterlin Paradox' by Betsey Stevenson and Justin Wolfers." *Brookings Papers on Economic Activity*, Spring, 88-95.

Bentham, Jeremy. 1789. *An Introduction to the Principle of Morals and Legislations*.

Bernheim, Douglas B. 2009. "Behavioral Welfare Economics." *Journal of the European Economics Association*, 7(2-3), 267-319.

Blanchflower, David G. and Andrew J. Oswald. 2004. "Money, Sex and Happiness: An Empirical Study." *Scandinavian Journal of Economics*, 106(3), 393-415.

Bronsteen, John, Christopher J. Buccafusco, and Jonathan S. Masur. Forthcoming. "Well-being analysis." *Georgetown Law Journal*.

Cook, J. R. and Leonard A. Stefanski. 1994. "A simulation extrapolation method for parametric measurement error models." *Journal of the American Statistical Association*, 89, 1314-1328.

Diener, Ed, Richard Lucas, Ulrich Schimmack, and John Helliwell. 2009. *Well-Being for Public Policy*. New York: Oxford University Press.

DiTella, Rafael, Robert J. MacCulloch, and Andrew Oswald. 2003. "The Macroeconomics of Happiness." *Review of Economics and Statistics*, 85(4), 809-827.

Ellison, Christopher G. 1991. "Religious involvement and subjective well-being." *Journal of Health and Social Behavior*, 32, 80–99.

Easterlin, Richard A. 1974. "Does economic growth improve the human lot? Some empirical evidence." In Paul A. David and Melvin W. Reder (eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*. New York: Academic Press, Inc.

Easterlin, Richard A. 1995. "Will raising the incomes of all increase the happiness of all?" *Journal of Economic Behavior and Organization*, 27, 35-47.

Finkelstein, Amy, Erzo F.P. Luttmer, and Matthew J. Notowigdo. 2008. "What good is wealth without health? The effect of health on the marginal utility of consumption." *NBER W.P. 14089*, June.

Frey, Bruno S., Simon Luechinger, and Alois Stutzer. 2008. "The life satisfaction approach to valuing public goods: The case of terrorism." *Public Choice*, 138, 317-345.

Gilbert, Daniel T. 2006. *Stumbling on happiness*. New York: Knopf.

Gruber, Jonathan, and Sendhil Mullainathan. 2005. "Do cigarette taxes make smokers happier?" *Advances in Economic Analysis & Policy*, 5(1) Article 4.

Heffetz, Ori, and Robert H. Frank. Forthcoming. "Preferences for Status: Evidence and Economic Implications." *Handbook of Social Economics*, Jess Benhabib, Alberto Bisin, Matthew Jackson, eds., Elsevier.

John, Oliver P., and Sanjay Srivastava. 1999. "The Big Five trait taxonomy: History, measurement, and theoretical perspectives." In Lawrence A. Pervin and Oliver P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York: Guilford.

Laibson, David I. 1997. "Golden eggs and hyperbolic discounting." *Quarterly Journal of Economics*, 112(2), 443-477.

Layard, Richard. 2005. *Happiness*. Penguin Press.

Liddell, F. D. K. 1983. "Simplified exact analysis of case-referent studies: matched pairs; dichotomous exposure." *Journal of Epidemiology and Community Health*, 37(1), 82-84.

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

Kahneman, Daniel, and Angus Deaton. 2010. "High income improves evaluation of life but not emotional well-being." *Proceedings of the National Academy of Sciences*, 107(38) 16489-16493.

Kahneman, Daniel, Alan B. Krueger, David A. Schkade, Nobert Schwarz, and Arthur A. Stone. 2004. "A survey method for characterizing daily life experience: The Day Reconstruction Method." *Science*, 306, 1776-1780.

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. "Back to Bentham? Explorations of experienced utility." *Quarterly Journal of Economics*, 112(2), 375-405.

Kelly, William. 2004. "Sleep-Length and Life Satisfaction in a College Student Sample." *College Student Journal,* 38(3).

Kimball, Miles S., Helen Levy, Fumio Ohtake, and Yoshiro Tsutsui, 2006. "Unhappiness After Hurricane Katrina." University of Michigan mimeo.

Kimball, Miles S., Ryan Nunn, and Daniel Silverman, 2010. "Accounting for Adaptation in the Economics of Happiness." Univeristy of Michigan mimeo.

Kimball, Miles S., and Robert J. Willis. 2006. "Happiness and Utility." University of Michigan mimeo.

Koszegi, Botond, and Matthew Rabin. 2008. "Choices, Situations, and Happiness." *Journal of Public Economics*, 92, 1821-1832.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.

Nussbaum, Martha. 2000. *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.

O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing it now or later." *American Economic Review*, 89(1), 103-124.

Oreopoulos, Philip. 2007. "Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling." *Journal of Public Economics*, 91, 2213–2229.

Sen, Amartya. 1985. *Commodities and Capabilities*. Oxford: Oxford University Press.

Stevenson, Betsey and Justin Wolfers. 2008. "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox." *Brookings Papers on Economic Activity*, Spring, 1-87.

Stevenson, Betsey and Justin Wolfers. 2009. "The Paradox of Declining Female Happiness." *American Economic Journal: Economic Policy*, 1(2), 190-225.

Stutzer, Alois and Bruno S. Frey, 2008. "Stress that Doesn't Pay: The Commuting Paradox," *Scandinavian Journal of Economics,* 110(2), 339-366.

Stutzer, Alois. 2007. "Limited Self-Control, Obesity and the Loss of Happiness," IZA Discussion Papers 2925, Institute for the Study of Labor (IZA).

Truglia, Ricardo Nicolás Pérez. 2010. "A Samuelsonian Validation Test for Happiness Data," Harvard University mimeo, August 17.

Tsutsui, Yuichi, Miles Kimball, and Fumio Ohtake. 2010. "Koizumi Carried the Day: Did the Japanese Election Results Make People Happy and Unhappy?" *European Journal of Political Economy*, 26, 12-24.

Tversky, Amos, and Dale Griffin. 1991. "Endowments and Contrast in Judgments of Well-Being." In Richard J. Zeckhauser (ed.), *Strategy and Choice*, Cambridge, MA: MIT Press. Reprinted in Kahneman, Daniel, and Amos Tversky (eds.), *Choices, Values, and Frames*, Cambridge, UK: Cambridge University Press.

Welsch, Heinz. 2009. "Implications of happiness research for environmental economics." *Ecological Economics*, 68, 2735-2742.

White, Mathew P., and Paul Dolan. 2009. "Accounting for the richness of daily activities." *Psychological Science*, 20(8), 1000-1008.

Table 1: Study-specific Information

Study Location	**Denver**		**CNSS**	**Cornell**	
Choice vs. SWB: Within- or Between- Subjects	Within	Between	Within	Within	Between
Sample Population	Volunteers at a doctor's waiting room		Nationally representative	Cornell Students	
Observations	497	569	1000	432†	201
Scenarios used	1, 3, 4, 11, 12, 13	1, 2, 3, 4, 12, 13	1	1-10	
SWB Question Format	Observations for each SWB question variant in parenthesis				
(*i*) Life Satisfaction					
(Isolated)	(164)	(569)			
(*ii*) Happiness with Life as a Whole					
(Isolated)	(162)		(1000)		
(*iii*) Felt Happiness					
(Isolated)	(171)				
(*iv*) Own Happiness with Life as a Whole					
Isolated				(107)	(201)
First/Last In Series				(107)	
(*v*) Immediately Felt Own Happiness					
Isolated				(110)	
First/Last In Series				(108)	
SWB Response Scale	6-point		Binary	7-point	
Choice Response Scale	6-point		Binary	6-point	
Meta-Choice Question?	Yes	No	No	Yes	No
Order variations					
Scenario order	4-1-11-12-13-3 3-13-12-11-1-4	1-2-12-13-3-4 3-13-12-2-1-4‡	1	1-2- … -9-10	
Question order	Choice-Meta-SWB SWB-Choice-Meta		SWB-Choice	Choice-SWB	
Aspects of life order				Two opposite orderings of 11 non-SWB aspects	
Summary: number of questionnaire versions	12	4	1	8	4

Notes: See section I for details and for the framing of the choice, SWB, and meta-choice questions; see the Appendix for the list of scenarios (1-13).
† Of these, 230 were surveyed twice, allowing us to conduct measurement-error-corrected estimation.
‡ Scenario 4 is always presented last because it is followed by both a choice and a SWB question. In order to have a clean between-subjects design, we did not want subjects to know we were interested in both choice and SWB until after subjects were done with the rest of the scenarios. We also note that this scenario is presented in four different order-versions, so strictly speaking the Denver between-subjects study includes the four questionnaire versions reported in the table's bottom row, times four (sixteen versions in total). We report the number four because in terms of the between-subjects part of this study, there are four versions.

Table 2: Choice and SWB Responses Across Studies and Scenarios *(continued on next page)*

		Denver Study							CNSS
	Choice Scenario	1	2	3	4	11	12	13	1
	For exact phrasing see Appendix	Sleep vs Income	Concert vs Birthday	Abs. Inc. vs Rel. Inc.	Legacy vs Income	Concert vs Duty	Low Rent vs Short Commute	Friends vs Income	Sleep vs Income
Within	Higher SWB: Option 1 Chosen: Option 1	58%		48%	24%	16%	52%	50%	74%
	Higher SWB: Option 2 Chosen: Option 2	29%		42%	60%	65%	32%	34%	18%
	Higher SWB: Option 2 Chosen: Option 1	1%		6%	2%	12%	11%	2%	1%
	Higher SWB: Option 1 Chosen: Option 2	12%		4%	14%	7%	5%	14%	7%
	p-value from Liddell Exact Test	0.000		0.350	0.000	0.024	0.002	0.000	0.000
		n = 425		n = 420	n = 422	n = 422	n = 425	n = 422	n = 972
	Average p-value from 500 repetitions of Fisher Exact Test treating within-subject data *as if* they were between-subjects	0.111		0.527	0.044	0.359	0.315	0.071	
		n = 427		n = 423	n = 427	n = 423	n =427	n = 424	
					Version 2		Version 2		
Between	Higher SWB: Option 2	34%	86%	51%	***		54%	54%	
	Chosen: Option 2	44%	84%	48%	***		55%	53%	
	p-value of Fisher Test	0.020	0.716	0.433	***		0.930	0.793	
		n = 525	n = 524	n = 525			n = 526	n = 525	

Notes: Response distribution by study and scenario. For the complete text of each scenario, see the Appendix. If a scenario's phrasing changed meaningfully between surveys, the version of the scenario is indicated in the first row of the study block. For between-subjects data, we report the Fisher Exact Test *p*-value testing the null hypothesis that mean response to choice question = mean response to SWB question (an unpaired equality-of-proportions test). For within-subject data, we report the analogous Liddell Exact Test *p*-value (a paired equality-of-proportions test). In cases where respondents could indicate SWB indifference, responses indicating indifference were dropped from these tests. To conduct hypothesis tests with equal power for the within- and between-subjects data, we treat the within-subject data as if they were between-subjects by randomly assigning half of the observations to choice and half to SWB, and only looking at each respondent's assigned response. The reported *p*-value and *n* are the average across 500 repetitions of this algorithm. See details in section II. Since Scenario 4 ("legacy vs income") was not presented in a between-subjects design in spite of being included in the between-subjects part of the Denver study, its results from that part are not reported. For a description of this scenario's results, see section IV.A.

Table 2: Choice and SWB Responses Across Studies and Scenarios *(continued from previous page)*

Cornell Study

	Choice Scenario *For exact phrasing, see Appendix*	1 Sleep vs Income	2 Concert vs Birthday	3 Abs. Inc. vs Rel. Inc.	4 Legacy vs Income	5 Apple vs Orange	6 Money vs Time	7 Socialize vs Sleep	8 Family vs Money	9 Education vs Social life	10 Interest vs Career
		Version 2									
Within	Higher SWB: Option 1 Chosen: Option 1	29%	29%	41%	44%	45%	44%	62%	68%	53%	27%
	Higher SWB: Option 2 Chosen: Option 2	46%	49%	43%	31%	50%	37%	15%	15%	22%	35%
	Higher SWB: Option 2 Chosen: Option 1	1%	7%	14%	8%	2%	14%	17%	5%	22%	3%
	Higher SWB: Option 1 Chosen: Option 2	23%	15%	2%	17%	3%	5%	6%	12%	3%	35%
	Indifference for SWB	8%	14%	13%	10%	37%	22%	10%	5%	6%	6%
	p-value of Liddell Exact Test	0.000	0.002	0.000	0.001	0.424	0.000	0.000	0.001	0.000	0.000
		n = 397	n = 368	n = 375	n = 387	n = 270	n = 333	n = 385	n = 409	n = 402	n = 402
	Average p-value from 500 repetitions of Fisher Exact Test.	0.025	0.387	0.234	0.348	0.531	0.333	0.184	0.377	0.035	0.002
		n = 197	n = 196	n = 193	n = 196	n = 168	n = 194	n = 198	n = 198	n = 198	n = 198
		Version 2									
Between	Higher SWB: Option 2	46%	67%	49%	38%	56%	34%	26%	14%	16%	61%
	Chosen: Option 2	66%	71%	40%	38%	51%	34%	18%	26%	29%	50%
	p-value of Fisher Test	0.006	0.643	0.195	1.000	0.537	1.000	0.234	0.048	0.027	0.117
		n = 197	n = 196	n = 193	n = 196	n = 168	n = 194	n = 198	n = 198	n = 198	n = 198

Notes: Response distribution by study and scenario. For the complete text of each scenario, see the Appendix. If a questions phrasing changed meaningfully between surveys, the version of the question is indicated in the first row of the study block. For between-subjects data, we report the Fisher Exact Test *p*-value testing the null-hypothesis that mean response to choice question = mean response to SWB question (an unpaired equality-of-proportions test). For within-subject data, we report the analogous Liddell Exact Test *p*-value (a paired equality-of-proportions test). In cases where respondents could indicate SWB indifference, responses indicating indifference were dropped from these tests. To conduct hypothesis tests with equal power for the within- and between-subjects data, we treat the within-subject data as if they were between-subject by randomly assigning observations to choice or to SWB in numbers that match the between-subjects test, and only looking at each respondent's assigned response. The reported *p*-value and *n* are the average across 500 repetitions of this algorithm. See details in section II.

Table 3: Regressions of Choice on Aspects of Life

	OLS				Ordered Probit	Probit	
Measurement error correction	None	None	None	SIMEX Corrected Additive	None	None	SIMEX Corrected Additive
Own happiness	0.54***		0.46***	0.59***	0.37***	0.37***	0.48***
	(0.009)		(0.010)	(0.014)	(0.009)	(0.012)	(0.019)
Family happiness		0.15***	0.08***	0.11***	0.06***	0.09***	0.13***
		(0.017)	(0.015)	(0.026)	(0.012)	(0.017)	(0.032)
Health		0.07**	0.00	0.00	0.01	0.01	0.02
		(0.021)	(0.019)	(0.031)	(0.016)	(0.022)	(0.042)
Life's level of romance		-0.00	-0.01	0.01	-0.00	-0.00	0.04
		(0.024)	(0.021)	(0.033)	(0.018)	(0.025)	(0.045)
Social life		-0.01	-0.03	-0.05	-0.02	-0.02	-0.04
		(0.020)	(0.018)	(0.028)	(0.015)	(0.021)	(0.036)
Control over your life		0.17***	0.08***	0.11***	0.06***	0.09***	0.13***
		(0.017)	(0.015)	(0.025)	(0.012)	(0.017)	(0.028)
Life's level of spirituality		-0.08**	-0.02	-0.04	-0.02	-0.04	-0.05
		(0.024)	(0.021)	(0.036)	(0.018)	(0.025)	(0.047)
Life's level of fun		0.13***	0.05*	0.03	0.04*	0.04*	0.03
		(0.021)	(0.018)	(0.031)	(0.015)	(0.021)	(0.036)
Social status		0.07***	0.06***	0.07**	0.05***	0.07***	0.10***
		(0.016)	(0.014)	(0.023)	(0.012)	(0.016)	(0.027)
Life's non-boringness		0.07***	-0.01	-0.01	0.00	0.00	0.01
		(0.020)	(0.017)	(0.030)	(0.014)	(0.020)	(0.037)
Physical comfort		0.09***	0.04**	0.03	0.04**	0.05**	0.04
		(0.017)	(0.014)	(0.023)	(0.012)	(0.017)	(0.030)
Sense of purpose		0.21***	0.12***	0.13***	0.10***	0.12***	0.14***
		(0.015)	(0.013)	(0.022)	(0.011)	(0.015)	(0.025)
Observations	6285	6220	6217	6217	6217	6217	6217
(pseudo) R^2	0.38	0.21	0.41		0.19	0.35	

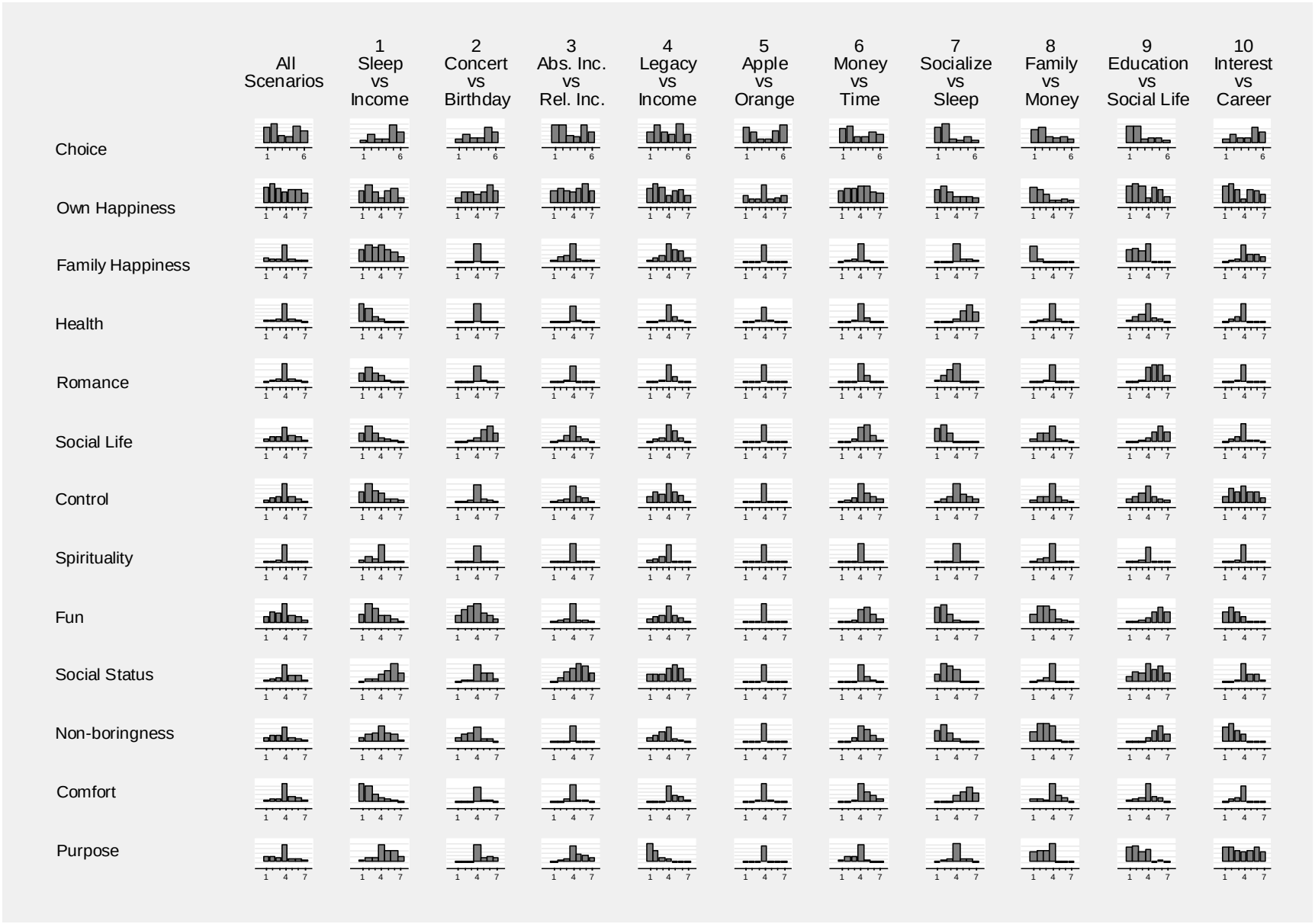
Notes: Standard errors in parentheses. In the OLS and ordered probit regressions, the dependent variable is 6-point choice. In the probit regressions the dependent variable is binary choice. All regressions use 7-point ratings of aspects. Based on 633 Cornell respondents. Each observation is a respondent's ratings for one scenario; there are 10 observations per respondent corresponding to the 10 scenarios in the questionnaires. Probit and ordered probit regressions include (unreported) scenario fixed effects. OLS regressions' variables are demeaned at the scenario level, generating coefficients equivalent to including scenario fixed effects. Measurement error corrections are done using the simulation extrapolation method described in section III, under the assumption of additive measurement error. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: OLS Regressions of Choice on All Aspects of Life, by Scenario

Choice Scenario *For exact phrasing, see Appendix*	All questions	1 Sleep vs Income	2 Concert vs Birthday	3 Abs. Inc. vs Rel. Inc.	4 Legacy vs Income	5 Apple vs Orange	6 Money vs Time	7 Socialize vs Sleep	8 Family vs Money	9 Education vs Social life	10 Interest vs Career
Own happiness	0.46***	0.38***	0.44***	0.52***	0.44***	0.73***	0.53***	0.31***	0.53***	0.35***	0.27***
	(0.010)	(0.031)	(0.031)	(0.032)	(0.031)	(0.036)	(0.036)	(0.032)	(0.033)	(0.029)	(0.030)
Family happiness	0.08***	0.07*	0.01	0.16***	0.05	0.16	0.15*	-0.09	0.05	0.14***	0.21***
	(0.015)	(0.032)	(0.071)	(0.046)	(0.041)	(0.159)	(0.059)	(0.053)	(0.050)	(0.037)	(0.041)
Health	0.00	-0.05	-0.07	-0.11	-0.04	0.05	0.06	0.18***	0.05	-0.03	-0.06
	(0.019)	(0.055)	(0.076)	(0.077)	(0.058)	(0.065)	(0.075)	(0.054)	(0.057)	(0.044)	(0.063)
Life's level of romance	-0.01	0.08	-0.02	0.07	-0.00	-0.67**	-0.10	0.02	-0.03	0.01	0.01
	(0.021)	(0.059)	(0.064)	(0.078)	(0.066)	(0.228)	(0.086)	(0.054)	(0.068)	(0.053)	(0.072)
Social life	-0.03	-0.02	0.02	-0.01	0.00	0.02	0.04	-0.00	-0.05	-0.04	0.01
	(0.018)	(0.055)	(0.043)	(0.056)	(0.058)	(0.225)	(0.071)	(0.065)	(0.053)	(0.053)	(0.054)
Control over your life	0.08***	0.02	0.05	0.04	0.08*	-0.00	0.07	0.15***	0.05	0.06	0.07*
	(0.015)	(0.042)	(0.053)	(0.056)	(0.039)	(0.093)	(0.052)	(0.043)	(0.049)	(0.038)	(0.035)
Life's level of spirituality	-0.02	-0.04	-0.00	-0.16	0.13*	0.31	-0.15	-0.01	-0.15*	-0.00	-0.01
	(0.021)	(0.049)	(0.061)	(0.090)	(0.055)	(0.221)	(0.091)	(0.076)	(0.062)	(0.054)	(0.068)
Life's level of fun	0.05*	0.06	0.15**	0.04	0.05	-0.08	0.13	-0.03	0.03	0.06	-0.00
	(0.018)	(0.042)	(0.051)	(0.066)	(0.047)	(0.127)	(0.068)	(0.073)	(0.059)	(0.057)	(0.057)
Social status	0.06***	-0.00	0.04	0.05	0.04	-0.27	-0.01	0.06	0.11	0.06*	0.16***
	(0.014)	(0.036)	(0.045)	(0.040)	(0.036)	(0.227)	(0.061)	(0.059)	(0.060)	(0.029)	(0.043)
Life's non-boringness	-0.01	0.05	-0.03	0.22**	-0.01	0.09	-0.03	0.18**	-0.05	-0.02	0.05
	(0.017)	(0.037)	(0.054)	(0.078)	(0.047)	(0.121)	(0.060)	(0.062)	(0.061)	(0.055)	(0.055)
Physical comfort	0.04**	0.09*	0.00	-0.05	0.00	0.21**	-0.00	0.05	-0.10*	0.06	-0.02
	(0.014)	(0.036)	(0.060)	(0.054)	(0.042)	(0.066)	(0.049)	(0.048)	(0.041)	(0.040)	(0.049)
Sense of purpose	0.12***	0.17***	0.12**	0.12**	0.12**	0.29*	0.05	0.04	0.09*	0.17***	0.17***
	(0.013)	(0.038)	(0.047)	(0.044)	(0.041)	(0.119)	(0.050)	(0.044)	(0.046)	(0.037)	(0.029)
Observations	6217	615	621	620	624	624	619	622	625	626	621
R^2	0.41	0.46	0.43	0.53	0.41	0.58	0.42	0.32	0.38	0.43	0.37
Incremental R^2	0.03	0.06	0.03	0.04	0.04	0.02	0.02	0.07	0.02	0.08	0.13

Notes: Standard errors in parentheses. OLS regressions of 6-point choice on 7-point aspects of life. Based on 633 Cornell respondents. The leftmost column aggregates data across choice scenarios; each of the other columns corresponds to a specific scenario. Each observation is a respondent's ratings for one scenario; there are 10 observations per respondent corresponding to the 10 scenarios in the questionnaires. All variables are demeaned at the scenario level, generating coefficients equivalent to including scenario fixed effects. "Incremental R^2" is the difference in R^2 between the reported multivariate regression and a univariate regression of choice on happiness; it represents the increased percentage of variation in choice that can be explained by including the additional aspects. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1: Raw Response Distributions (Choice and Aspects of Life)



Notes: Based on 633 Cornell respondents. See details in text. The histograms show the distribution of 6-point responses to the choice question (top row) and 7-point responses to the aspect questions (bottom twelve rows). The leftmost column aggregates data across choice scenarios; each of the other columns corresponds to a specific scenario.

Appendix: Scenarios Presented in Surveys

Scenario 1: Sleep vs. Income

Say you have to decide between two new jobs. The jobs are exactly the same in almost every way, but have different work hours and pay different amounts.

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.

Scenario 2: Concert vs. Birthday

Suppose you promised a close friend that you would attend his or her 50th ["21st" in student samples] birthday dinner. However, at the last minute you find out that you have won front row seats to see your favorite musician, and the concert is at the same time as the dinner. This is the musician's last night in town. You face two options:

Option 1: Skip your friend's birthday dinner to attend the concert.

Option 2: Attend your friend's birthday dinner and miss the concert.

Scenario 3: Absolute Income vs. Relative Income

Suppose you are considering a new job, and have offers from two companies. Even though all aspects of the two jobs are identical, employees' salaries are different across the two companies due to arbitrary timing of when salary benchmarks happened to be set. Everyone in each company knows the other employees' salaries. You must choose one of the two companies, which means you must decide between the following two options:

Option 1: Your yearly income is $105,000, while on average others at your level earn $120,000.

Option 2: Your yearly income is $100,000, while on average others at your level earn $85,000.

Scenario 4: Legacy vs. Income

(*Phrasing in Denver within-subject study)***:** Suppose you are a skilled artist, and you have to decide between two career paths for your life.

Option 1: You devote yourself to your own style of painting. This would require a number of sacrifices, such as having less time for friends and family, and making less money. For example, you expect that selling your paintings will give you an income of $40,000 a year. If you choose this path, you don't expect that your work will be appreciated in your lifetime, but posthumously you will make an impact on the history of art, achieve fame, and be remembered in your work.

Option 2: You become a graphic designer at an advertising company. This would give you more money and more time with friends and family than Option 1. The company is offering you a salary of $60,000 a year, which will afford you a much more comfortable lifestyle, but you will have no impact and leave no legacy to be remembered.

(*Phrasing in Denver between-subject study and Cornell study***) :** Suppose you are a skilled artist, and you have to decide between two career paths for your life. There are two styles of painting that you consider to be your own style, and you enjoy both equally. Style 1 happens to be much less popular than Style 2 today, but you know it will be an important style in the future.

Option 1: You devote yourself to Style 1. You expect that selling your paintings will give you an income of $40,000 a year. If you choose this path, you don't expect that your work will be appreciated in your lifetime, but posthumously you will make an impact on the history of art, achieve fame, and be remembered in your work.

Option 2: You devote yourself to Style 2. You expect that selling your paintings will give you an income of $60,000 a year, but you will have no memorable impact. [In the Denver between-subject study, each subject saw this question three times, with different salaries in option 2. This

number took a value of either $42,000, $60,000, $80,000, or $100,000.]

Scenario 5: Apple vs. Orange

Suppose you are checking out a new supermarket that just opened near where you live. As you walk by the fresh fruit display, you are offered your choice of a free snack:

Option 1: A freshly sliced apple.

Option 2: A freshly sliced orange.

Scenario 6: Money vs. Time

Suppose that due to budget cuts, the school implements a "student activities fee" of $15 dollars a week to help pay for maintenance of facilities used for extracurricular student activities. However, the school allows you to not pay the fee if instead you put in 2 hours of service a week shelving books at the library. You face two options:

Option 1: Spend 2 hours a week shelving books.

Option 2: Pay $15 a week.

Scenario 7: Socialize vs. Sleep

Say you are hanging out with a group of friends at your friend's room. You are having a really good time, but it is getting to be late at night. You have to decide between two options.

Option 1: Stay up another hour. It is likely you will feel tired all day tomorrow, but this particular evening you are having an especially fun time.

Option 2: Excuse yourself from the group, and go to bed. You will be disappointed to miss the fun, but you know you will feel better the next day and be more productive at paying attention in class and doing your homework.

Scenario 8: Family vs. Money

Imagine that for the first time in three years, your parents (or if your parents are gone, your closest relatives who are older than you) have arranged for a special family gathering that will happen the day after Thanksgiving, with everyone also invited to Thanksgiving dinner. You face two options. Would you choose to go to the family gathering the day after Thanksgiving (and maybe to Thanksgiving dinner) if getting there required a $500 roundtrip plane ticket for plane flights that were 5 hours each way?

Option 1: Go to the thanksgiving gathering, which requires a $500 round trip plane ticket.

Option 2: Miss the thanksgiving gathering, but save the money.

Scenario 9: Education vs. Social Life

Suppose you have decided to leave Cornell, and are transferring to a new school. You have been accepted to two schools, and are deciding where to go. The first school is extremely selective and high quality, but is in a small town out in the country with a less active social scene. The second school is in a major city with a great social scene, but is slightly less renowned. Which would you choose?

Option 1: Highly selective school, isolated socially and geographically.

Option 2: Less selective school, socially active and in a major city.

Scenario 10: Interest vs. Career

Suppose you are considering two summer internships. One is extremely interesting and involves work you are passionate about, but does not advance your career. The other will likely be boring, but will help you get a job in the future. Which would you choose?

Option 1: Interesting internship which does not advance career.

Option 2: Boring internship which will help you get a job.

Scenario 11: Concert vs. Duty

Say you are driving by yourself to see your favorite musician in concert on their last day in town.

You are five minutes away, and the concert starts in ten minutes. On the drive, you witness a truck hit a parked car, causing roughly $500 in damages, and then drive away without leaving their information. You notice the truck's license plate, and you are the only witness. You face two options:

Option 1: Keep driving and get to the concert on time.

Option 2: Call the police, in which case you will have to wait around the parked car to give a testimony. This would take about half an hour. You would have trouble finding a seat and might miss the whole concert.

Scenario 12: Low Rent vs. Short Commute

(*Phrasing in Denver within-subject study):* Say you are moving to a new town. You are trying to decide between two similar apartments which you could rent. The two apartments are identical in almost everything – including floor plan, amenities, neighborhood character, schools, safety, etc. However, they have different rents and are located at different distances from your work.

Option 1: An apartment which requires a 45-minute drive to work. The rent is about 20% of your monthly income.

Option 2: A similar apartment, with only a 10-minute drive. The rent is about 40% of your monthly income.

(*Phrasing in Denver between-subject study):* Say you are moving to a new town. The new town is known for its terrible traffic jams, and driving there is widely considered to be unpleasant. You are trying to decide between two similar apartments which you could rent. The two apartments are identical in almost everything – including floor plan, amenities, neighborhood character, schools, safety, etc. However, they have different rents and are located at different distances from your work.

Option 1: An apartment which requires a 45-minute drive each way to work. The commute has

heavy traffic almost the whole way. The rent is about 20% of your monthly income.

Option 2: A similar apartment which requires a 10-minute drive each way to work. The commute has heavy traffic almost the whole way. The rent is about 40% of your monthly income.

Scenario 13: Friends vs Income

Say you have been reassigned at your job, and will be moved to a new location. There are two offices where you could request to work. One office is in a city where many of your friends happen to live, and pays 20% less than your current salary. The other office is in a city where you don't know anyone, and pays 10% more than your current salary. Your job will be exactly the same at either office. You must decide between the following two options:

Option 1: Make 20% less than your current salary and move to the city with your friends.

Option 2: Make 10% more than your current salary and move to a city where you do not know anyone.