BEYOND HAPPINESS AND SATISFACTION:
TOWARD WELL-BEING INDICES BASED ON STATED PREFERENCE

Daniel J. Benjamin
Ori Heffetz
Miles S. Kimball
Nichole Szembrot

Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference
Daniel J. Benjamin, Ori Heffetz, Miles S. Kimball, and Nichole Szembrot
NBER Working Paper No. 18374
September 2012
JEL No. A13,D69,E01,I31

ABSTRACT

This paper helps provide foundations for survey-based tracking of well-being. First, we propose a theory in which utility depends on "fundamental aspects" of well-being, measurable with surveys. Second, drawing from psychologists, philosophers, and economists, we compile a comprehensive list of such aspects. Third, to estimate the aspects' marginal utilities—a necessary input for constructing an individual-level well-being index—we conduct a survey in which ~4,600 U.S. respondents state their preference between pairs of aspect bundles. We estimate high relative marginal utilities not only for traditional happiness and life satisfaction measures, but even more for aspects related to family, health, security, values, and freedoms.

Daniel J. Benjamin
Economics Department
Cornell University
480 Uris Hall
Ithaca, NY 14853
and NBER
daniel.benjamin@gmail.com

Ori Heffetz
324 Sage Hall
Cornell University
Ithaca, NY 14853
oh33@cornell.edu

Miles S. Kimball
Department of Economics
University of Michigan
Ann Arbor, MI 48109-1220
and NBER
mkimball@umich.edu

Nichole Szembrot
Economics Department
Uris Hall, Cornell University
Ithaca, NY 14853
nes66@cornell.edu

An online appendix is available at:
http://www.nber.org/data-appendix/w18374

The cornerstone of neoclassical welfare economics is the principle of revealed preference, according to which the ultimate criterion for judging what makes a person better off is what she chooses, in a situation in which she is well-informed about the consequences of her options. Yet for most policy decisions, a government cannot directly infer an individual's welfare from her choices over policies because individuals rarely make such choices.[1] Hence in practice economists often rely on revealed preference indirectly, evaluating policy options by how they affect indicators—most prominently, GDP—that can be viewed as summarizing, under some assumptions, a set of generally-desired outcomes. But because GDP and other available indicators have known limitations as well-being measures, economists have been seeking additional indicators that go "beyond GDP" (for a recent survey, see Fleurbaey, 2009). In this paper, we focus on developing one such indicator: an individual-level index that combines together different aspects of well-being that may be measured by survey questions.

As candidate measures of individuals' well-being, economists and psychologists have recently been investigating survey measures of "subjective well-being" (SWB); while we use this term to refer to any subjective assessment of some aspect of well-being, economists have primarily focused on questions about own happiness or life satisfaction. Because responses to such questions reflect a wide range of experiences, including those unrelated to market exchange (Diener and Seligman, 2004; Kahneman and Krueger, 2006), many researchers have advocated conducting nation-wide SWB surveys and using the responses to calculate indicators alongside GDP-like measures (e.g., Diener, 2000; Diener, 2006, signed by 50 researchers; Layard, 2005; Stiglitz, Sen, and Fitoussi, 2009).

Although these proposals are controversial among economists, policymakers have begun to embrace them. For example, starting in April 2011, the U.K. Office of National Statistics began including the following SWB questions in its Integrated Household Survey, a survey that reaches 200,000 adults annually (ONS, 2011):

> *Overall, how satisfied are you with your life nowadays?*
> *Overall, how happy did you feel yesterday?*

---

[1] Holding a referendum on every issue would incur prohibitively high transaction costs. Moreover, for many issues, even a direct vote would not reveal preferences because voters lack full information about— and might systematically mispredict—the consequences of alternative policy options (see Gilbert, 2006, for evidence on systematic misprediction of happiness).

Overall, how anxious did you feel yesterday?

Overall, to what extent do you feel the things you do in your life are worthwhile?

According to Prime Minister David Cameron, "it's time we focused not just on GDP but on GWB—general wellbeing."[2] Other governments around the world have expressed similar intentions to field SWB surveys and use the responses to guide policy.[3]

Notwithstanding this recent enthusiasm, there are many open questions regarding the endeavor of tracking well-being with surveys. Among the most urgent practical questions is, first, which SWB questions should governments ask? It is increasingly recognized that more than one question is likely to be needed because SWB is multi-dimensional (e.g., Ryff, 1989; Kahneman and Deaton, 2010) and because factors not captured by widely-used SWB measures may matter for utility (Benjamin, Heffetz, Kimball, and Rees-Jones, 2012). Current proposals for survey questions, however, rely on experts' reading of the SWB literature rather than on a systematic method.[4] Second, how should responses to different questions be weighted relative to each other? Current proposals are virtually silent on relative weighting (in some cases purposefully so). But in practice, due to an apparently-inevitable demand for summary indicators, *ad hoc* weights often end up being applied implicitly by users or explicitly in published indices (Micklewright, 2001).

This paper has two overarching purposes. First, we propose a framework, grounded in a preference-based theory, for conceptualizing and discussing survey-based measurement of well-being. Second, we demonstrate a disciplined approach, anchored in revealed preference—albeit based on hypothetical choices—to applying our framework to the development of well-being surveys and indices. We emphasize that—relative to the many decades of theoretical and practical work that underlies the present well-developed state of measures such as GDP—efforts

---

[2] "David Cameron aims to make happiness the new GDP." *The Guardian*, November 14, 2010. http://www.guardian.co.uk/politics/2010/nov/14/david-cameron-wellbeing-inquiry as accessed on May 13, 2011.

[3] For example, in December 2011, the U.S. National Academy of Sciences's Committee on National Statistics convened the first in a series of meetings of a "Panel on Measuring Subjective Well-Being in a Policy Relevant Framework." As part of its Better Life Initiative, the OECD has held four conferences on "Measuring Well-Being for Development and Policy Making" since May, 2011, with more planned. Moreover, this interest is not limited to rich, Western countries; indeed, Bhutan is considered the pioneer of Gross National Happiness, a concept conceived there in the 1970s.

[4] For prominent—and conflicting—proposals to the U.K. Office of National Statistics, see Dolan, Layard, and Metcalfe (2011) and Deaton, Kahneman, Krueger, Schkade, Schwarz, and Stone (2011).

to construct and apply survey-based well-being indicators are still in their infancy. Hence, we view this paper primarily as proposing an agenda for a new approach, and we view our specific contributions as first steps to be improved upon by future work.

In Section I, we present our theoretical framework. We assume that utility, $u(\mathbf{w})$, depends on a vector $\mathbf{w}$ of fundamental aspects of well-being, for example those that can be measured with survey questions similar to the four above. The vector of marginal utilities, $D_{\mathbf{w}}u(\mathbf{w})$, can then be used as a vector of relative weights for combining the components of $\mathbf{w}$ into an individual-level index that tracks small changes in well-being. While we do not make novel contributions regarding how to aggregate well-being indices across individuals, our framework could be used in conjunction with existing approaches to aggregation.

In Section II, we describe our attempt to identify the major components of $\mathbf{w}$. We compile a list—of 136 aspects of well-being—that is aimed at including all the main factors that have been proposed as important components of well-being in a sample of major works in philosophy, psychology, and economics, from Maslow (1946) to Stiglitz, Sen, and Fitoussi (2009) and beyond. The list includes aspects that have been considered as fundamental, as well as broader, composite aspects that may single-handedly capture many fundamental aspects. While it includes SWB measures widely used by economists (e.g., happiness and life satisfaction), it also includes other items, such as goals and achievements, freedoms, engagement, morality, self-expression, relationships, and the well-being of others. While far from exhaustive, our list represents, as far as we know, the most comprehensive effort to date to construct such a compilation (cf., Alkire, 2002). We additionally include on the list, for comparison purposes, "objective" measures that are commonly used as indicators of well-being (e.g., GDP, unemployment, inflation).

Next, as described in Section III, we design and conduct a survey to estimate the vector of marginal utilities, $D_{\mathbf{w}}u(\mathbf{w})$. We present more than 4,600 Internet survey respondents— recruited by a private firm to match the U.S. population reasonably well on observable demographics—with sets of hypothetical-choice scenarios. In each scenario we elicit respondents' stated preference between two options that differ only on how they rate on a small set of aspects. For example:

4

Imagine you are **making a personal decision**, and that you face a choice between two options: Option 1 and Option 2. The two options are predicted to have different effects over the next four years but to have the same effects after that. The table below lists these predicted differences in the next four years. Please assume that anything not listed in the table would be marked "about equal" if it were listed.

Click here to see the instructions again

	OPTION 1				OPTION 2		
	much higher	somewhat higher	slightly higher	about equal	slightly higher	somewhat higher	much higher
how happy **you** feel						X	
you not feeling anxious			X				

Between these two options, which do you think you would choose?

OPTION 1			OPTION 2		
Much prefer Option 1	Somewhat prefer Option1	Slightly prefer Option 1	Slightly prefer Option 2	Somewhat prefer Option 2	Much prefer Option 2
◎	◎	◎	◎	◎	◎

In our marginal utility estimation procedures, the dependent variable is the response to the choice question above, and the independent variables are the relative ratings of the aspects (the "X"s) in the table above. Because we randomly assign which aspects vary between the options and by how much, we can identify the marginal utility of each aspect. We call this stated-preference survey for estimating marginal utilities our *SP survey* to distinguish it from a *SWB survey* that a government might conduct for measuring individuals' levels of aspects of well-being.[5]

Section IV presents our main survey findings. Using personal choice scenarios similar to the one above we find, among other things, that while commonly-measured aspects of well-being such as happiness, life satisfaction, and health are indeed among those with the largest marginal utilities, other aspects that are measured less commonly have marginal utilities that are at least as large. These include aspects related to family (well-being, happiness, and relationship quality),

---

[5] Broadly speaking, our SP survey can be viewed as an application of conjoint analysis (Green and Rao, 1971). In the context of assessing welfare, our survey design is closely related to the method proposed in Adler and Dolan (2008), who argue that policymakers' weighting of different aspects of well-being should be informed by how survey respondents rank alternative "lives" that vary in the aspects. They illustrate their method with an exploratory study of 72 undergraduate respondents and four aspects: income, health, happiness, and life expectancy. Relatedly, Adler (2012) proposes a conceptual framework similar to ours, viewing w as a hybrid of "mental" (e.g., emotions) and "non-mental" (e.g., freedoms) aspects. He encourages using stated-preference surveys, rather than SWB surveys, for learning about $u(w)$.

security (financial, physical, and with regard to life and the future in general), values (morality and meaning), and having options (freedom of choice, and resources). Using policy-choice scenarios, in which respondents vote between two policies that differ in how they affect aspects of well-being for *everyone* in the nation, we continue to find the patterns above and in addition find high marginal utilities for aspects related to political rights, morality of others, and compassion towards others, in particular the poor and others who struggle. While we find some differences across demographic-group and political-orientation subpopulations of our respondents, most of these main results hold across the subpopulations we examine.

We present a long list of robustness checks in section V. These include exploring the sensitivity of our findings to: estimating alternative econometric specifications; excluding alternative candidates for composite aspects from the estimation; examining subsets of our respondents based on the time they took to complete the survey; and varying survey design elements that we randomly manipulated.

In section VI we return to the two practical questions above—which items to include on a SWB survey, and how to weight them?—and discuss potential solutions, both in theory and in practice.[6] We also consider, as a pragmatic future direction for our approach, these two practical questions with the added constraint that a government can only ask a small number of survey questions.

Our paper contributes to a long line of research on social welfare measures, recently surveyed by Fleurbaey (2009).[7] Our approach has the appealing feature that it accommodates several traditions that are often considered conflicting. It is, at the same time, "super-liberal" (Fleurbaey, 2009) since we weight aspects based *solely* on our respondents' stated preferences; "welfarist" (Sen, 1979) since utility can be viewed as our exclusive criterion for judging well-being;[8] and an application of Sen's (1985, 1992) "capability" approach since the vector w can be

---

[6] Besides the immediate practical value of our framework for selecting questions and weights, the use of systematic procedures could help allay concerns, discussed by Frey and Stutzer (forthcoming), that political agents might otherwise manipulate national well-being surveys for their own benefit.

[7] Fleurbaey's (2009) title, "Beyond GDP: The Quest for a Measure of Social Welfare," inspired the title of our paper.

[8] Sen (1979) defines welfarism as the imposition of the following "informational constraint": "[t]he judgment of the relative goodness of alternative states of affairs must be based exclusively on, and taken as an increasing function of, the respective collections of individual utilities in these states." Sen criticizes this notion and argues that information beyond individual utilities may be needed for making moral

considered as including both functionings (achieved states, e.g., material standard of living, health) and capabilities (opportunities for such achievements, e.g., freedoms).[9] Indeed, as described above, we find that our respondents put a large weight on capabilities, especially in policy scenarios.[10]

The central assumption underlying our SP survey is that a person's stated preference in our abstract scenarios is an unbiased measure of her ("true") preference.[11] This assumption is surely wrong; indeed, there are known ways in which stated preference is biased relative to incentivized choice, for example when one choice option is viewed as more socially desirable (Camerer and Hogarth, 1999; Ding, Grewal, and Liechty, 2005). Nonetheless, we believe it is more attractive to rely on what people's own stated preferences suggest about what they themselves care about than to paternalistically rely on the opinions and introspections of "experts" (such as researchers and policymakers) regarding which aspects to track and how to weight them. Moreover, some of the objections to using stated preferences as if they are descriptive of incentivized choice may have less force when stated preferences are used normatively. In particular, while hypothetical choices in abstract scenarios may elicit meta-preferences (the preferences people *want* themselves to have), rationalized preferences (more deliberated, internally consistent preferences), or otherwise laundered preferences (e.g., omitting "dirty" preferences such as racism), it is sometimes argued that these are more relevant for evaluating welfare than the preferences that describe actual behavior. For example, our SP survey may put respondents in a deliberative frame of mind, causing them to weight emotional

judgments. To partially address this concern, our approach attempts to include important specific examples of such information as additional arguments in the utility function (i.e., as elements of w).

[9] From the capabilities-approach perspective, our empirical effort is an attempt to address the "index problem," namely the problem of choosing weights for functionings and capabilities. This problem has been considered a key obstacle to systematically applying the capabilities approach (Fleurbaey, 2009) and is in general central to the construction of multi-dimensional well-being indices (Decancq and Lugo, 2013). Although Sen initially opposed combining measures of capabilities and functionings into an index, he later seems to have become more sympathetic (Micklewright, 2001).

[10] This also gives some support to Rawls's (1971) contention that most people would prioritize basic rights ahead of other goods.

[11] While we interpret our results in light of this assumption, we believe that our empirical contribution can be useful also in the context of alternative theoretical perspectives, such as that of Fleurbaey, Schokkaert, and Decancq (2009), who single out a specific survey question—a life satisfaction question—and propose to use it as the primary source of information about preferences, and those explored by Decancq, Van Ootegem, and Verhofstadt (2011), who pre-select nine survey questions and evaluate alternative weighting schemes by directly asking survey respondents how important they consider each of the nine to be. In particular, our approach can inform researchers regarding the link between different survey questions and stated preferences.

factors less than they do in "real life"—but doing exactly that is common prescriptive advice for avoiding emotion-induced mistakes (e.g., Camerer, Issacharoff, Loewenstein, O'Donoghue, and Rabin, 2003, pp.1238-1240). We discuss related points when interpreting our results (in section IV).

In section VII, we mention some concerns regarding tracking well-being with surveys that we do not address. We conclude by discussing possible extensions of our approach, and we point to a few readily actionable steps suggested by our findings. Throughout, we highlight the limitations of our specific implementation and point out promising directions for further developing the agenda we propose.

I. Theoretical Framework

We start with the standard framework for aggregating an individual's consumption of different commodities. This framework underlies empirical expenditure- and income-based measures of well-being, including GDP; below we adapt it for conceiving a well-being index.

In the consumption context, an agent's well-being is represented by a continuously-differentiable utility function $u(\boldsymbol{c})$. The vector, $\boldsymbol{c} = (c_1, \ldots, c_M)'$, represents the quantities of M market goods. Following a change in the consumption vector, $\Delta\boldsymbol{c}$, the change in utility can be approximated up to an arbitrary multiplicative scale as

$$(1) \qquad \Delta u \approx \left(D_{\boldsymbol{c}} u(\boldsymbol{c})\right)' \cdot \Delta\boldsymbol{c} = \sum_{m=1}^{M} \frac{\partial u(\boldsymbol{c})}{\partial c_m} \Delta c_m \propto \sum_{m=1}^{M} p_m \, \Delta c_m.$$

The proportionality follows because, at the optimum, as long as the consumer chooses a strictly positive amount of each good c_m, each marginal utility $\frac{\partial u(\boldsymbol{c})}{\partial c_m}$ is equal to a Lagrange multiplier times the market price p_m. By fixing prices, $\overline{p_1}, \ldots, \overline{p_M}$, at their levels in some base period and measuring the agent's consumption vector over time, the government can track a quantity index of real consumption, $\sum_{m=1}^{M} \overline{p_m} \, c_m$. For small changes in $\boldsymbol{c}$, changes in this index are approximately proportional to changes in utility.

Perhaps the biggest limitation of this consumption-based approach is that it only considers a narrow set of determinants of well-being. To broaden its scope, we follow other researchers in shifting attention away from standard consumption goods (for example, rice, TVs,

train rides) and toward more fundamental aspects of well-being (for example, health, emotional states, freedoms). This is intended to be more general in that these fundamental aspects include all objects of desire for individuals regardless of which specific consumption goods are in an agent's choice set at a given time and place. In this framework, consumption matters for well-being through its effects on these more fundamental aspects of well-being, but non-consumption determinants of well-being are also accounted for via their effects on the fundamental aspects.[12]

Consider the utility function $u(\mathbf{w})$, where $\mathbf{w} = (w_1, \dots, w_J)'$ represents the quantities of J fundamental aspects. In analogy to the consumption formula above (equation 1), a change in utility resulting from a change in the fundamental aspects can be approximated as

$$(2) \qquad \Delta u \approx \left(D_{\mathbf{w}} u(\mathbf{w})\right)' \cdot \Delta \mathbf{w} = \sum_{j=1}^{J} \frac{\partial u(\mathbf{w})}{\partial w_j} \Delta w_j.$$

Instead of measuring the agent's consumption vector $\mathbf{c}$, the government would measure her fundamental-aspects vector $\mathbf{w}$; and instead of tracking a quantity index for standard consumption goods, the government would track $\sum_{j=1}^{J} \frac{\overline{\partial u(\mathbf{w})}}{\partial w_j} w_j$, with the marginal utilities fixed at a base period. We term $\sum_{j=1}^{J} \frac{\overline{\partial u(\mathbf{w})}}{\partial w_j} w_j$ the agent's *well-being index*. Since the marginal utilities are defined only up to an arbitrary multiplicative constant, so is the index. From the perspective of this theoretical framework, the purpose of a SWB survey is to measure the w_j's.

Because there are no observable prices that can be used in place of the marginal utilities, the government will need direct estimates of the (relative) marginal utilities in order to calculate the agent's well-being index. The purpose of our SP survey is to demonstrate a method for generating such estimates. We envision governments applying such a method on nationally representative samples and, because the marginal utilities may change as $\mathbf{w}$ changes, doing so on a regular basis (just as prices are currently re-measured on a regular basis).[13]

---

[12] This perspective has parallels with hedonic pricing (Rosen, 1974); the fundamental aspects can be thought of as the ultimate underlying characteristics that people care about when making choices.
[13] As with real consumption indices, there will be a tradeoff between using the correct weights and using weights comparable to those used in previous years. We imagine that some compromise analogous to chain-weighting may emerge. In addition, one could in principle make the approximation in equation (2) more accurate by including higher-order terms. For example, a more complicated version of our survey could estimate the second derivative matrix, $D_{\mathbf{ww}} u(\mathbf{w})$, in addition to the first derivative vector, $D_{\mathbf{w}} u(\mathbf{w})$.

Although in this paper we focus primarily on survey-based measures of the fundamental aspects that comprise w, our theoretical framework also applies when some or all of the aspects are measured objectively. For example, some dimensions of health could be measured with physiological tests, and, arguably, some freedoms could be quantified. Indeed, widely-used indices such as the U.N.'s Human Development Index and Okun's "misery index" (the sum of the inflation and the unemployment rates) consist entirely of objective measures that are combined by using *ad hoc* weights. Regardless of whether the components of w are measured objectively or subjectively, we propose using a data-driven method for estimating the marginal utilities; our SP survey is an attempt at implementing such a method.

While our discussion above focuses on constructing a well-being index for a single agent, in order to construct a national index, governments would need to, first, construct indices for many people and, second, aggregate them. To construct indices for many people, a method such as our SP survey could, in principle, be used to estimate marginal utilities for each person in a representative sample. In practice, if w includes many aspects, doing so would require an impractically long survey, and a more realistic approach may be to partition the population into types based on observable characteristics and pool across individuals to estimate the marginal utilities of each type.[14] Our main specification in section IV treats all respondents as a single type; in the Web Appendix we report results estimated separately for different demographic groups.

How to aggregate utility across individuals is a central question of welfare economics and an active area of research in the literature on social choice. Although this paper focuses on constructing a well-being index for individuals within a single type, in principle our framework and empirical contribution are compatible with a variety of approaches to aggregation.[15]

Then changes in the index could be adjusted by $\frac{1}{2}\Delta w' \cdot D_{ww}u(w) \cdot \Delta w$ to provide a more accurate assessment of changes in well-being. We do not pursue these extensions in this paper.

[14] "Type" is defined by the local slope of the indifference surface. Thus, even if two respondents have the same preferences, if their w's differ in a way that leads to indifference surfaces with different slopes, they should be considered different types. On the other hand, even if two respondents interpret an aspect differently—for example, "being a good, moral person" could mean different things to different people—they are still of the same type as long as that aspect enters their preferences in the same way.

[15] Aggregation requires making additional and controversial assumptions, for example, regarding the extent to which utility can be compared interpersonally; recent contributions include Fleurbaey, Schokkaert, and Decancq (2009) and Fleurbaey and Maniquet (2011). In earlier work, many researchers seem to have implicitly adopted a position that assumes that the social welfare function is utilitarian, i.e., equals to the sum of individual utilities, and that individuals' utility levels can be measured by responses

II. Aspects of Well-Being

A major obstacle to any real-world application of our theoretical framework is that no one knows which fundamental aspects comprise the vector w, nor which broader, composite aspects might capture much of the variation in u.[16] Indeed, different authors have proposed different sets of aspects as important components of well-being.

Our approach in this paper is to construct as comprehensive a list of aspects as we practically can. Doing so has three advantages. First, it reduces the risk of missing important components of w. Second, it allows us to minimize the influence of our own *ex ante* beliefs on the set that emerges as important from our analysis.[17] Third, it renders our results as broadly useful as possible, since different researchers can focus on the subset of our results that pertains to the aspects they believe comprise w—or to those they happen to have data on in an existing social survey. On the other hand, our attempt to be comprehensive has the drawback of increasing the likelihood that different aspects on our list overlap with each other. Our analysis in section V, in which we exclude some potentially overlapping aspects from the analysis, suggests that such possible overlap does not meaningfully affect our marginal utility estimates. However, in selecting which aspects to include in a well-being index, judgment will be required in order to ensure that fundamental aspects are not double-counted.[18,19]

---

to a specific survey question (e.g., a happiness or a life satisfaction question; for discussion see Layard, 2005; Diener et al., 2009). We note, however, that in our framework, small changes in individuals' w's induce small changes in individuals' utilities, and it is these *changes* (rather than *levels*) that need to be aggregated. Hence under the assumption of interpersonal comparability, the change in *any* well-behaved social welfare function (a function of the utilities of individuals)—not just utilitarian—can be approximated to first order by the change in a weighted sum of $\sum_{j=1}^{J} \frac{\overline{\partial u(w)}}{\partial w_j} w_j$ across individuals.

[16] By a "composite aspect" we mean a facet of well-being that is not fundamental but is correlated with many fundamental aspects. Under restrictive conditions that seem unlikely to hold in practice (analogous to the assumptions underlying the composite commodity theorem), a composite aspect can be understood as a fixed bundle of fundamental aspects—e.g., a few positive emotions together, or a set of freedoms— that has a well-defined marginal utility; in that case, there is a formal justification for substituting, in the well-being index in equation (2), the composite aspect for its constituent fundamental aspects.

[17] That said, in compiling the list we draw exclusively on English-speaking sources, introducing a cultural bias—but one compatible with our respondents'. In addition, "dirty" aspects (such as racist objectives) are typically absent from our sources, but as we discuss in section IV, the few aspects on our list that could be considered "dirty" (e.g., social status) are ranked very low by our SP survey respondents anyway.

[18] Such judgment could be based on a combination of data and theory. For example, one could collect panel data from a SWB survey and use factor analysis to identify groups of aspects that co-move. Theory would be required for determining whether the aspects in a group are conceptually distinct.

In this section we summarize our method for compiling our master list of 136 aspects of well-being (see table 2 for a version of the list). The list, as well as much more detail regarding the process of compiling it, dividing it into different sub-lists, and creating different versions of it, is available in the Web Appendix.

Our list draws from six classes of survey measures. First, we include single-question SWB measures, modeled after the SWB questions most commonly asked in large-scale social surveys (for example, those asked in or proposed for the U.K. survey discussed above). These include mostly measures that are considered evaluative/cognitive (e.g., life satisfaction) or hedonic/affective (e.g., felt happiness).

Second, we include measures based on items in multi-question survey measures of SWB. These are primarily drawn from scales commonly used in psychology, such as the PANAS (Positive and Negative Affect Scale).

The third and largest class contains aspects of well-being proposed by prominent economists, psychologists, and philosophers, that are not typically elicited on large surveys. We drew on proposals from the Stiglitz Commission (Stiglitz, Sen, and Fitoussi, 2009), the related systematic compilation effort by Alkire (2002), as well as many classic sources (e.g., Maslow, 1946; Sen, 1985; Nussbaum, 2000) and more recent contributions (Seligman and Diener, 2004; Loewenstein and Ubel, 2008; Graham, 2011). A complete list of the works we reviewed, including references by aspect, is given in the Web Appendix. The many aspects in this class include some that would be considered "eudaimonic" SWB measures (e.g., having a meaningful life; Ryff, 1989), as well as some that can be understood as capabilities, i.e., access to resources, choice sets, and freedoms.

The fourth class of measures resulted from our own introspection and discussion. Some of our proposed additions were confirmed to be important in our past work (Benjamin, Heffetz, Kimball, Rees-Jones, 2012), and others resulted from extensive discussions among ourselves and with colleagues. Some reflect our attempt to break down important parts of life into more fundamental aspects. For example, while many writers have proposed that religion is important for well-being, we refined "religion" into 15 aspects of well-being that may help explain the

---

[19] A related but distinct issue is whether a well-being index should encompass other-regarding preferences; if not, then aspects of others' lives should be omitted.

value of religion but that are also valued by many non-religious people (e.g., "you having people around you who share your values, beliefs and interests").

While these four classes of measures represent our attempt to include fundamental aspects in our list, our fifth and sixth classes serve different purposes. The fifth class represents our attempt to formulate composite-aspect survey questions that might serve as good empirical proxies for u itself—thus possibly obviating the need for an index. These include novel, broad, evaluative well-being measures, such as "how much you like your life" and "the overall well-being of you and your family."

Finally, as a sixth class, we crafted survey versions of "objective" indicators that are widely-used as measures of well-being, such as the rates of GDP growth, inflation, and unemployment, or income inequality. The weights respondents put on such measures can serve as a benchmark against which we can compare measures from our first five classes.

After compiling an initial list, we revised it according to several criteria. To make its length more manageable, we combined similar items into single measures (but we preserved commonly-used survey questions close to their original form). To reduce subject confusion and response error, we oriented all items so that rating higher would conventionally be considered desirable; for example, "not feeling anxious." We further edited items in order to use vocabulary that would be understandable by most respondents in a national sample.

The final list of 136 aspects includes 113 "private good" aspects—relating to an individual's own well-being (e.g., "your health")—and 23 "public good" aspects (also labeled *public*-aspects)—relating to an entire society's well-being (e.g., "equality of opportunity in your nation"). Among the private-good aspects, we distinguish between what we label *you*-aspects— 108 that pertain to the respondent but could in principle pertain to everyone (e.g., "your health")—and *you-only*-aspects—5 that pertain to the respondent and could *not* meaningfully and distinctly pertain to everyone at the same time, for example due to their being inherently relative (e.g., "your social status").

Finally, from each of the 108 *you*-aspects we constructed two additional aspect versions: an *everyone*-aspect that pertains to everyone in a nation (e.g., by replacing "your health" with "people's health"); and an *others*-aspect that pertains to typical others (e.g., "others' health"). In the next section, we explain the purpose of these aspect versions and discuss them further.

III. Survey Design

The core of our online SP survey consists of 30 hypothetical-choice scenarios, one per screen. They are preceded by a screen of detailed instructions and followed by a multi-screen exit questionnaire.

III.A. Scenario Screens

An example scenario, as it appears on the screen, was reproduced in the Introduction. Each such screen has three components. First, the *preamble* frames the scenario as a choice between two options, neutrally labeled "Option 1" and "Option 2," that have different consequences over the next four years. Second, the *aspect table* describes the difference in consequences between the two options. Finally, the *choice question* elicits a participant's stated preference between the two options.

Preamble. The preamble appears in one of two versions. The first version (reproduced in the Introduction) introduces *personal-choice* scenarios. Since much of the discourse regarding SWB surveys is focused on private-good aspects, personal choice seems the relevant setting for eliciting the aspects' weights through respondents' pairwise decisions that trade them off.

The second preamble version introduces *policy-vote* scenarios: the opening clause "Imagine you are **making a personal decision**" is replaced with the clause "Imagine that you and everyone else in your nation are **voting on a national policy issue**."[20] Policy-vote scenarios have two purposes. First, as with standard public goods, our 23 *public*-aspects cannot typically be affected by one individual's personal choice—but are routinely traded off in policy-vote contexts. Second, even for *you*-aspects, if a national SWB survey is to be used for evaluating policy, it may be useful to elicit the relative weights also in a setup where the aspects pertain to everyone (and are traded off by policy in the same way for everyone). Due to other-regarding preferences, for example, these *everyone*-weights could differ from their corresponding *you*-

---

[20] We chose this voting frame for two reasons. First, it is designed to imitate a situation that, while hypothetical, is as close as possible to the policy choice set-up people have real-world experience with; while most people are never in a position to actually choose a policy that would get implemented, people often face choices over real policies in the voting booth. Second, the frame is designed to elicit participants' *own* preferences. While these preferences might well incorporate a concern for others, we did *not* want the choices to put extra weight on others out of concern that these others would be left out of the decision-making process. For this reason the question explicitly states that "you and everyone else in your nation" are voting.

weights elicited in personal-choice scenarios. While our empirical effort is focused on personal scenarios, we also explore such personal-vs.-policy comparisons below.

The rest of the preamble is identical across all scenarios. Designed to elicit participants' "single-period" utility, it explicitly limits the duration of the predicted difference between the options. While four years is a somewhat arbitrary duration, it does not seem unreasonable as a time frame for assessing policy (for example, it is the length of the term of the U.S. President) as well as personal choices. The preamble ends with a sentence that effectively asks participants to imagine that anything not explicitly stated to differ is held constant.

Aspect table. Each row of the aspect table compares the two options in terms of one aspect, with an "X" positioned to indicate that either Option 1 or 2 rates "much higher," "somewhat higher," or "slightly higher" on that aspect. The "about equal" column in the middle never contains an "X"; it serves as a reminder to participants that unlisted aspects are to be considered as affected equally by the two options.

Personal scenarios draw aspects randomly from the set of 113 private aspects (which consists of 108 *you*- and 5 *you-only*-aspects). The policy scenarios we analyze in this paper instead draw aspects randomly from the set of 108 *everyone*- and 23 *public*-aspects; these 131 aspects are effectively public goods because they affect everyone in the same way. Each respondent faces, in random order, 11 personal scenarios, 5 such policy scenarios, and 14 additional exploratory versions of policy scenarios that we do not analyze in this paper.[21]

The number of aspects (or rows) in a scenario's aspect table is randomly drawn from the set {2, 3, 4, 6}. While a shorter table may be easier for respondents to read and think through, a longer one improves statistical power for identifying marginal utilities.

The rating of each aspect—i.e., the location of the "X" in each row—is randomly drawn from the six feasible ratings. However, we place two restrictions on the combination of ratings within a scenario. First, scenarios with an even number of aspects must be *balanced*: exactly half of the aspects favor Option 1 and exactly half favor Option 2. Second, scenarios with 4 or 6 aspects must additionally be *symmetric*: each rating in favor of Option 1—i.e., "much higher,"

---

[21] These additional versions are designed to explore issues related to other-regarding preferences that are beyond the scope of this paper. (For example, to what extent do people's votes on tax policy reflect their willingness to sacrifice their own utility for increasing others'?) These policy scenario versions draw aspects from sets that consist of different combinations of *you*-, *you-only*-, *others*-, and *public*-aspects.

"somewhat higher," or "slightly higher"—is matched by a rating of the same intensity in favor of Option 2 on another aspect.[22]

Choice question. The choice response scale is identical across all scenarios. It is designed to elicit intensity of preference on a six-point scale ("Much prefer Option 1," "Somewhat prefer Option 1," "Slightly prefer Option 1," "Slightly prefer Option 2," etc.). To discourage "lazy" responses, we omitted an "indifferent" option; our "slightly" options are intended to allow for nearly-indifferent choices.

III.B. Instructions and Questionnaire

The instructions screen is reproduced in the Web Appendix. Respondents could re-open it from every scenario screen by clicking a hyperlink. It includes an example aspect table that is more complex than that shown in the Introduction, illustrating and explaining more possibilities. The instructions emphasize a number of scenario design points, including: the distinctions between personal-choice and policy-vote scenarios; that the two options differ *only* on the consequences listed in the aspect table; and that "The items and their rankings in the tables are randomly chosen by the computer so that we [the researchers] learn as much as possible from your choices."[23]

In addition, the instructions explain that in each row of the aspect tables, one word is emphasized in boldface type (in the example in the Introduction, **you** and **your**). Participants are asked to "pay careful attention to the emphasized words, and interpret a consequence with the following emphasized words as affecting the following people:"[24]

---

[22] We require 2-aspect scenarios to be balanced because, otherwise, half of them are expected to elicit a trivial choice, as one of the options would rank higher on all aspects. By placing no restrictions on 3-aspect scenarios, we allow such trivial cases to occur (with 25% probability), and we use them as a secondary robustness check to identify respondents who might have answered at random (see footnote 37 below). Finally, we require 4- and 6-aspect scenarios to be balanced and symmetric because our pre-tests suggested that otherwise, participants faced with these longer scenarios may adopt "visual shortcut" heuristics—such as choosing an option that ranks better on most aspects—and pay less attention to the identities of the aspects. We compare results estimated separately from 2-, 3-, 4-, and 6-aspect scenarios in section V and in the Web Appendix.

[23] This randomization disclosure is intended to prevent respondents from perceiving random combinations of aspects as reflecting an intentional design decision to construct a particular scenario. Such mistaken inferences could give rise to undesired experimenter-demand effects.

[24] We included the bolding and this "legend" (reproduced here as it appears on the screen) because pre-tests without the bolding indicated that participants might quickly skim the aspect tables, mistakenly assuming that, for example, all aspects apply to only them personally. The bolding scheme allows us to

you/your: affects only you (and, when stated, your family or close friends).
others/others': does not affect you, your family, or your friends—but does affect other people. The table indicates the average effect on other people in your nation.
nation/society/people: affects everyone in your nation (including you, your family, and your friends).
world/humanity: affects everyone in the world.

Once participants complete the scenarios, they fill out an exit questionnaire. We ask participants, in both a multiple-choice question and an open-ended question, whether they understood what we were asking them to do (in the multiple choice, 92.5% answered "always" or "mostly"; 7.5% answered "not really"). We also ask basic demographic questions, as well as questions about ideology, political party affiliation, and religiosity.

IV. Empirical Results: Marginal Utility Estimates

Our SP survey respondents were recruited during December 2011 by Clear Voice Research, a private firm that invites individuals to "start sharing their voice" and "make a little money" by participating in online surveys. We aimed at a sample that, although not a random sample, would resemble the adult (20+) U.S. population on the demographics listed in table 1.

IV.A. Respondent Demographics

Table 1 reports the demographic distribution of the 7,391 respondents who began our survey (the "All" column); the 5,397 who completed it ("Completes"); and the 4,608 included in our main analysis below ("Primary Sample").[25] This primary sample excludes respondents who completed the survey in less than eight minutes; our robustness analysis in section V below suggests that including them does not affect our results qualitatively but increases measurement error. The table shows that the three groups are similar on observables.

The rightmost column reports figures from the 2010 American Community Survey, 2010 Census, and 2011 Current Population Survey (see Web Appendix for details). While our respondents are more likely to be married, college-educated, and white, and less likely to have very high income, be Hispanic, and live alone, they are reasonably similar in the other cells.

---

keep the aspects short and simple and at the same time visually clear and coherent. The legend provides respondents with a quick reference that clarifies the bolded words' intended meaning.

[25] Due to a programming error, the first 1,936 primary-sample respondents faced scenarios in which aspects were unintentionally drawn from only 108 of the intended 113 personal and 131 policy aspects. As a result, we have more data—and tighter estimates—on some aspects. Excluding these 1,936 early respondents has very little effect on our main estimates (see Web Appendix table WA1).

IV.B. Personal Choices: Benchmark Specification and Results

Our main results are reported in the "Personal" panel of table 2. Recall from section III that each respondent faced eleven personal scenarios where the two choice options differ on 2, 3, 4, or 6 of the 113 personal aspects. Pooling all such scenarios across all respondents, we report results from the following OLS regression:

(3) $$\text{StatedPreference}_s = \alpha + \textbf{AspectRatings}'_s \cdot \boldsymbol{\beta} + \varepsilon_s.$$

Each observation s captures the information from a single scenario faced by a respondent, corresponding to a single survey screen like the example in the Introduction. $\text{StatedPreference}_s$ encodes the response to the choice question. $\textbf{AspectRatings}_s$, a 113-element vector, encodes the differences between the two options; all of its entries are 0 except for the 2–6 entries representing the aspects on which the options differ. We cluster standard errors at the respondent level.

To the six points on the choice scale ("Much prefer Option 1," etc.) we assign the six numerical values (-1, -0.47, -0.14, +0.14, +0.47, +1), and to the seven columns in the aspects table ("Option 1 much higher," etc.) we assign the values (-1, -0.83, -0.75, 0, +0.75, +0.83, +1). As described in section V below, these numerical values were estimated from the data using a non-linear ordered probit model, and then constraining the scales to be symmetric and range from -1 to +1. The resulting choice scale captures respondents' interpretations of the verbal labels; and the aspects-rating scale captures a combination of respondents' interpretations of the labels, higher derivatives of utility (averaged across the aspects), and any "focusing effect," that is, respondents paying less attention to an aspect's rating than to its direction in favor of one of the options. While we prefer using these estimated scales, misspecifying the scales should have little effect on the estimated aspect coefficients relative to each other. Indeed, as reported in the Web Appendix, for the personal scenarios, the correlation between the coefficients reported in table 2 and those estimated with linear codings (i.e., choice scale: (-3, -2, -1, +1, +2, +3); aspect ratings: (-3, -2, -1, 0, +1, +2, +3)) is 0.998. We find similarly high correlations (0.99) with coefficients estimated from a probit or logit model where we collapse the choice scale to a binary variable (prefer Option 1 vs. Option 2) and use the estimated scales for the aspect columns.

Here we report OLS for maximum transparency, and for simplicity we assume for now that respondents are identical in both their marginal rates of substitution and their use of the

response scale. As reported in section V below, our results are essentially unaffected when we relax many of the restrictions imposed by this specification. We explore heterogeneity in marginal rates of substitution across respondent subpopulations later in this section.[26]

We interpret the regression (3) as estimating equation (2), where in our empirical analysis, Δw is the difference in aspect ratings across the two options in a scenario. The β vector estimates $D_w u(w)$, and ε_s captures response error.[27] We interpret α, which we estimate to be -0.02 (s.e. = 0.003), as picking up a very small respondent bias in favor of "Option 1" despite the fact that the content of the two options is randomly drawn from the same distribution.

The Personal panel in the table reports, for each of the 113 aspect regressors, its coefficient and standard error. The "Rank" column orders the *you*-aspects by coefficient size (1– 108) and, additionally, places the 5 *you-only*-aspects relative to these by assigning to them rank numbers with letter suffixes (e.g., "74a" lies between 74 and 75).[28] Since the independent and dependent variables are coded over ranges of the same length (–1 to +1), a coefficient of, for example, +0.46 means that on average, changing the relevant aspect from the extreme rating "Option 1 much higher" to the other extreme of "Option 2 much higher" causes choice to move 46% of the entire choice scale in the same direction.

Figure 1 summarizes the coefficient and rank information in table 2 graphically, for the 113 personal aspects (x's) and the 131 policy aspects (triangles; we discuss these below), sorted by their respective within-panel rank. Of the 113 personal aspects, all but two are positive, almost all statistically significantly so. This confirms that, as intended by our wording of the aspects, an option rating higher on an aspect is, *ceteris paribus*, generally considered preferable.

---

[26] We note that applying our estimates as weights in a SWB-survey-based index requires neither the assumption that a given aspect is interpreted identically across respondents (see footnote 14) nor that a given respondent uses the rating scales identically across aspects. Rather, we assume that a given respondent: (*i*) interprets the meaning of a given aspect identically across our SP survey and a SWB survey; and (*ii*) translates the SP-survey aspect rating scale into a SWB-survey aspect rating scale the same way across aspects and, up to a multiplicative constant, the same as others who are treated as being of the same type. (Recall that in our main specification, we have only one type.)

[27] Equation (2) holds as a first-order approximation around some reference aspect vector. If respondents interpret the scenarios as describing changes in a neighborhood around their current aspect vector, then β from equation (3) estimates the current marginal utilities. While our scenarios do not explicitly frame the choice as being relative to respondents' current aspect levels, we believe that this is the natural interpretation (see language in the example scenario in the Introduction).

[28] This ranking scheme facilitates comparing the 108 *you*-aspect ranks with their corresponding 108 *everyone*-aspect ranks in the "Policy" panel, accommodating the fact that the *you-only*-aspects do not have counterparts in policy scenarios (and, similarly, that the *public*-aspects do not have counterparts in the "Personal" panel).

(We discuss the two coefficients that are negative below.) The table and the figure show that the greatest variation in coefficient size across aspects occurs among those at the top (for example, the top 10 coefficients range from +0.46 to +0.32), and at the bottom (the bottom 10 range from +0.09 to –0.09); coefficients vary more slowly among middle-ranking aspects. The standard errors on the coefficients are typically below 0.017. These features of our estimates should be borne in mind when reading our discussion below, which focuses on ranks.

Looking at specific aspects, those involving family (well-being [rank 1], happiness [2], and relationship quality [5]); health (general [3] and mental [7]); security (financial [6], about life and the future [8], and physical [21]); values (morality [4] and meaning [10]); and options (freedom of choice [9] and resources [12]) are conspicuous in their predominance at the top of the table, along with—reassuringly—some measures of happiness and life satisfaction (these are discussed below). In contrast, at the very bottom of the table, we find all four *you-only* aspects that involve relative position—power over other people [108c], social status [108b], high relative income [108a], and postmortem fame [106a]—with coefficients that are either negative or close to zero.[29] Since much evidence seems to imply a *high* marginal utility to status and relative position (see Heffetz and Frank, 2011, for a survey), we conjecture that the low ranks of these aspects may reflect respondents' answering our stated preference question in terms of their meta-preferences or laundered preferences.[30]

While we do not view the potential for meta- or laundered preference elicitation as a disadvantage, our estimates may also be sensitive to specific details of our survey design and the underlying respondent population. Later in this section, the next section, and the Web Appendix, we compare estimates based on alternative design details and on alternative subpopulations. Since our general method is held fixed, however, we can only speculate on its effect on our results. The deliberative frame of mind induced by our setup may make evaluative SWB aspects easier for respondents to consider than affective aspects; the double-negative framing of negative emotions may make them harder to think about than positive emotions; and the instruction to

[29] The fifth *you-only*-aspect is non-positional: "the happiness of your friends" [74a]. At the same time, two of the other aspects at the bottom—"your enjoyment of winning, competing, and facing challenges" [108] and, to a lesser extent, a "ladder" aspect modeled after Cantril's Self-Anchoring Scale [103]—could have been interpreted by respondents as involving relative position, although we did not perceive them as such when compiling our aspect list. Indeed, we expected the ladder aspect to rank high, along with other evaluative SWB measures (see IV.C below).

[30] Another possibility is that the low rank of these aspects reflects experimenter-demand effects. While we cannot rule out this concern, we believe it is less likely in an anonymous web-survey like ours.

hold other aspects constant may be more difficult to follow when broader, composite aspects are varied in a scenario than when only narrow aspects are varied. To the extent that our fixed design choices amplify estimated coefficients on some categories of aspects relative to others, comparisons of aspects *within* an aspect category may be more generalizable to alternative design choices than comparisons *across* categories. These caveats should be borne in mind throughout our discussion below.

IV.C. Personal Choices: Discussion

Evaluative and affective SWB. Among happiness and life satisfaction measures, the more evaluative ones—family happiness[31] [2] and life satisfaction [11]—are among the highest-ranking aspects, and rank higher than the more affective ones—"how much of the time you feel happy" [31] and "how happy you feel" [39]. Other measures that past work has classified as positive affect measures, such as "how often you smile or laugh" [35] (e.g., Kahneman and Deaton, 2010), rank similarly to these affective happiness measures.

Negative emotions. Recently, Deaton et al. (2011) suggest that national SWB surveys focus also on measuring negative emotions. In our data, the six measures they recommend (in their "rough order of preference") get the following ranks: pain [49], stress [45], worry [52], anger [76], tired (not on our list, but we have: feeling full of energy [42] and quality of sleep [77]), and sad [64]. This group of measures lies in the middle of our table, with coefficients in the range 0.19–0.25. Other negative emotions, such as frustration [67], are also in this range. Of particular interest because it is the only negative emotion among the four U.K. questions from the Introduction, anxious [92] lies somewhat *below* this range; its coefficient (0.13) is roughly half the coefficients of stress (0.25) and pain (0.24).

Eudaimonic SWB. While evaluative and affective measures dominate the policy discourse on national SWB surveys, researchers increasingly recognize the importance of eudaimonic dimensions of well-being. In our data, eudaimonic aspects such as being a good, moral person and living according to personal values [4] (coeff. 0.40) and having a life that is meaningful and

---

[31] If our respondents interpret our family happiness aspect [rank 2] as including self, then it is more inclusive than typical happiness measures (that refer only to self). As a result, we do not know if its place at the top of the table is due to it being a relatively evaluative happiness measure or to its inclusion of family. (We view this aspect as a relatively evaluative measure since reporting one's family happiness is likely to require more evaluative effort than reporting about one's own emotions.)

has value [10] (coeff. 0.32) indeed rank among the highest aspects. The aspect modeled after the only eudaimonic question in the U.K. four—feeling that the things you do in your life are worthwhile [20]—has a reasonably high coefficient (0.28), yet lower than the two above.

In summary, our agnostic, stated-preference-based approach yields high marginal utility estimates on measures—such as life satisfaction—that have been at the center of the discussion about well-being indices. At the same time, other measures that have received recent attention— such as those of positive and negative affect, and certain eudaimonic measures—have lower coefficients than aspects that have received less attention in this context, in particular, about family, health, security, values, and options. The two measures with the largest coefficients— "the overall well-being of you and your family" and "the happiness of your family"—are, as far as we know, survey questions we invented that have not previously been asked in large-scale surveys. Our results suggest they deserve attention in future data collection efforts.[32]

IV.D. Policy Choices

The "Policy" panel of table 2 reports estimates from a specification identical to that used in the Personal panel but uses data from the five policy scenarios each respondent faced. Recall that in such scenarios, respondents vote on policy, trading off 131 aspects that include *everyone*-aspects (personal aspects that pertain to everyone in the nation) and *public*-aspects (public goods that pertain to the entire nation or, when stated, to the entire world). To make the coefficient magnitudes comparable across the two panels, we use the same numerical scales as in the personal regression (rather than re-estimating them). As mentioned above, to further facilitate such comparison between the *you*-aspects and their corresponding *everyone*-aspects, the "Rank" column ranks the *everyone*-aspects by coefficient size (1−108) and, in addition, places the 23 *public*-aspects relative to these by assigning to them rank numbers with letter suffixes.

Since we collected less data in these policy scenarios than in personal scenarios, standard errors are larger, typically in the 0.023–0.035 range. Nonetheless, the correlation between the 108 *you*- and *everyone*-coefficient pairs is fairly high (0.81).[33] Figure 2 conveys the comparison

---

[32] For example, to capture our overall top-ranked aspect, surveys could ask: "On a scale from 0 to 10, how would you rate the overall well-being of you and your family?"

[33] Here and later, we report the correlation between two vectors of estimated (sample) coefficients. Due to sampling error, this reported correlation is a lower bound on the correlation between the vectors of true (population) coefficients.

graphically by replicating figure 1 for only the 108 *you*- and 108 *everyone*-aspects, both sorted by the rank of the *you*-aspects in the personal scenarios. The dashed curve reports a locally-weighted linear regression of the *everyone*-coefficients (triangles). The figure suggests that on average, *everyone*-coefficients in the policy scenarios are attenuated versions of their counterpart *you*-coefficients in the personal scenarios. This may reflect respondents' greater uncertainty regarding others' preferences, perhaps causing respondents to state preferences with weaker intensity in policy scenarios.

Consistent with the high correlation, some of the high-ranking *you*-aspects retain their high rank as *everyone*-aspects. These include overall well-being of you and your family [personal rank 1; policy rank 3], health [3; 6], personal values [4; 2], and financial security [6; 8]. At the same time, aspects related to freedom and to avoiding abuse seem to rank higher as policy aspects. These include the freedom to choose [9; 1]; your ability to pursue your dreams [36; 14]; being treated with dignity [26; 11]; and, among double negatives, avoiding deception [23; 5], pain [49; 10], and emotional abuse [68; 16].

Perhaps most importantly, several of the 23 *public*-aspects (included in the policy but not the personal scenarios) have among the largest coefficients. These include freedom from corruption, injustice, and abuse of power [0a] (coeff. 0.39), society helping those who struggle [5a], the morality of other people [5b], freedom of speech and of political participation [6a], and the well-being of the people in your nation [6b].[34] High coefficients on freedom from corruption and freedom of speech seem consistent with the tendency for respondents to weight heavily, in a policy context, aspects that expand individuals' choice sets—reducing the need to choose specific outcomes for others. We interpret these findings as providing empirical evidence that when making a policy choice, our respondents put high value on capabilities (Sen, 1985) and basic rights (Rawls, 1971).

Finally, we discuss aspects modeled after widely-used "objective" well-being measures. Some such aspects have large coefficients and have already been discussed above, including freedom from corruption, injustice, and abuse of power; financial security; health; and freedom of speech and of political participation. At the same time, standard macroeconomic indicators have lower ranks: low unemployment [18a], low inflation [43a], GDP growth [57a], GDP per

---

[34] Note that the *public*-aspect "the well-being of people in your nation" [6b] is quite similar to the *everyone*-aspect "the well-being of people and their families" [3]. The higher rank of the latter is consistent with the idea that highlighting "families" may raise the perceived importance of an aspect.

capita [61a], and GDP [88a].[35] The tendency to prefer aspects that increase the choice set (rather than focusing exclusively on outcomes) may help explain why, among the "objective" aspects, the coefficient on equality of opportunity [19b] is relatively large (0.24), larger, for example, than that on equality of income [66a] (0.16).

IV.E. Cross-Group Heterogeneity

To examine cross-group differences in marginal utilities, we produce versions of table 2 for subpopulations of our respondents. Web Appendix tables WA2–WA6 report coefficients, respectively, for men and women; those above and below an income of $50,000/year; liberals, moderates and conservatives; those who do and do not attend religious services at least monthly; and those younger and older than 45. The tables allow for many comparisons and tests; here we briefly summarize only some of the main findings that emerge.

Overall, aspects in the personal scenarios rank similarly across the subpopulations we examine. The sets of 113 coefficients are highly correlated across pairs of disjoint groups, with correlations ranging from 0.86 (liberals vs. moderates) to 0.91 (men vs. women). Moreover, there seems to be a broad consensus across the subpopulations that the highest-ranking aspects include those related to family, health, and security.

To explore which aspects' coefficients change most dramatically across groups, we first normalize our marginal utility estimates to comparable units across groups by dividing each group's set of 113 coefficients by its mean (effectively treating the "average aspect" as numeraire); we then examine, for each aspect and each pair of disjoint groups, the ratio of the two normalized coefficients, dividing the smaller coefficient by the larger (so all ratios are ≤ 1). Tables WA2–WA6 report these ratios. We highlight here aspects that rank in the top ten in one group but drop sufficiently in the complementary group to yield a ratio of less than 0.8. We caution that not only our estimates in table 2, but also these cross-group differences, may not generalize to a more representative sample (cf., Heffetz and Rabin, 2012).

---

[35] Our finding that low unemployment and low inflation have the two highest coefficients *among* macroeconomic indicators may be viewed as consistent with the exclusive focus of Okun's misery index on the two. Furthermore, that the two coefficients are somewhat similar—0.24 and 0.19, respectively—may provide some support for the equal weights of inflation and unemployment in the index, under the assumption that percentage point differences in inflation and unemployment translate similarly into verbal ratings like those in our SP survey, and with the caveats discussed in section VI below.

Men rank higher: "your sense that your life is meaningful and has value" [men rank 5, women rank 29, normalized coefficient ratio (*cr*) = 0.77]. *Women rank higher*: "your mental health and emotional stability" [women 6, men 23, *cr* = 0.71].

High-income rank higher: life satisfaction [high income 5, low income 41, *cr* = 0.70], and having a meaningful life [high 7, low 24, *cr* = 0.75].

Liberals rank higher: having enough time and money [liberals 4, conservatives 30, *cr* = 0.78]. *Conservatives rank higher*: being a good, moral person and living according to personal values [co. 4, li. 6, *cr* = 0.75], family happiness [co. 1, li. 3, *cr* = 0.75], and family relationships [co. 5, li. 9, *cr* = 0.77].

More religious rank higher: sense of purpose [more religious 9, less religious 72, *cr* = 0.60], having people around you who share your values, beliefs and interests [more 10, less 60, *cr* = 0.69], feeling grateful [more 8, less 50, *cr* = 0.74], being good and moral [more 2, less 7, *cr* = 0.74], and making a difference and making the world a better place [more 7, less 35, *cr* = 0.76]. *Less religious rank higher*: life satisfaction [less 9, more 36, *cr* = 0.77].

Older rank higher: having many options and possibilities in life and the freedom to choose among them [older 5, younger 35, *cr* = 0.65].

In policy scenarios, correlations between pairs of sets of the 131 coefficients are lower—although still reasonably high—and range from 0.60 (liberals vs. conservatives) to 0.81 (more vs. less religious). Using the same method as above, all ratios are reported in tables WA2–WA6. Here we highlight aspects with normalized ratios below 0.6. *Women rank higher*: "people being good, moral people…" [women 1, men 35, *cr* = 0.50]. *Men rank higher*: "people getting the rewards and punishments they deserve" [men 5, women 75, *cr* = 0.58]. *Liberals rank higher*: the condition of animals, nature, and the environment [li. 1, co. 113, *cr* = 0.27]. *Conservatives rank higher*: "people's ability to have and raise children" [co. 8, li. 111, *cr* = 0.33]. *Older rank higher*: being treated with dignity and respect [older 6, younger 76, *cr* = 0.53].

V. Robustness and Additional Results

In the previous section we reported results: from (*i*) a simple OLS specification; including (*ii*) our entire personal and policy aspect lists; and pooling responses (*iii*) across scenarios with different numbers of aspects, (*iv*) across all but the speediest-to-answer

respondents, and (*v*) across scenarios faced earlier and later in the survey. In this section, we briefly revisit these points, summarizing analyses reported in more detail in the Web Appendix.

(*i*) *Econometric Specification*: In our OLS specification we coded the verbal scales of the independent and dependent variables as exogenously imposed numeric scales. These were estimated from the following nonlinear ordered probit specification:

$$(4) \quad \text{StatedPreference}_s^* = \left(\mathbf{Option}_s \circ \left[\gamma_{\text{slightly}}\mathbf{1}_{\{\text{slightly}\},s} + \gamma_{\text{somewhat}}\mathbf{1}_{\{\text{somewhat}\},s} + \mathbf{1}_{\{\text{much}\},s}\right]\right)' \cdot \boldsymbol{\beta} + \varepsilon_s.$$

$\text{StatedPreference}_s^*$ is the latent dependent variable; $\mathbf{Option}_s$, $\mathbf{1}_{\{\text{slightly}\},s}$, $\mathbf{1}_{\{\text{somewhat}\},s}$, and $\mathbf{1}_{\{\text{much}\},s}$ are vectors (whose length is the number of aspects) that jointly encode the differences between the two options; and $\circ$ is the entry-wise vector product. Each entry of $\mathbf{Option}_s$ is equal to -1, $+1$, or 0, depending on whether the aspect table rates that aspect higher on Option 1, Option 2, or neither. The entries of $\mathbf{1}_{\{\text{slightly}\},s}$, $\mathbf{1}_{\{\text{somewhat}\},s}$, and $\mathbf{1}_{\{\text{much}\},s}$ are indicators of whether the aspect is rated slightly, somewhat, or much higher. We assume that ε_s is normally distributed and use maximum likelihood to estimate the parameters: γ_{slightly}, γ_{somewhat}, the coefficient vector $\boldsymbol{\beta}$, and the five cutpoints that link $\text{StatedPreference}_s^*$ to the observed choice, $\text{StatedPreference}_s$. Results are reported in Web Appendix table WA7.

Note that specification (4) normalizes "much" to be -1 or $+1$ depending on whether it favors Option 1 or Option 2, and the numerical values for "slightly" and "somewhat" used for the OLS specification in section IV are $\hat{\gamma}_{\text{slightly}}$ and $\hat{\gamma}_{\text{somewhat}}$. To obtain values for the choice scale in the OLS specification, we use the standard normal cdf to calculate the expected value of latent preference intensity conditional on observed preference intensity category; linearly rescale these conditional expectations to lie in the $(-1, +1)$ interval; and symmetrize them around zero by taking the average of the absolute value of each pair of corresponding conditional expectations.

Not surprisingly, since the numerical scales in the OLS regressions reported in section IV are estimated from the nonlinear ordered probit, the correlations between the 113 personal β's across tables 2 and WA7, as well as between the 131 policy β's across the tables, are virtually 1. As an alternative, re-estimating the OLS regressions in table 2 using linear scales as described in section IV.B yields personal and policy coefficients (table WA8) whose correlations with those in table 2 are above 0.99. As a variant that allows respondents to differ in their interpretation of the choice intensities, we first normalize the choice scale at the respondent level by stretching the

linear scale so that the variance across each respondent's 30 choices is 1, and only then estimate the OLS (table WA8). The correlations between the coefficients estimated with and without this normalization are at least 0.98. For probit and logit models, see table WA9.

(*ii*) *Fundamental and Composite Aspects*: If a composite aspect and a fundamental aspect appear in the same scenario, then the presence of the composite aspect might affect respondents' interpretation of the fundamental aspect. For example, if "the overall well-being of you and your family" is a function of "your family's happiness," then a respondent asked to trade them off might—contrary to our intention—interpret the former as meaning overall well-being exclusive of family happiness. Depending on the prevalence of such situations and on how respondents interpret them, our estimated coefficients might be biased.

To probe the robustness of our results to this potential concern, we re-estimate our benchmark OLS model leaving out scenarios containing aspects that seem most likely to be functions of other aspects on our list. For example, for personal choices we exclude all aspects that we view as evaluative SWB measures—including both commonly used ones (e.g., life satisfaction) and our novel proposals (e.g., how desirable your life is).[36] The estimated coefficients on the 94 remaining aspects (table WA10) are broadly similar to those reported in table 2 (correlation 0.99). Results are similar in other specifications, for example, excluding the macroeconomic indicators in policy vote scenarios (table WA11).

(*iii*) *Number of Aspects per Scenario*: As explained in section III.A, respondents face 2-, 3-, 4-, and 6-aspect scenarios. Re-estimating our OLS model separately for each of these four scenario designs, we find that the range of coefficient sizes roughly halves from 2- to 6-aspect scenarios. The correlations between the four sets of 113 and of 131 coefficients range from 0.84 to 0.92 in personal and from 0.52 to 0.70 in policy scenarios (but note that the latter are more attenuated because the policy coefficients' estimation errors are larger). Our general conclusion is that while respondents allow each aspect to influence their stated preference less intensely *per aspect* when the number of aspects is larger, the rank order of coefficients remains rather stable in personal scenarios and somewhat stable in policy scenarios. Further details are provided in table WA12.

---

[36] As our data and code will be available on our websites upon acceptance at a journal, an interested reader can readily re-estimate our model using her or his preferred subset of aspects.

(*iv*) *Respondents' Effort and Comprehension*: Respondents may exert little effort on unincentivized surveys. This may bias the estimated coefficients away from the true marginal utilities, for example by "compressing" aspects' coefficients toward each other if respondents pay less attention to each aspect's identity, or by biasing coefficients toward zero if respondents answer randomly. Using amount of time to complete the survey as a proxy for effort level, and re-estimating our main OLS specification separately by approximate sextiles (table WA13), we indeed find that the coefficients of the speediest sextile (less than 8 minutes) are severely attenuated relative to other sextiles'. Outside the speediest sextile, coefficient sizes seem to increase with completion time less dramatically, peaking at the second-slowest sextile (21–31 minutes). Furthermore, the correlations between the personal coefficients of the speediest sextile and of other sextiles range from 0.23 to 0.32, much lower than the correlations between pairs of other sextiles' coefficient sets, which range from 0.76 to 0.89. For these reasons, our benchmark OLS specification reported in table 2 excludes respondents who took less than 8 minutes to complete the survey.[37] Further excluding those who reported in the exit questionnaire that they did "not really" understand what they were asked to do (see III.B above) yields virtually identical estimates (table WA15).

(*v*) *Early vs. Late Scenarios*: Responses made later in the survey may be less reliable due to tiredness or boredom. Alternatively, they may be *more* reliable due to practice. Also, respondents' interpretations of particular aspects may change over the course of the survey. For example, respondents may interpret life satisfaction as a broad SWB measure early on, but as they face new examples of affective SWB measures, they may interpret it to exclude feelings.

To assess these possibilities, we estimate an augmented version of our OLS specification by including a dummy for whether a scenario appeared in the earlier half of the survey and interacting it with each aspect's rating (table WA16). We find no evidence of systematic differences between estimates from scenarios in earlier and later halves of the survey.

VI. Pragmatics

---

[37] As an additional respondent-effort sensitivity check, we examine respondents who, in at least one three-aspect scenario where one option happens to rate higher on all three aspects, choose the other option. The rankings of aspects in this sample are similar to our benchmark, but the coefficients are greatly attenuated (see table WA14). We like this sensitivity check less because it may fail to drop random responders (who choose the higher-rated option by chance) and, at the same time, may unduly drop respondents who prefer *less* of certain aspects.

We now return to the two practical issues we posed in the Introduction: which questions should a SWB survey ask, and how should the responses be weighted? In theory, the well-being index formula, $\sum_{j=1}^{J} \frac{\overline{\partial u(w)}}{\partial w_J} w_j$, provides clear solutions: ask about the levels of each of the J fundamental aspects comprising w, and weight the responses by their marginal utilities. In practice, both our proposed list of fundamental aspects and our marginal-utility estimates are only first passes in need of further development. Also, since our SP survey provides only one source of evidence, the appropriate weights from a Bayesian point of view should be a compromise between our estimates and one's priors. Nonetheless, our estimates can provide a point of departure for a disciplined discourse in which rival proposals would explicitly communicate priors for the marginal utilities and for biases that our SP survey estimates need to be adjusted for.

Because governments may only be able to afford to ask a limited number of SWB questions on national surveys, the rest of this section revisits these two practical questions with such a constraint: which items should a government include in a SWB survey that is limited to only $N < J$ questions? And, conditional on asking N questions (not necessarily the optimal ones), how should the responses be weighted? (The second question is also relevant for applied researchers who want to use existing SWB survey data to proxy for utility.) The added constraint introduces new considerations, both theoretical and pragmatic.

Conditional on a SWB survey eliciting responses to a given set of N questions, $\{r_1, r_2, \ldots, r_N\}$, a natural weighting approach is to adopt weights, $\alpha_1, \alpha_2, \ldots, \alpha_N$, such that the "abridged index" $\sum_{n=1}^{N} \alpha_n r_n$ is the best predictor of the full index in an R^2 sense. The optimal weights will generally differ from the aspects' marginal utilities because included aspects will proxy for excluded ones with which they covary. If, in addition, the N questions can be selected, then the government could select those that, weighted optimally, best predict the full index. The optimal set may include composite-aspect questions.

As a concrete example of weighting issues, imagine that a SP survey conducted on a U.K. representative sample obtained the estimates in table 2, and consider weighting the four U.K. questions using our estimates: index = 0.31 *satisfied* + 0.28 *worthwhile* + 0.26 *happy* – 0.13 *anxious*. Since stress (coefficient = 0.25), for example, is omitted from the index and is likely more strongly correlated with anxiety than with the other three included aspects, the optimal weight on anxiety could be more negative than –0.13. As to optimal selection of questions, due

to such covariances between the items, the optimal N questions are not necessarily those with the largest marginal-utility estimates. Our SP survey does not provide information concerning these covariances because we varied the aspects' levels exogenously.

To construct an abridged index, it would therefore be useful to first collect data in what we call a SWB "pre"-survey, which asks a smaller sample of respondents a broader set of questions. Intuitively, the optimal abridged index could be calculated by finding the regression of the full index on a set of N questions that maximizes R^2. In practice, to reduce overfitting, it will generally be necessary to impose additional structure on the method of choosing the N questions.[38]

Two additional complexities arise with an abridged index. First, *which* covariance matrix of the aspects' levels to use when seeking to maximize R^2 depends on the index's intended use. For example, if the abridged index is used mainly to track individuals' well-being over time, then the individual-level, time-series covariance matrix should be used.[39] Alternatively, if it is used to guide policy, then question selection and weighting should account for the fact that some aspects may vary a great deal but be relatively immune to policy, while others may move little unless changed by policy.[40]

Second, if the aspects' covariance matrix shifts over time, then the optimal questions and weights may change. This point is related to the well-known Lucas critique from macroeconomics; even if the underlying structural model is fixed, the best-fitting reduced-form equation may be unstable as circumstances shift. For this reason, the aspects' covariance matrix will need to be re-estimated regularly.

---

[38] Measurement error in the SWB survey presents further complications. For a SWB survey that elicits all J fundamental aspects, the optimal index weights are affected to the extent that measurement error differs across aspects. If the survey is constrained to N questions, then even identical measurement error can affect the optimal weights (and optimal set of questions) by affecting how well included questions proxy for omitted aspects. Ideally, the SWB pre-survey would (in addition to asking the SWB questions at regular intervals) sometimes ask the same respondents the same questions at very short time intervals. This would enable one to estimate, and correct for, the aspects' measurement error covariance matrix.

[39] For cross-group SWB comparisons—currently the focus of much of the SWB literature—the relevant covariance is cross-sectional. We caution that such comparisons rely on much stronger assumptions; for example, that different groups use the response scales similarly and that utility differences are interpersonally comparable.

[40] Formally, to calculate weights and select questions that are optimal for some class of policies, one would use the covariance matrix of the effects of such policies on the aspects (while applied economists routinely estimate the effect of policy on a set of outcomes, here the *covariance* of the effects across outcomes is relevant).

VII. Concluding Remarks

Our current system of national accounts has been continually refined over many decades. Converging on how well-being should be tracked is likely to be similarly arduous. Much work has already been done. But that work, and our own contributions—the theory, aspect list, and marginal utility estimates—are only a bare beginning. We address only a fraction of the issues in constructing well-being indices that we are aware of, and there are surely many other issues we have not thought of.

Perhaps the most urgent unresolved theoretical issue is aggregation across individuals. This issue is not specific to our approach. As a workaround, researchers (say, using aggregate consumption or even GDP as a welfare measure) often take the leap of assuming a representative agent. If one were willing to assume a representative agent, a national well-being index could be constructed from marginal utility estimates like ours, together with average responses to a SWB survey. However, developing better approaches to aggregation is a high priority (see Fleurbaey, Schokkaert, and Decancq, 2009, for a recent contribution).

Regarding implementation, while we have assumed throughout that aspects of well-being can be meaningfully measured with a SWB survey, such measurement still faces major challenges (see Adler, 2012, for a comprehensive critical review). For example, traditional SWB measures seem over-sensitive to immediate context (e.g., Schwarz and Strack, 1991) and under-sensitive to lasting changes in life circumstances (e.g., Oswald and Powdthavee, 2008). Such issues are active topics of research.

We have focused on constructing a well-being index based on combining a SWB survey with a SP survey. But the framework and method we have developed could be applied in three additional directions. First, as mentioned above, one could use a SP survey to obtain weights for existing indices of objective measures, such as the Human Development Index (HDI), that currently use ad hoc weights. In that case, the objective measures—for the HDI: longevity, education, and GDP per capita—would replace the aspects in the SP survey.

Second, one could extend market-price-based indices, such as GDP, to incorporate other factors using "price" imputations. Rather than a function of fundamental aspects, utility would be modeled as a function of market goods as well as non-market goods such as leisure, social relationships, and the environment. These goods would replace the aspects in the SP survey.

Third, while our SP survey and analysis are based on the assumption that respondents' stated preferences can be used to assess welfare—an approach we find attractive on liberalist grounds—our methodology could be adapted to accommodate alternative assumptions. For example, if one assumed that life satisfaction equaled welfare, then one would replace our stated choice question with a predicted-life-satisfaction question. Alternatively, one might replace "you" with "someone like you" in the stated choice question if one believed that the latter would yield more reliable responses.

While we view this paper primarily as proposing a long-term agenda, our findings also point to a few readily actionable steps. First, our results suggest prioritizing the measurement of aspects related to family, health, and security; eudaimonic and especially evaluative SWB measures; and, especially in the policy context, freedoms and capabilities. Second, as discussed in section VI, for the purpose of selecting specific questions for a SWB survey, we highlight the value of gathering data on as many aspects as possible—at least initially and even if on a smaller sample (cf. Deaton et al., 2011). Third, along with conducting SWB surveys, we call for governments and researchers to devote resources to estimating aspects' marginal utilities; our SP survey illustrates one method for doing this.

References

Adler, Matthew D. 2012. "Happiness Surveys and Public Policy: What's the Use?" University of Pennsylvania Law School Research Paper 12–36, http://ssrn.com/abstract=2076539.

Adler, Matthew D., and Paul Dolan. 2008. "Introducing a 'Different Lives' Approach to the Valuation of Health and Well-Being." University of Pennsylvania Law School ILE Research Paper 08-05, available at http://ssrn.com/abstract=1112636, March.

Alkire, Sabina. 2002. **"**Dimensions of Human Development." *World Development*, 30(2): 181–205.

Benjamin, Daniel J., Ori Heffetz, Miles S. Kimball, and Alex Rees-Jones. 2012. "What Do You Think Would Make You Happier? What Do You Think You Would Choose?" *American Economic Review*, 102(5): 2083–2110.

Camerer, Colin F., and Robin M. Hogarth. 1999. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework." *Journal of Risk and Uncertainty*, 19(1-3): 7–42.

Camerer, Colin F., Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin. 2003. "Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism.'" *University of Pennsylvania Law Review*, 151(3): 1211–1254.

Deaton, Angus, Daniel Kahneman, Alan Krueger, David Schkade, Norbert Schwarz, and Arthur Stone. 2011. Memo to the Office of National Statistics' Advisory Group on Subjective Well-Being. Published in the Supporting Documents for the Meeting to Provide Guidance to the Organisation for Economic Co-operation and Development on its Plans to Measure Self-Reported Well-Being held on July 8, 2011.

Decancq, Koen, and María Ana Lugo. 2013. "Weights in Multidimensional Indices of Well-Being: An Overview." *Econometric Reviews*, 32(1): 7–34.

Decancq, Koen, Luc Van Ootegem, and Elsy Verhofstadt. 2011. "What if we voted on the weights of a multidimensional well-being index? An illustration with Flemish data." University of Antwerp CSB Working Paper No. 11/10, November.

Diener, Ed. 2000. "Subjective Well-Being: The Science of Happiness and a Proposal for a National Index." *American Psychologist*, 55(1): 34–43.

Diener, Ed. 2006. "Guidelines for National Indicators of Subjective Well-Being and Ill-Being." *Applied Research in Quality of Life*, 1: 151–157.

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

Diener, Ed, and Martin E.P. Seligman. 2004. "Beyond money: Toward an economy of well-being." *Psychological Science in the Public Interest*, 5: 1–31.

Ding, Min, Rajdeep Grewal, and John Liechty. 2005. "Incentive-Aligned Conjoint Analysis." *Journal of Marketing Research*, 42(1): 67–82.

Dolan, Paul, Richard Layard, and Robert Metcalfe. 2011. "Measuring Subjective Well-being for Public Policy." Office for National Statistics.

Fleurbaey, Marc. 2009. "Beyond GDP: The Quest for a Measure of Social Welfare." *Journal of Economic Literature*, 47(4): 1029–1075.

Fleurbaey, Marc, and François Maniquet. 2011. *A Theory of Fairness and Social Welfare*. New York: Cambridge University Press.

Fleurbaey, Marc, Erik Schokkaert, and Koen Decancq. 2009. "What good is happiness?" CORE Discussion Paper 17, March.

Frey, Bruno, and Alois Stutzer. Forthcoming. "The use of happiness research for public policy." *Social Choice and Welfare*. DOI 10.1007/s00355-011-0629-z

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

Graham, Carol. 2011. *The Pursuit of Happiness: An Economy of Well-Being*. Washington, D.C.: Brookings Institution Press.

Green, Paul E., and Vithala R. Rao. 1971. "Conjoint Measurement for Quantifying Judgmental Data." *Journal of Marketing Research*, 8(3): 355–363.

Heffetz, Ori, and Robert H. Frank. 2011. "Preferences for Status: Evidence and Economic Implications." In *Handbook of Social Economics Vol.* 1A, edited by Jess Benhabib, Alberto Bisin, and Matthew Jackson: 69–91. Amsterdam: North-Holland.

Heffetz, Ori, and Matthew Rabin. 2012. "Conclusions Regarding Cross-Group Differences in Happiness Depend on Difficulty of Reaching Respondents." Manuscript, Cornell University, available at http://ssrn.com/abstract=2101097, July.

Kahneman, Daniel, and Angus S. 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, and Alan B. Krueger. 2006. "Developments in the Measurement of Subjective Well-Being." *Journal of Economic Perspectives*, 20(1): 3–24.

Layard, Richard. 2005. *Happiness: Lessons from a New Science*. The Penguin Press.

Loewenstein, George, and Peter A. Ubel. 2008. "Hedonic Adaptation and the Role of Decision and Experience Utility in Public Policy." *Journal of Public Economics*, 92(8-9): 1795–1810.

Maslow, Abraham. 1946. "Theory of Human Motivation." In *Twentieth century psychology: recent developments in psychology*, ed. Philip Lawrence Harriman. New York, NY: Philosophical Library.

Micklewright, John. 2001. "Should the UK government measure poverty and social exclusion with a composite index?" Manuscript, UNICEF Innocenti Research Centre, Florence.

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

ONS. 2011. "Initial investigation into Subjective Well- being from the Opinions Survey." Office for National Statistics.

Oswald, Andrew, and Nattavudh Powdthavee. 2008. "Does happiness adapt? A longitudinal study of disability with implications for economists and judges." *Journal of Public Economics*, 92: 1061–1077.

Rawls, John. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition*." Journal of Political Economy* 82(1): 34–55.

Ryff, Carol D. 1989. "Happiness is Everything, or Is It? Explorations on the Meaning of Psychological Well-Being." *Journal of Personality and Social Psychology*, 57: 1069–1081.

Schwarz, Norbert, and Fritz Strack. 1991. "Evaluating one's life: a judgment model of subjective well-being." In *Subjective Well-Being*, ed. Fritz Strack, Michael Argyle, and Norbert Schwarz. Elmsford, New York: Pergamon Press.

Sen, Amartya. 1979. "Utilitarianism and Welfarism." *Journal of Philosophy*, 76(9): 463–489.

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

Sen, Amartya. 1992. *Inequality Reexamined*. New York: Russell Sage Foundation; Cambridge: Harvard University Press.

Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress*. www.stiglitz-sen-fitoussi.fr

Table 1. Respondent Demographics

		All (N = 7391)	Completes (N = 5397)	Primary Sample (N = 4608)	Census Etc.
Marital Status	Married	59.3	58.9	59.8	48.8
	Never Married	24.3	25.5	23.4	32.1
	Other	16.4	15.6	16.8	19.1
Highest Education Level Completed	High School Grad	23.2	21.9	21.4	42.9
	Some College	40.7	39.8	41.0	28.9
	Bachelor's Degree	24.3	25.4	25.5	17.7
	Graduate Degree	11.5	12.7	12.0	10.4
Age	20-29	19.7	21.1	18.0	18.9
	30-39	18.3	19.5	18.4	17.8
	40-49	19.3	20.0	20.0	19.3
	50-64	20.8	20.4	22.0	26.1
	65 and older	21.9	19.0	21.5	17.9
Income	less than $20,000	17.9	18.1	17.4	19.8
	$20,000-39,999	27.9	27.0	27.9	21.7
	$40,000-49,999	10.9	10.2	10.8	8.9
	$50,000-74,999	19.8	20.0	20.5	17.7
	$75,000-99,999	11.3	11.7	11.6	11.4
	$100,000 and above	12.1	13.0	11.8	20.4
Region	Midwest	23.0	23.2	24.3	21.7
	Northeast	19.8	20.0	19.3	17.9
	South	34.0	33.4	33.8	37.1
	West	23.2	23.4	22.6	23.3
Race	White	76.6	75.2	78.2	63.7
	Black	9.8	10.1	9.6	12.2
	Hispanic/Latino	7.7	8.2	6.4	15.4
	Asian	3.7	4.2	3.5	4.7
	Other	2.2	2.2	2.2	4.0
Household Size	1	18.0	17.7	18.2	26.7
	2	35.2	33.8	35.1	32.8
	3	18.8	19.1	18.6	16.1
	4 and above	27.8	29.2	27.9	24.4
Employment Status	Employed	51.2	53.0	50.7	57.0
	Unemployed	8.5	8.6	8.7	6.9
	Not in labor force	37.4	35.2	37.8	35.6

Notes: All numbers are percentages. "All": respondents who began the survey. "Completes": respondents who completed all scenarios. "Primary Sample": respondents who took at least 8 minutes to complete the survey. Sources: Authors' survey, 2010 American Community Survey, 2010 Census, 2011 Current Population Survey.

Table 2. Marginal Utility Estimates

Aspect	Personal			Policy		
	Coef.	S.E.	Rank	Coef.	S.E.	Rank
freedom from corruption, injustice, and abuse of power in your nation				0.39	0.026	0a
the overall well-being of you and your family	0.46	0.016	1	0.33	0.033	3
the happiness of your family	0.43	0.017	2	0.24	0.024	21
your health	0.42	0.017	3	0.29	0.025	6
you being a good, moral person and living according to your personal values	0.40	0.017	4	0.35	0.025	2
the quality of your family relationships	0.37	0.017	5	0.25	0.024	13
society helping the poor and others who struggle				0.30	0.024	5a
the morality, ethics, and goodness of other people in your nation				0.29	0.024	5b
your financial security	0.34	0.017	6	0.28	0.023	8
freedom of speech and people's ability to take part in the political process and community life				0.29	0.025	6a
the well-being of the people in your nation				0.29	0.024	6b
your mental health and emotional stability	0.34	0.016	7	0.25	0.025	15
your sense of security about life and the future in general	0.33	0.016	8	0.26	0.024	12
you having many options and possibilities in your life and the freedom to choose among them	0.32	0.017	9	0.35	0.034	1
the amount of freedom in society				0.27	0.025	9a
your sense that your life is meaningful and has value	0.32	0.017	10	0.27	0.023	9
how satisfied you are with your life	0.31	0.017	11	0.18	0.033	53
you feeling that you have enough time and money for the things that are most important to you	0.30	0.017	12	0.21	0.023	32
how much you like your life	0.30	0.017	13	0.19	0.031	46
how peaceful, calm, and harmonious your life is	0.29	0.017	14	0.24	0.024	18
your nation being a just society				0.25	0.023	14a
your feeling of independence and self-sufficiency	0.29	0.016	15	0.23	0.024	25
your pride and respect for yourself	0.29	0.017	16	0.19	0.024	44
your sense that you are standing up for what you believe in	0.29	0.017	17	0.21	0.025	33
your sense that you are making a difference, actively contributing to the well-being of other people, and making the world a better place	0.29	0.017	18	0.32	0.025	4
how low the rate of unemployment is in your nation				0.24	0.024	18a
how much you enjoy your life	0.29	0.016	19	0.24	0.025	17
trust among the people in your nation				0.24	0.023	19a
equality of opportunity in your nation				0.24	0.023	19b
the extent to which you feel the things you do in your life are worthwhile	0.28	0.016	20	0.24	0.032	23
your physical safety and security	0.28	0.016	21	0.24	0.023	22
the well-being of the people in the world				0.24	0.024	21a
you "being the person you want to be"	0.28	0.017	22	0.17	0.024	62
your freedom from being lied to, deceived, or betrayed	0.28	0.017	23	0.30	0.026	5
people getting the rewards and punishments they deserve				0.23	0.025	23a
you having people you can turn to in time of need	0.28	0.016	24	0.28	0.024	7
the extent to which you "have a good life"	0.28	0.016	25	0.24	0.032	19
the condition of animals, nature, and the environment in the world				0.22	0.026	25a
you having the people around you think well of you and treat you with dignity and respect	0.27	0.017	26	0.26	0.023	11
how grateful you feel for the things in your life	0.27	0.017	27	0.22	0.032	26
your sense of control over your life	0.27	0.017	28	0.24	0.023	20
how much love there is in your life	0.27	0.016	29	0.20	0.025	42
how much you appreciate your life	0.27	0.016	30	0.18	0.033	58

Aspect	Personal			Policy		
	Coef.	S.E.	Rank	Coef.	S.E.	Rank
how much of the time you feel happy	0.27	0.017	31	0.20	0.025	41
your sense that things are getting better and better	0.27	0.016	32	0.19	0.032	49
your sense that you know what to do when you face choices in your life	0.26	0.016	33	0.18	0.023	55
the extent to which humanity does things worthy of pride				0.21	0.024	33a
you having people around you who share your values, beliefs and interests	0.26	0.016	34	0.21	0.024	29
how often you smile or laugh	0.26	0.017	35	0.14	0.024	82
your ability to dream and pursue your dreams	0.26	0.016	36	0.25	0.024	14
your chance to live a long life	0.26	0.016	37	0.19	0.025	45
the amount of love in the world				0.20	0.025	37a
how fulfilling your life is	0.26	0.017	38	0.15	0.037	78
how happy you feel	0.26	0.016	39	0.23	0.034	24
how glad you are to have the life you have rather than a different life	0.25	0.016	40	0.15	0.034	73
your passion and enthusiasm about things in your life	0.25	0.016	41	0.15	0.024	77
you feeling alive and full of energy	0.25	0.016	42	0.15	0.023	76
your ability to fulfill your potential	0.25	0.016	43	0.21	0.022	28
how low the rate of inflation is in your nation's economy				0.19	0.024	43a
your ability to be yourself and express yourself	0.25	0.016	44	0.19	0.023	47
the absence of stress in your life	0.25	0.017	45	0.15	0.024	74
your ability to keep good perspective in your life	0.25	0.017	46	0.18	0.025	57
your sense of purpose	0.25	0.017	47	0.21	0.023	35
the amount of order and stability in your life	0.24	0.016	48	0.18	0.023	56
your freedom from pain	0.24	0.017	49	0.26	0.022	10
you feeling that things are going well for you	0.24	0.016	50	0.21	0.035	34
the quality of your romantic relationships, marriage, love life or sex life	0.24	0.017	51	0.16	0.026	69
the absence of worry in your life	0.23	0.016	52	0.20	0.024	38
your sense that you are competent and capable in the activities that matter to you	0.23	0.017	53	0.20	0.024	39
your physical comfort	0.23	0.015	54	0.12	0.023	88
the amount of order and stability in society				0.18	0.025	54a
how full of beautiful memories your life is	0.23	0.016	55	0.11	0.032	90
your success at accomplishing your goals	0.23	0.017	56	0.21	0.023	36
your ability to shape and influence the things around you	0.23	0.016	57	0.19	0.024	50
the rate of economic growth (GDP growth) over time in your nation				0.18	0.026	57a
you feeling that your life has direction	0.22	0.017	58	0.16	0.023	67
how rewarding the activities in your life are	0.22	0.016	59	0.16	0.034	68
you getting the things you want out of life	0.22	0.017	60	0.15	0.033	79
your sense of optimism about your future	0.22	0.016	61	0.22	0.034	27
the average income of people in your nation (GDP per capita)				0.17	0.024	61a
you feeling that you have been fortunate in your life	0.22	0.016	62	0.18	0.033	54
the extent to which your nation does things worthy of pride				0.17	0.024	62a
your knowledge, skills, and access to information	0.22	0.016	63	0.21	0.023	31
the absence of sadness in your life	0.22	0.017	64	0.20	0.024	37
how often you can feel relaxed instead of feeling your life is hectic	0.22	0.016	65	0.13	0.023	85
your sense of achievement and excellence	0.21	0.016	66	0.16	0.025	65
equality of income in your nation				0.16	0.024	66a
the absence of frustration in your life	0.21	0.016	67	0.17	0.025	61
your freedom from emotional abuse or harassment	0.20	0.017	68	0.25	0.025	16
you not feeling depressed	0.20	0.016	69	0.15	0.025	80

Aspect	Personal			Policy		
	Coef.	S.E.	Rank	Coef.	S.E.	Rank
your ability to have and raise children	0.20	0.017	70	0.18	0.024	59
you feeling that you are part of something bigger than yourself	0.20	0.016	71	0.17	0.024	60
you having many moments in your life when you feel inspired	0.20	0.016	72	0.15	0.025	71
the amount of pleasure in your life	0.20	0.017	73	0.19	0.024	43
your personal growth	0.19	0.016	74	0.14	0.024	81
the happiness of your friends	0.19	0.024	74a			
how often you are able to challenge your mind in a productive or enjoyable way	0.19	0.016	75	0.15	0.024	75
the absence of anger in your life	0.19	0.017	76	0.19	0.025	51
the quality of your sleep	0.19	0.016	77	0.09	0.025	96
you feeling that you understand the world and the things going on around you	0.19	0.016	78	0.15	0.025	70
your sense that everything happens for a reason	0.18	0.016	79	0.09	0.024	95
the absence of fear in your life	0.18	0.017	80	0.20	0.026	40
how easy and free of annoyances your life is	0.17	0.016	81	0.07	0.025	98
how desirable your life is	0.17	0.016	82	0.09	0.035	97
your ability to fully experience the entire range of healthy human emotions	0.17	0.016	83	0.19	0.024	48
your ability to use your imagination and be creative	0.17	0.017	84	0.16	0.023	66
your sense of discovery and wonder	0.16	0.016	85	0.11	0.024	91
freedom of conscience and belief in your nation				0.13	0.024	85a
how close your life is to being ideal	0.16	0.017	86	0.12	0.032	87
your sense of community, belonging, and connection with other people	0.16	0.017	87	0.21	0.024	30
you not being lonely	0.15	0.016	88	0.17	0.024	64
the total size of your nation's economy (GDP)				0.12	0.024	88a
you feeling that you are understood	0.15	0.017	89	0.19	0.024	52
your absence of internal conflict (conflict within yourself)	0.15	0.016	90	0.15	0.024	72
the absence of regret you feel about your life	0.14	0.016	91	0.02	0.033	105
you not feeling anxious	0.13	0.017	92	0.09	0.024	94
how interesting, fascinating, and free of boredom your life is	0.13	0.016	93	0.03	0.023	103
you having new things, adventure, and excitement in your life	0.12	0.016	94	0.13	0.024	86
the amount of fun and play in your life	0.12	0.016	95	0.11	0.024	89
your sense of connection with the universe or the power behind the universe	0.12	0.017	96	0.10	0.025	93
how much beauty you experience in your life	0.11	0.017	97	0.04	0.025	101
your material standard of living	0.11	0.016	98	0.06	0.026	100
the overall quality of your experience at work	0.10	0.017	99	0.14	0.023	83
you having a role to play in society	0.10	0.016	100	0.17	0.024	63
your opportunities to participate in ceremonies, cultural events, and celebrations that are meaningful to you	0.09	0.016	101	0.11	0.024	92
how often you become deeply engaged in your daily activities (so deeply engaged that you lose track of time)	0.09	0.017	102	0.01	0.023	106
your rating of your life on a ladder where the lowest rung is "worst possible life for you" and the highest rung is "best possible life for you"	0.09	0.016	103	0.04	0.030	102
the absence of shame and guilt in your life	0.07	0.016	104	0.00	0.025	107
you having a beautiful life story, or a life that is "like a work of art"	0.07	0.016	105	0.13	0.034	84
the absence of humiliation and embarrassment in your life	0.07	0.016	106	0.07	0.025	99
you having others remember you and your accomplishments long after your death	0.04	0.022	106a			
your ability to "be in the moment"	0.04	0.017	107	-0.02	0.023	108
your enjoyment of winning, competing, and facing challenges	0.04	0.016	108	0.02	0.024	104

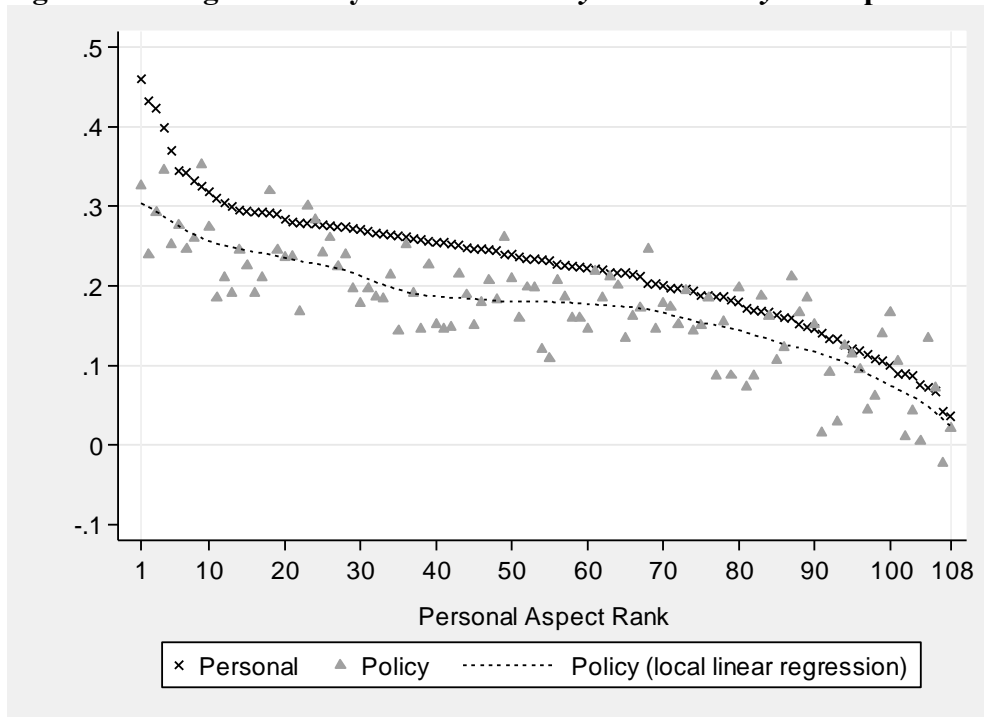| Aspect	Personal			Policy		
	Coef.	S.E.	Rank	Coef.	S.E.	Rank
how high your income is compared to the income of other people around you	0.03	0.022	108a			
your social status	-0.06	0.022	108b			
your power over other people	-0.09	0.022	108c			

Notes: Personal panel: OLS regression of stated preference on 113 personal aspects and a constant (const. = -0.02, s.e. = 0.003), using personal choice scenarios ($N = 50,688$). Policy panel: OLS regression of stated preference on the 131 policy aspects and a constant (const. = -0.01, s.e. = 0.004) using the policy vote scenarios ($N = 23,040$). Standard errors clustered at the respondent level. For the 108 aspects that have versions that appear in both types of scenarios, the text used in the personal choice scenarios is shown; in policy choice scenarios, "people" and "people's" replace "you and "your." For each scenario type, the numbers 1–108 are used to rank, by coefficient size, aspects that appear in both scenario types. An aspect that appears in one scenario type receives a rank with a letter: 5a, for example, indicates that the coefficient was between the aspects ranked 5 and 6.

Figure 1. Marginal Utility Estimates
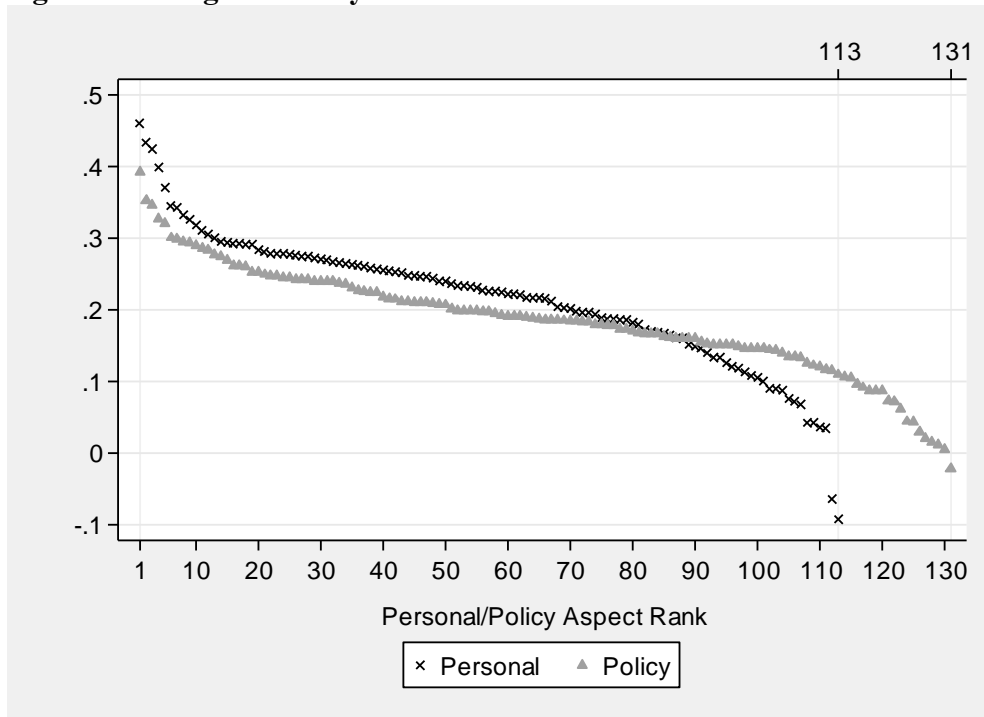
Notes: Aspect coefficients by rank, from benchmark OLS regressions (table 2), separately for 113 personal-scenario aspects (x's) and 131 policy-scenario aspects (triangles).

Figure 2. Marginal Utility Estimates: 108 *you*- and *everyone*-aspects

Notes: Aspect coefficients from benchmark OLS regressions (table 2) for the 108 *you*-aspects (from personal scenarios, x's) and their corresponding 108 *everyone*-aspects (from policy scenarios, triangles), by rank in the personal scenarios. Dashed curve: local linear regression of *everyone*-coefficients (Epanechnikov kernel, bandwidth = 6).