### NBER WORKING PAPER SERIES

### WHAT DO PEOPLE WANT?

Daniel J. Benjamin Kristen B. Cooper Ori Heffetz Miles S. Kimball Tushar Kundu

Working Paper 33846 http://www.nber.org/papers/w33846

## NATIONAL BUREAU OF ECONOMIC RESEARCH 1050 Massachusetts Avenue Cambridge, MA 02138 May 2025

We are grateful for NIH/NIA grants R01-AG065364 to The Hebrew University of Jerusalem, and R01-AG051903 to the University of California Los Angeles; to Dmitry Taubinsky and conference participants at the 2025 ASSA Annual Meeting for helpful comments and discussion; and to Tal Asif, Keren-Or Barashi Gortler, Yehonatan Caspi, Colby Chambers, Arshia Hashemi, Dominic Kassirra, Mattar Klein, Dimitriy Leksanov, Rosie Li, Lev Maresca, Josef McCrum, Tuan Nguyen, Jeffrey Ohl, Shenhav Or, Ofri Piltz, Yonatan Rahimi, Becky Royer, Hannah Solheim, Pierre-Luc Vautrey, Shira Zadik, and Ruthie Zussman for outstanding research assistance. This work utilized the Summit supercomputer and the Alpine high performance computing resource at the University of Colorado Boulder. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or other funding bodies. The authors received IRB approval from the relevant institutions and have no material financial interests that relate to the research described in this paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2025 by Daniel J. Benjamin, Kristen B. Cooper, Ori Heffetz, Miles S. Kimball, and Tushar Kundu. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

What Do People Want? Daniel J. Benjamin, Kristen B. Cooper, Ori Heffetz, Miles S. Kimball, and Tushar Kundu NBER Working Paper No. 33846 May 2025 JEL No. D12, D90, I31

### ABSTRACT

We elicited over a million stated preference choices over 126 dimensions or "aspects" of well-being from a sample of 3,358 respondents on Amazon's Mechanical Turk (MTurk). Our surveys also collected self-reported well-being (SWB) questions about respondents' current levels of the aspects of well-being. From the stated preference data, we estimate relative log marginal utilities per point on our 0-100 response scale for each aspect. We validate these estimates by comparing them to alternative methods for estimating preferences. Our findings provide empirical evidence that both complements and challenges philosophical perspectives on human desires and values. Our results support Aristotelian notions of eudaimonia through family relationships and Maslow's emphasis on basic security needs, yet also suggest that contemporary theories of well-being may overemphasize abstract concepts such as happiness and life satisfaction, while undervaluing concrete aspects such as family well-being, financial security, and health, that respondents place the highest marginal utilities on. We document substantial heterogeneity in preferences across respondents within (but not between) demographic groups, with current SWB levels explaining a significant portion of the variation.

Daniel J. Benjamin University of California, Los Angeles Anderson School of Management Economics Department and NBER daniel.benjamin@anderson.ucla.edu

Kristen B. Cooper Gordon College kristen.cooper@gordon.edu

Ori Heffetz Cornell University Johnson Graduate School of Management and The Hebrew University of Jerusalem and also NBER oh33@cornell.edu Miles S. Kimball University of Colorado Boulder and NBER miles.kimball@colorado.edu

Tushar Kundu Columbia University tk2859@columbia.edu

A data appendix is available at http://www.nber.org/data-appendix/w33846

What do people want? This age-old question has intrigued philosophers, psychologists, and economists alike.<sup>1</sup> Ancient philosophers such as Aristotle emphasized eudaimonia (flourishing) achieved through virtuous living, while Epicureans and Stoics debated the importance of pleasure versus tranquility. Religious thinkers focused on divine connection, while Enlightenment philosophers such as John Locke emphasized liberty and property. The modern literature on self-reported well-being implicitly assumes that people's primary goals are happiness and life satisfaction (Benjamin et al. 2023a). However, a more recent literature has demonstrated that not only will people sacrifice happiness and life-satisfaction for other aspects of life, but they often value other aspects more highly (Benjamin et al. 2012; Benjamin et al. 2014b; Adler, Dolan, and Kavestsos 2017). Understanding what people want is central to understanding how people allocate resources, make decisions, and respond to policy, and it is crucial for designing interventions that improve welfare.

Economists have developed a sophisticated theoretical and empirical apparatus for studying preferences. However, economists often focus exclusively on preferences over market goods, ignoring critical non-market dimensions of well-being, such as family relationships, health, and personal fulfilment. In this paper, we bring to bear tools from economics to study people's preferences over a wide array of both market and non-market goods, with the goal of painting a comprehensive picture of what people want in life.

A simple but important first insight from the economics lens is that what people want will generally depend on what they currently have. For most decision-making and policy purposes, the relevant notion of "what people want" is its (relative) marginal utility. In this paper, we estimate marginal utilities of different aspects of life by surveying respondents about what they would choose when making trade-offs between improvements (or between worsenings) in the aspects. Figure 1 displays an example: the respondent chooses between improving *Your cultures and traditions being honored* or *The happiness of your family*, holding all other aspects fixed. This stated preference method allows us to quantify the relative importance of diverse aspects of well-being, ranging from income and financial security to mental and physical health.<sup>2</sup>

<sup>&</sup>lt;sup>1</sup> Philosophers distinguish between the related concepts of value, desire, preference, well-being, and the good life (i.e., what people should strive for). The concept we focus on in our main analyses is preference, as revealed by hypothetical choices ("stated preferences"), but in the context of our surveys, we generally expect differences between these concepts to be small.

<sup>&</sup>lt;sup>2</sup> Evidence from field experiments supports the validity of stated preference methods for estimating marginal values. List, Sinha, and Taylor (2006) examine choice experiments across two field settings and find that while hypothetical





A second insight is that *any* comparison of value across heterogenous life domains requires a common unit. Consider the relative marginal utilities of watermelons and grapes. The relative marginal utility of one grape will be quite different from the relative marginal utility of a pound of grapes, and whether watermelons and grapes are valued more may depend on whether relative marginal utility is measured per piece of fruit or per pound. For market goods, money metrics—e.g., a dollar's worth of grapes—are a natural unit, but there is no analog for nonmarket goods that have subjective units, such as family relationships. We address this unit problem by conducting our analysis with several different reasonable choices of units. In our main analysis, we adopt respondents' own 0-100 scale as the common unit and compare marginal utilities of aspects of well-being per point. This choice of units relies on individuals' own judgments of how to translate the magnitude of a "point" across aspects of well-being. As an alternative, we examine marginal utilities in units of sample standard deviations. With those

choices may overestimate absolute purchase decisions, they can provide credible estimates when combined with appropriate design features. Crucially for our purposes, Lusk and Schroeder (2004) demonstrate that while hypothetical choices may overestimate total willingness-to-pay, marginal willingness-to-pay, which is our focus, is not statistically different between hypothetical and actual payment settings.

units, we are studying marginal utilities for increasing an individual's placement in the distribution by the same amount. We find that our qualitative conclusions are largely the same for these alternative units because the variation in marginal utilities across aspects of well-being, using any of these measures, is much larger than the effect of changing units.

A third insight from economics is a framework for understanding the variation across aspects' marginal utilities. According to the household-production framework, each respondent produces non-market aspects using resources that may include money but also non-tradeable resources, such as time, cognitive effort, and emotional energy. When the resource cost of producing an additional unit is high, the aspect has low "supply." Since individuals have a low level of the aspect, its marginal utility is high. For example, elderly individuals may have lower levels of physical mobility due to supply constraints, which is why their lower levels are accompanied by *higher* marginal utility. Conversely, when production is inexpensive, supply is high, so the level of the aspect will be high which, for the same preferences tends to make marginal utility low.

On the other hand, given two individuals who face the same constraints—that is, the same supply—one may care more about the aspect. Being willing to sacrifice more for it, that individual would in fact end up sacrificing more to get a higher quantity, with the extra sacrifice being greatest for the last increment. (A higher marginal utility.) This difference in preferences is a difference in "demand"—that is, taste for the aspect.

Thus, differences in levels of aspects of well-being can be due to either differences in supply or differences in demand, or a combination. Differences in supply are associated with marginal utility moving in the opposite direction, while differences in demand are associated with marginal utility moving the same direction. Heterogeneity across individuals, in both levels and marginal utilities (including demographic differences), can be understood through this "supply and demand" framework.<sup>3</sup>

Our sample is over a million tradeoffs we collected from 3,358 respondents on Amazon Mechanical Turk (MTurk) across 126 aspects of well-being. Although large and diverse, our sample is not representative of the US population: they are more likely to be white, be college-

<sup>&</sup>lt;sup>3</sup> Here, we are treating "supply and demand" as a synonym for "constraints and preferences." Supply and demand are already focused on marginal changes, while constraints and preferences imply consequences for marginal changes that can be readily derived.

educated, be middle-aged, have lower incomes, have children, and be actively employed. We go beyond prior work (in particular, Benjamin, Heffetz, Kimball, and Szembrot, 2014; hereafter BHKS) by collecting measures of both levels and relative values for 126 aspects distilled from a much longer list – a list of over 2000 aspects of well-being designed with the primary aim of being comprehensive, from a relatively agnostic stance, rather than relying on predetermined philosophical or theoretical frameworks. We chose aspects of well-being that capture human desires at a more fundamental level than market goods, asking about *ends*, such as financial security or happiness, rather than *means*, such as income or specific consumption choices. Crucially for our analyses, having data on both levels and tradeoffs allows us to look not only at individual differences, but also at the *reasons* for those individual differences.<sup>4</sup>

We find that respondents' marginal utilities are highest for aspects related to family wellbeing, financial stability, and health. These aspects typically have marginal utility estimates 5–20 times larger than the average aspect (the aspect with the largest, *Your children's health*, is an outlier, with an estimate 67 times larger than the average aspect). By contrast, the standard wellbeing measure *How satisfied you are with your life* has a marginal utility only about twice the average. Status signals, prosocial aspects, and global public goods are in the bottom decile. That children's health is at the top, while status concerns are at the bottom, raises the concern that our estimates are driven by social desirability biases. But as we discuss below, additional analyses suggest that the rankings are affected by social desirability concerns only to a moderate, quantifiable extent. The striking difference in marginal utilities highlights a major limitation of focusing solely on life satisfaction as a comprehensive measure of well-being.

We observe substantial heterogeneity in marginal utilities across individuals, yet these differences are far greater *within* demographic groups than *between* them. This finding challenges some psychological theories about preference evolution across the lifespan. Carstensen's socioemotional selectivity theory, for instance, predicts that as time horizons shrink, older adults should prioritize emotionally meaningful experiences such as positive daily affect, stress reduction, and close relationships over future-oriented goals such as status or skillbuilding. Our data reveals a more nuanced pattern. The clearest gap appears for *physical health*, whose relative marginal utility is higher for older respondents, likely reflecting supply

<sup>&</sup>lt;sup>4</sup> BHKS collected no data on respondents' current levels of the well-being aspects. It also used only qualitative scales (e.g., "slightly more", "much more").

constraints as health naturally declines. Surprisingly, aspects such as *feeling calm and peaceful*, *the absence of stress*, and *spending enough time with loved ones* in fact show little variation across age groups. This suggests that the appearance of age-specific preferences may be driven to an important degree by changes in production capabilities (supply), not simply by fundamental shifts in peoples' tastes (demand).

To gain some empirical traction on the reasons for differences in marginal utilities across aspects, we conducted an "aspect flagging survey," where respondents evaluated the characteristics of different aspects, such as breadth, susceptibility to social desirability biases, and the reasons for differences in people's valuation (to quantify the extent of supply-driven versus demand-driven differences). We quantify these characteristics, allowing us to assess supply and demand, as well as test other relationships between aspect characteristics and their marginal utilities. For example, we document a positive relationship between estimated marginal utilities and social desirability scores for aspects, suggesting that social considerations influence stated preferences. However, this relationship explains only 15% of the variation in marginal utility estimates, indicating that while social desirability matters, substantial meaningful variation in preferences across aspects remains.

We also find that the supply-and-demand framework contributes to understanding the empirical relationship between aspect levels and marginal utilities, with aspects respondents find to be supply-constrained showing steeper negative relationships between individual marginal utilities and levels. These findings help clarify whether observed marginal utility heterogeneity reflects variation in tastes or in life circumstances, providing empirical validation for our interpretive framework. The result also has significant implications for how we interpret individual differences in well-being. For aspects where supply differences dominate (such as many health and financial aspects), individuals with lower levels likely face greater constraints in production rather than simply having different demand. This suggests that interventions targeting these constraints could generate substantial welfare improvements. In contrast, for aspects where taste differences dominate, the diversity in levels may primarily reflect differences in individual values rather than inequalities in capabilities or resources.

Our analysis brings empirical evidence to bear on the centuries-old question: *What do people want*? While we find support for some views—for example, Aristotle's emphasis on family and virtue and Maslow and Locke's on security—we find evidence against others—such

6

as the contemporary emphasis on life satisfaction. Taken together, our results provide a much richer picture of people's preferences over the full spectrum of aspects of well-being.

# 1. Survey Design

We utilize data from three surveys; our baseline (henceforth, *Baseline*), follow-up (*Bottomless*), and *Aspect Flagging* survey. The two key types of questions in Baseline and Bottomless are SWB questions, described in Section 1.1, and stated-preference questions (tradeoffs), described in Section 1.2. The survey flow is: (i) consent form; (ii) basic demographic questions (age, gender, household income, ZIP code); (iii) instructions; (iv) SWB and stated-preference questions and (v) calibration questions designed to capture scale-use (not analyzed in this paper) ((iv) and (v) in randomized order); (vi) additional demographic, behavioral, and psychological questions; and (vii) exit questions about how the respondent approached the survey.

## 1.1 SWB Questions

Our SWB questions elicit respondents' ratings of various aspects of life, over the past year, using a slider. Figure 2 shows an example.





Response options are integers from 0 (labeled "Lowest level possible") to 100 (labeled "Highest level possible"). The default slider position is at 50. To give a rating, the respondent moves the slider, and then clicks "Confirm Rating." To prevent lazy default responses, this button appears only after the slider has been moved.

The quasi-continuous 0-100 integer scale makes conclusions less susceptible to untestable assumptions about a latent variable that are needed when there are only a few

response categories (Bond and Lang, 2019).<sup>5</sup> We specify the timeframe of the past year to reduce heterogeneity in interpretation of the question across respondents (Benjamin et al. 2023b). The endpoint labels "Lowest/Highest level possible" result in response options that can be used for SWB questions about any potential aspect, and are meant to sound extreme in order to reduce potential top- and bottom-coding.

### 1.2 Tradeoff Questions

Tradeoff questions ask, "Which option do you think you would choose?" Below the question are two panels. The top of the left panel has a large button labeled "Option 1;" the top of the right panel has a large button labeled "Option 2." Below each option button is an aspect of well-being followed by the word "increases" or "decreases" and a verbal representation of the change (e.g., "+4 from 58 to 62" or "-5 from 68 to 63"). Figure 1 shows an example.

A key innovation of our design is the tight integration between these tradeoff questions and the previously collected SWB levels data. The starting point for each aspect's change is precisely the level that the respondent had reported earlier in their rating question. This integration enables us to analyze how preferences relate to current SWB levels, providing a more accurate measurement of marginal utilities relative to each individual's specific circumstances rather than abstract hypotheticals. We explain how the directions and magnitudes of the changes are determined further below. The representation of the change ("+4" or "-5") appears in a contrasting circle for further emphasis. The final element of each panel is an image of the slider bar from the rating questions. The slider bar is colored up to the level of the starting point for the change (this mimics what the respondent saw at the end of the rating question for this aspect). A darker shade of the same color fills the area between the final and initial levels of the change, and a circle appears around this area. We also emphasize the shaded area of the change by putting a black arrow that points in the direction of the change within the slider bar on that area.

<sup>&</sup>lt;sup>5</sup> While psychological research has traditionally advocated for 5–7-point scales based on respondents' discrimination capabilities, we adopt a quasi-continuous 0-100 scale for several reasons. First, this approach minimizes information loss that occurs when forcing responses into widely-spaced categories (as when response options are separated by a standard deviation). Second, the granular scale reduces ceiling and floor effects by providing ample room at both extremes. Third, although evidence suggests that finer scales may be rated as less quick to use (Preston and Coleman, 2000), they better allow respondents to express nuanced distinctions in their feelings, thereby reducing the need for assumptions about latent variables when analyzing responses. We acknowledge that respondents may engage in rounding behavior on such scales, a limitation that future work should address following approaches like Giustinelli, Manski, and Molinari (2022).

To progress through the survey, a respondent must click on the Option 1 button or the Option 2 button. Below both panels, respondents also have the option of checking a box next to the statement, "This decision does not make sense." If the respondent checks the box, a question appears: "Why does this decision not make sense?" Based on pilot testing, where we had an open text box for replies to this question, the following options are given: The aspect on the left does not apply to my life; The aspect on the right does not apply to my life; Both aspects do not apply to my life; These two aspects of life are too similar; These two aspects of life go hand in hand (you can't change one without the other changing as well); I am indifferent between these two options. Other: [text box].<sup>6</sup>

The SWB questions and tradeoffs are arranged in a "triple" format, with three SWB questions (for aspects A, B, C) followed by six tradeoff questions between those aspects (A vs B, B vs A, B vs C, C vs B, A vs C, C vs A, in randomized order).

The magnitudes of the changes in levels depicted in the tradeoffs range from 1 to 8. The changes are chosen as follows:

- 1. For a given triple, the changes in the tradeoff questions all go in the same direction, and that direction is randomized with 50% probability.<sup>7</sup>
- 2. At the level of each pairwise trade-off, there is a 2% probability that the aspects move in opposite directions. We use this "opposite direction trap" as a data quality check.
- 3. For each aspect in a tradeoff, the magnitude of the change is randomly chosen from integers 1 to 8, with equal probability.

<sup>&</sup>lt;sup>6</sup> In pilot testing, we found that respondents who checked "This decision does not make sense" had two main concerns: some wanted to express that they did not prefer either option over the status quo (plausible for about half of all tradeoffs since scenarios show both aspects decreasing with 50% probability), while others had specific reasons why the decision framework didn't apply to them. We captured the latter in multiple-choice options based on open-text responses. To address the former in the Baseline survey, we added a note below the "Why does this decision not make sense?" question stating: "Note: If you dislike both options, please choose the option you dislike the least." Respondents can progress to the next screen either by selecting one of the two options, or by clicking the "submit" button after checking the box.

<sup>&</sup>lt;sup>7</sup> There are some exceptions. If a respondent has rated one of the aspects in the tradeoff below 8 or above 92, it would be possible for the realized value of the tradeoff to put the rating below 0 or above 100, respectively. To avoid this outcome, if one of the three aspects in the triple was rated above 92, we "force" the direction of the changes in the tradeoffs to be negative (i.e., we do not randomize the direction of changes for the triple); if one of the three aspects was rated below 8, we "force" the direction of tradeoffs to be positive. If one of the aspects was rated below 8 *and* another one of the aspects was rated above 92, we skip the tradeoff questions for that triple. We do not tell respondents about this feature of the survey. In our data for X respondents covering Y triples, only Z% of triples were skipped.

Note that in the practice questions (in the Baseline survey instructions), the tradeoffs are not randomized in the same way as the real survey. The magnitudes of the tradeoffs are always 4 and 5; we randomize which aspect has the bigger change. The first example tradeoff (between X and Y) shows decreases of -4 and -5. The second and third example tradeoffs (between X and Y and Z and A, respectively) show increases of +4 and +5. Exceptions can arise if a respondent rates an aspect below 8 or above 92 in the practice questions; if both aspects in a practice tradeoff question were rated below 8 (above 92), the relevant tradeoff question is always shown as an increase (decrease). If a respondent has rated one aspect below 8 and the other above 92, the practice tradeoff screen says, "There is no practice decision about the two aspects you just rated since your ratings were extraordinarily high and extraordinarily low."

### 1.3 Baseline and Bottomless Surveys

Using Amazon's Mechanical Turk (MTurk) platform, we collected data between June 13 and December 7, 2022.<sup>8</sup> We restricted eligibility for all surveys to MTurk workers located in the United States (as identified by MTurk), with a HIT approval rating of at least 95%, and at least 100 previously approved HITs. MTurk restricts all participation to workers who are at least 18 years of age. Respondents were recruited to Baseline with a HIT titled "Academic survey about what is important in life," after passing an initial prescreening survey. Respondents were compensated \$4.50 for completing Baseline. For those who passed our stringent quality control checks, the median completion time was 36 minutes.

After completing Baseline, respondents who passed quality control were invited to a follow-up survey (Bottomless). Bottomless features the same SWB and Tradeoff questions for additional aspects of well-being. We organize Bottomless into "blocks", each one featuring questions about a set of 33 aspects. Within a block, the aspects' order (i.e., assignment to triples within the block) is always randomized. The aspect list for regular blocks always includes the 16 aspects from which we plan to construct a welfare measure in a forthcoming paper<sup>9</sup>. The

<sup>&</sup>lt;sup>8</sup> Note that ChatGPT was launched on November 30, 2022, so while LLM-generated responses are currently a major issue for online surveys, they were not at the time of our data collection.

<sup>&</sup>lt;sup>9</sup> Including the aspects for the welfare measure in every block achieves two aims for our other work which estimates the aspects' relative marginal utilities. First, this design allows us to compare the relative MU's within a list that includes all the Bottomless survey aspects. The tradeoffs where aspects in the welfare measure are compared to other aspects create the "interlocking" triples required to do this. Second, for other work, we want to be able to estimate individual-level marginal utilities for these aspects, to test the robustness of our welfare measure to individual-level prices. We will not be able to estimate individual prices for all respondents because of the high data

remaining 17 aspects in each block have been chosen for a variety of reasons, such as finding a more comprehensive set of aspects and testing hypotheses about the wording of aspects.

In Block 2, we tried to develop an aspect list which with maximal "comprehensiveness." Based on pilot data for 2000+ aspects of well-being, we chose the 17 non-welfare-measure aspects for Block 2 with a stepwise procedure; we maximized an R<sup>2</sup> criterion with marginal utility weights to see which of the other aspects added the most explanatory power to the 33 aspects already in the Baseline list. We also did the procedure with marginal utility-squared weights and without marginal utility weights; the aspects identified with the alternate procedures which did not already appear in the Bottomless survey are included in later Blocks (6-8).

In the other Blocks, we included other aspects of well-being from the pilot survey's long list of aspects, for various reasons. The following list of rationales corresponds to aspects in Blocks 3 through 12. For the first 12 blocks of the Bottomless survey, every respondent sees the same 33 aspects in that block. (1) Two aspects are included to study aspects of well-being with low importance; they were identified by low marginal utilities in pilot data. (2) Fifteen aspects are included because we hypothesize that they are closely related to choices; we plan to study these in another project studying the multidimensionality of well-being. (3) Four aspects are included because they reflect local public goods; their ratings can be compared with external data. (4) Three aspects are included because they correspond to questions we ask about expected future aspect levels; we plan to study these for studying theories about the shape of well-being over the life course. (5) We include 35 aspects because they correspond, as nearly as possible, to questions asked by major SWB surveys (Gallup and statistical agencies). (6) We include 22 aspects as variant wording of Baseline aspects, to test the sensitivity of small wording changes. (7) We include 13 aspects because they correspond to the aspects used for personal calibration questions. This is another source of data for exploring general vs. specific scale use. (8) We include 53 aspects which reflect extremes in different categories – such as social desirability, ease of purchasing with money, or "objectivity" – based on another preliminary pilot survey.

Across the remaining 17 regular Blocks, the order of the 289 aspects is randomized. These aspects were also chosen from the long pilot list for various reasons. We include 7 aspects

requirement, but it is feasible for respondents who complete a lot of Bottomless blocks. Across all surveys, roughly 1/4 of tradeoffs are between two aspects within the welfare measure list; we will pool these tradeoffs at the individual level, across all available blocks, to estimate the individual-level prices.

related to political efficacy and life satisfaction which correspond as closely as possible to the vignette literature. We include 66 aspects which are considered local public goods, in order to further probe potential correlations with external "objective" measures. We include 47 aspects which are considered national or global public goods; from among all such aspects in the long list, we looked for the best possible ones to use as "quasi-calibration questions": hypothesized to have minimal interpersonal differences in definition and perception (for example, trying to avoid aspects connected to political ideology). We include 82 aspects with simple and compound forms – such as *Your feeling of independence, Your feeling of self-sufficiency*, and *Your feeling of independence – to test hypotheses about wording and conceptual overlap between aspects.* We also include 14 aspects which we wrote since the pilot to test conceptual overlap related to trust. Finally, we include 130 new aspects – developed since the pilot – which are hypothesized to have high marginal utilities either directly or indirectly (through having a high marginal product), and 10 new aspects which are hypothesized to be highly predictive of other aspect levels.

### 1.4 Aspect Selection

Out of the full list of aspects on Baseline and Bottomless, we focus on 126, chosen for the following reasons:

- 1. Commonly used and studied in the subjective well-being literature (e.g., *How satisfied you are with your life*)
- 2. Collected by large statistical agencies (e.g., *How happy you feel*)
- 3. Those selected to provide comprehensive coverage of anything one may care about (e.g., *Your financial security*)
- 4. Represent aspects people regularly trade off with other aspects of life in common decisions (e.g., *You getting enough sleep*)

While we only report estimates for the 126 aspects chosen based on these criteria, our methods can in theory be applied to the full set of aspects from Baseline and Bottomless, and ignoring computational constraints, on our full list of 2000+ aspects. Our final sample contains respondents who completed SWB and tradeoff questions for all 126 aspects.

### 1.5 Sample

In total, 5,970 respondents completed Baseline. Table 1 summarizes the demographics among those who answered the demographic questions used in our main applications (5,466 respondents) and among those who also passed quality control (3,358 respondents). Though similar to the U.S. population (according to the 2020 Census) on most demographics, as in other MTurk studies, our sample differs from the U.S. population in several ways: respondents are more likely to have completed college, be younger than 50, and be unemployed (other than MTurk); and less likely to have annual household income above \$120,000 and to be Black or Hispanic/Latino.

For our main study sample, we further narrowed the respondents to include only those who answered tradeoffs in the first nine blocks. This ensures that every respondent in this sample provided a marginal utility estimate for our 126 aspects of interest. By fixing the study sample in this manner, we can more confidently analyze heterogeneity in preferences and make robust comparisons across the complete set of aspects. The resulting subsample consisted of 896 respondents, whose demographic shares are nearly identical to those of the broader Baseline passed quality control sample.

Category	Population	Baseline, 2022 Sample, (N=5466)	Baseline, 2022 Sample, Passed QC (N=3358)	Bottomless, 2022 Sample, Blocks 1-9 (N=896)	Census
Gender	Male	49.6	43.5	45.6	49.0
	Female	50.4	56.5	54.4	51.0
Marital status	Married	65.7	50.4	52.0	51.1
	Never married	24.6	35.4	34.2	33.5
	Other	9.6	14.3	13.8	15.4
Highest education level completed	High school grad	7.9	10.4	10.8	28.3
	Some college	16.6	25.7	24.1	27.1
	Bachelor's degree	52.6	46.2	48.0	22.2
	Graduate degree	22.6	17.2	16.9	12.8
Age	18–29	18.7	14.9	12.2	20.6
	30–39	35.3	33.6	32.8	17.5
	40-49	21.6	23.4	25.3	16.0
	50-64	19.2	20.7	24.3	24.9
	65 and older	5.2	7.4	5.4	21.1
Household Income	Less than \$30,000	14.2	17.3	15.6	22.1
	\$30,000-\$49,999	21.7	21.1	23.7	15.7
	\$50,000-\$69,999	22.2	20.1	19.5	13.4
	\$70,000-\$89,999	18.4	15.4	15.0	10.8
	\$90,000-\$119,999	13.5	12.8	12.6	12.0
	\$120,000 and above	10.0	13.3	13.6	26.0
Region	Midwest	22.8	22.9	22.9	20.8
	Northeast	16.6	18.1	19.9	17.4
	South	38.0	38.5	37.6	38.1
	West	22.6	20.5	19.6	23.7
Race	White, and other	79.1	78.6	80.1	63.3
	Black	6.8	8.1	6.9	12.1
	Hispanic/Latino	9.8	7.3	6.8	18.7
	Asian	4.3	6.0	6.1	5.9
Household size	1	15.6	20.7	21.0	28.5
	2	27.9	30.3	28.5	35.0
	3	19.0	19.5	20.3	15.0
	4 and above	37.6	29.5	30.2	21.5
Employment status	Employed	86.6	79.7	80.0	57.6
	Unemployed	6.4	10.0	9.8	3.8
	Not in labor force	7.0	10.3	10.2	38.5

# Table 1 - Demographics

Notes: All numbers are percentages. 'Passed QC' refers to everyone who passed quality control and 'Blocks 1-9' refers to everyone who completed tradeoffs in each of blocks 1 to 9. Sources: Authors' surveys, 2020 Census.

# 1.6 Aspect Flagging Survey

The *Aspect Flagging* survey aims to understand how respondents characterize various aspects of life. By eliciting these characteristics, we are better able to understand what aspect features drive the marginal utility estimates we derive from the tradeoff data in Baseline and Bottomless. This survey was administered to all MTurk respondents who completed *Baseline* and passed quality control, with a final sample of 1088 respondents. Respondents rated aspects on various characteristics (which we call "flags") using a 9-point integer scale. As shown in Figure 3, each flag contained a headline characteristic and detailed question text, followed by the list of aspects to be rated.



Figure 3 – Annotated Screenshot of Aspect Flagging Survey

To make the survey manageable, we divided 172 aspects (which includes all 126 of our aspects of interest) into four separate surveys of 43 aspects each. These were fielded across multiple waves, with each wave containing 7 flags. We fielded Waves 2-4 (Wave 1 was a pilot wave) between September 2024 and January 2025 to our sample of respondents. Median completion times ranged between 26 and 35 minutes. Each flag included two additional "calibrator" aspects designed to represent very low and high scores for that specific question.

For quality control, we dropped responses from duplicate survey takes and removed respondents who failed attention checks or "straight-lined" their responses (gave identical ratings for all aspects on a flag). After cleaning, 86.1% to 93.8% of complete responses were retained across surveys.

To account for scale-use heterogeneity, we distinguished between two types of flags: "calibration-question-like" flags, and "aspect-like" flags. Calibration-question (CQ)-like flags are designed such that ratings for these flags are assumed to be unimpacted by respondents' own levels of these aspects. An example would be the flag designed to capture social desirability concerns: "Would someone downplay how much they want this aspect because openly wanting it might make them look bad?" Aspect-like flags, on the other hand, are designed such that a respondent's own aspect level can impact their rating of that aspect for that flag. An example would be the flag we designed to capture marginal value: "Think about your own life as it is. Given how much you have of this aspect, how eager would you be to get more? In particular, compared to other aspects of life, how valuable would it be to have a little bit more of this aspect?".

We standardize ratings within person, only using responses to "CQ-like" flags (20 out of 28 flags) and not "aspect-like" flags. Given that the underlying truth is assumed to be the same across people for a "CQ-like" flag, differences in responses can be attributed to how individuals use the scale. Therefore, responses to "CQ-like" flags are useful for scale-use correction, while responses to "aspect-like" flags, which capture both scale-use heterogeneity and an individual's situation, are not. For each person *i*, their standardized rating *r* for aspect *a* and flag *f* was calculated as:

$$\widehat{r_{afi}} = \frac{r_{afi} - E_{af}[r_{afi}]}{\sqrt{Var_{af}[r_{afi}]}}$$

where expectations and variances were taken across all ratings for CQ-like flags. The resulting standardized means across respondents help characterize how aspects score on various aspects while controlling for individual scale-use differences.

# 2. Theoretical framework

In this section, we introduce our theoretical framework and underlying assumptions, and introduce our econometric model.

### 2.1 Binary choice model

Denote  $w_{ij}$  as individual *i*'s level of aspect *j*, and  $w_i$  as the vector of levels for each individual. Let  $u_i(w_i)$  be individual *i*'s utility function so that changes in utility that result from a change in the level of aspect *j* for individual *i*,  $\Delta w_{ij}$ , can be expressed as

$$\frac{\Delta u_i}{\Delta w_{ij}} = \frac{u_i (w_{ij} + \Delta w_{ij}, \boldsymbol{w}_{i,-j}) - u_i (w_{ij}, \boldsymbol{w}_{i,-j})}{\Delta w_{ij}}$$

Let  $M_{ij}$  denote MU ( $m_{ij}$  for log MU). Then for small changes  $\Delta w_{ij}$ , we have that

$$\frac{\Delta u_i}{\Delta w_{ij}} \approx \frac{du_i}{dw_i} = M_{ij} \Rightarrow$$
$$u_i (w_{ij} + \Delta w_{ij}, \boldsymbol{w}_{i,-j}) - u_i (w_{ij}, \boldsymbol{w}_{i,-j}) \approx M_{ij} \Delta w_{ij}$$

For small changes, the change in utility induced by a change  $\Delta w_{ij}$  is, by a first-order approximation, the marginal utility,  $M_{ij}$ , times the amount of the change. In order to estimate this model, we ask hypothetical choice questions (*tradeoffs*), as described in further detail in Section 1.2. Tradeoff questions ask respondents to choose between two options: a (small) increase in aspect *j*,  $\Delta w_{ij}$  and an increase in a different aspect *j*',  $\Delta w_{ij'}$ . The amount of each change is randomized. We assume that a respondent chooses the option changing aspect *j* if and only if  $M_{ij}\Delta w_{ij} > M_{ij'}\Delta w_{ij'}$ . Taking logs and rearranging yields

$$m_{ij} - m_{ij'} + \ln \left(\Delta w_{ij}\right) - \ln(\Delta w_{ij'}) > 0$$

and adding in response error we have

$$m_{ij} - m_{ij'} + \ln\left(\frac{\Delta w_{ij}}{\Delta w_{ij'}}\right) + e_{ijj'q} > 0$$

where q is a specific tradeoff question.

### 2.2 Hierarchical model

To study the heterogeneity of preferences, we add additional structure where individual draws of log marginal utilities are drawn according to a hierarchical process. Let  $\mu_i$  be hyperparameters

representing the mean log marginal utility for each aspect, *j*. Then individual log marginal utilities are drawn:

$$m_{ij} \sim N(\mu_j, \sigma_j),$$

where  $\sigma_j$  varies by aspect. Since the scale of marginal utilities is not identified, we impose the normalization that the average across log marginal utility means,  $\bar{\mu}$ , equals 0. Therefore, all log marginal utility estimates can be interpreted as being relative to the average aspect.

We model individual response error as being drawn from a log-normal distribution:

$$\sigma_{e_i} \sim \text{lognormal}(\mu_e, \sigma_e),$$

where  $\sigma_{e_i}$  is the standard deviation of response error,  $e_{ijj'q}$ , for individual *i*. Note that by allowing the variance of the error term to vary by individual, our estimates of aspect hyperparameters (the mean and standard deviation of log marginal utility draws) will weight more heavily responses from respondents who make more consistent choices.

Under this model, the likelihood of observing any given binary choice, conditional on  $\sigma_{e_i}$ , is given as

$$\mathbb{P}[r_{ijj'} = 1 | \Delta w_{ij}, \Delta w_{ij'}, m_{ij}, m_{ij'}, \sigma_{e_i}] = \Phi\left[\frac{m_{ij} - m_{ij'} + \ln\left(\frac{\Delta w_{ij}}{\Delta w_{ij'}}\right)}{\sigma_{e_i}}\right]$$

where  $r_{ijj'}$  is an indicator for choosing the option associated with increasing aspect *j* and  $\Phi[\cdot]$  is the standard normal cumulative distribution function.

We estimate the model using Hamiltonian Markov Chain Monte Carlo (HMC). Since this estimation method imposes a Bayesian framework, we follow standard practice (Gelman, 2006) and specify uninformative priors for our hyperparameters

$$\mu_j \sim N(0, 10)$$
  

$$\sigma_j \sim \text{cauchy}(0, 2)$$
  

$$\mu_e \sim N(0, 10)$$
  

$$\sigma_e \sim \text{cauchy}(0, 2).$$

One possible concern in estimating marginal utilities for individual aspects is the potential for overlap between aspects. Some aspects may represent combinations of underlying components, much like a market good might actually be a bundle of goods: 2 bananas and 4 oranges or 1 orange and 5 bananas. Importantly, our method does not rely on the assumption that aspects are strictly non-overlapping or mutually exclusive. Instead, the marginal utility we

estimate for a given aspect reflects the aggregate value across its individual components. Even if aspects share common elements, their contributions will be appropriately captured in the overall valuation as long as a one-point change in an aspect is interpreted by respondents as *only* changing that aspect, holding everything else constant. In the market goods analogy, we are estimating the marginal utilities of the bundles, not the marginal utilities of oranges or bananas.

# 3. Marginal Utility Estimates for Aspects of Well-Being

### 3.1 Aspects People Want Most

Our main estimates reveal the relative importance that individuals place on different aspects of well-being. For each aspect *j*, we estimate  $\mu_j$ , which represents the population mean of log marginal utilities across individuals. Given our normalization that the average of these means across aspects equals zero,  $M_j = e^{\mu_j}$  can be interpreted as the relative marginal utility of aspect *j* compared to the geometric mean across all aspects. Table 2 presents these estimates for all 126 aspects ordered from highest to lowest mean log marginal utilities, along with our subjective grouping of aspects in column 7 that is intended to help summarize broad patterns in the findings.<sup>10</sup>

The aspects with the highest marginal utilities cluster around three key domains: family well-being, financial security, and health (both mental and physical). Family-related aspects rank particularly high, with the top ten aspects involving children, spouses, or loved ones, a finding that aligns with Aristotle's claim that flourishing is lived 'among friends and family' (Aristotle, ca. 350 BCE) and more contemporary surveys such as Pew's finding that family is today's prime source of meaning (Pew Research Center, 2021). The top aspect, "Your children's health", is an outlier, with an estimated relative marginal utility approximately 67 times larger than the average aspect. This means that a one-point improvement on our 0-100 scale for "Your children's health" is valued as much as a 67-point improvement for an average aspect.

Importantly, for aspects that pertain to specific relationships (such as those involving children or partners), we treat responses differently based on respondent circumstances. When estimating relative marginal utilities reported in our main results, we only include tradeoffs from respondents for whom these aspects are directly relevant (e.g., only parents' responses for child-related aspects). We accomplish this by modeling these as totally distinct aspects for respondents without the relevant relationship, effectively allowing the model to estimate separate parameters

<sup>&</sup>lt;sup>10</sup> These groupings should be taken with a grain of salt; they represent one imperfect way to organize the aspects and many could reasonably be classified differently. For example, *Your ability to protect your loved ones* is categorized as "Financial" but could equally belong to "Relationships" or "Family Well-Being." Readers who find these groupings distracting should feel free to ignore them and focus on the individual aspects and their rankings, as our qualitative conclusions remain unchanged regardless of the specific categorization choices. We discuss these groupings in more detail in Section 3.2.

that we subsequently exclude from our main reported findings. This approach ensures our primary estimates reflect valuations from those who actually experience these aspects in their lives, while still maintaining the full structure of our hierarchical model.

# Table 2 – Relative Marginal Utility Estimates

(1) Aspect	(2) <i>M</i> <sub>j</sub> (LB - UB)	(3) Rank (Mj)	(4) $\mathcal{M}_{j}$ per SD (LB - UB)	(5) Rank ( <b>M</b> j per SD)	(6) e <sup>σ</sup> j	(7) Grouping
Your children's health	66.9 (42.4 - 115.5)	1	1731.5 (1098 - 2990.5)	1	25.02	Family Well-being
Your children being able to live happy lives	18.6 (13.7 - 26.3)	2	461.7 (340.3 - 650.5)	2	10.94	Family Well-being
Your spouse/partner's health	16.2 (11.5 - 23.8)	3	400.7 (284.1 - 587.8)	3	17.03	Family Well-being
You having a strong bond with your children	10.8 (7.8 - 15.8)	4	291.6 (208.8 - 424.4)	4	12.40	Relationships
The health of your parents and siblings	8.6 (6.8 - 11)	5	185.8 (146 - 236.1)	5	13.26	Family Well-being
The people who have been most important to you in your life still being alive	8 (6.4 - 10.1)	6	176.7 (140.9 - 222.9)	6	12.31	Relationships
The overall well-being of your family	7.2 (5.9 - 8.6)	7	119.1 (98.9 - 142.8)	9	6.31	Family Well-being
Your family living a wonderful life	5.8 (5 - 6.8)	8	119.3 (102.4 - 139.9)	8	5.01	Family Well-being
The hanning a conternal me	5.8 (5 - 6.8)	9	108 (93.2 - 125.3)	13	4.36	Summum Bonum
Your ability to protect your loved ones	56(46-68)	10	107.9 (89 - 130.2)	14	7.81	Financial
The overall well-being of you and your family	5.6 (4.8 - 6.7)	11	100.9 (85.7 - 119.8)	16	5.33	Summum Bonum
Vour beelth	5.0 (4.8 - 0.7)	12	00.7 (85.1 - 116.0)	17	5.55	Health
The base of some families	5 (4.5 - 5.9)	12	99.7 (85.1 - 110.9)	17	5.05	Freedul
You being the to support your family forentially	4.9 (4.4 - 5.5)	15	124.1 (111.5 120.1)	7	4.52	Family well-being
You being able to support your family financially	4.0 (4.1 - 5.1)	14	124.1 (111.5 - 139.1)	/	4.40	Financial
Your financial security	4.3 (3.7 - 5.2)	15	117.9 (100.2 - 141.5)	10	5.33	Financial
You having enough money to pay for healthcare and/or medicines that you or your family need	4.2 (3.5 - 5.1)	16	109.8 (91.8 - 133.5)	12	7.83	Financial
Your ability to take care of your family	4 (3.5 - 4.6)	17	90.5 (78.4 - 104.4)	18	4.79	Financial
You not worrying about money	3.8 (3.3 - 4.5)	18	112.5 (97.7 - 130.7)	11	4.50	Financial
You having enough money to buy food that you or your family need	3.8 (3.2 - 4.6)	19	81.4 (68.2 - 98.5)	22	7.99	Financial
Your physical health	3.7 (3.3 - 4.1)	20	72 (64.9 - 79.9)	26	4.18	Health
Your children being kind and caring about other people	3.6 (2.9 - 4.6)	21	86.3 (68.2 - 110.8)	20	7.35	Family Well-being
You having enough financial resources for your retirement years	3.3 (2.8 - 3.8)	22	104 (89.8 - 121.4)	15	5.19	Financial
Your mental health	3.2 (3 - 3.5)	23	78.1 (71.7 - 85)	23	3.17	Mental Health
Your relationship with your spouse, partner, or closest friend being stronger than ever	3.2 (2.7 - 3.9)	24	82.8 (70.1 - 99.4)	21	4.70	Relationships
You having enough material assets and resources for the future	2.9 (2.6 - 3.3)	25	77.1 (67.6 - 88.2)	24	4.03	Financial
You and your family having enough to eat	2.9 (2.3 - 3.6)	26	49.9 (40.1 - 62.1)	36	14.65	Financial
You having enough money to do everything you want to do	2.7 (2.3 - 3.1)	27	74.5 (64.8 - 86.1)	25	4.73	Financial
You not having to worry about where your family's next meal will come from	2.6 (2.2 - 3.1)	28	57.8 (48.5 - 69.4)	31	8.05	Financial
Your children growing up to follow the golden rule— treating others as they would like to be treated	2.5 (2 - 3.2)	29	59.1 (46.9 - 75)	30	7.34	Family Well-being
The quality of your spouse/partner as a mate	2.5 (2.1 - 3)	30	70.6 (59 - 84.9)	27	5.18	Relationships
You not having depression	2.3 (2 - 2.6)	31	64.8 (56.6 - 73.9)	28	4.26	Mental Health
You feeling that you have enough money for the things that are most important to you	2.3 (2 - 2.6)	32	61.7 (54 - 69.8)	29	3.69	Financial
Your freedom from physical pain	2.1 (1.8 - 2.4)	33	47.1 (41.1 - 54)	37	4.98	Health
The absence of worry in your life	2.1 (1.9 - 2.3)	34	57 (50.8 - 64.3)	32	3.11	Feelings
How happy you feel	2(19-21)	35	45.6 (42.9 - 48.4)	38	2.13	Feelings
How satisfied you are with your life	2(1.9 - 2.1)	36	45.4 (43.2 - 47.8)	30	1.84	Summum Bonum
How much you ariov your life	2(1.9 - 2.1)	37	42.6 (40.4 - 45)	41	1.04	Feelings
You not feeling downhearted and depressed	19(17-21)	39	50 8 (45 7 - 56 4)	35	2.56	Feelings
The absence of stress in your life	1.9 (1.6 - 2.1)	20	51.1 (45.2 - 57.9)	34	2.30	Time
Your rating of your life on a lader where the lowest rung is "worst possible 16.6 for rough and the highest margin "hart possible life for rough"	1.8 (1.7 - 2)	40	41.2 (38.4 - 44.2)	43	2.59	Summum Bonum
The surflive of users remention relationships, manifest have life as sen life	19(15 21)	41	54 5 (45 2 (57)	22	0.71	Deletionshine
The quality of your romantic relationships, marriage, love life of sex life	1.8 (1.5 - 2.1)	41	54.5 (45.2 - 65.7)	33	8.71	Relationships
You enjoying every day	1.7 (1.5 - 1.8)	42	37.6 (34.3 - 41.5)	46	2.16	Feelings
You getting the health care you need	1.7 (1.4 - 1.9)	43	37.8 (32.3 - 44.1)	44	5.22	Financial
How high your income is compared to the income of other people around you	1.6 (1.4 - 2)	44	42.1 (35.2 - 50.4)	42	7.65	Status
You not feeling anxious	1.6 (1.5 - 1.7)	45	43.5 (39.8 - 47.5)	40	3.52	Feelings
Your physical safety and security	1.6 (1.4 - 1.9)	46	25.2 (21.8 - 29.4)	56	5.41	Local Public Goods
The absence of sadness in your life	1.5 (1.4 - 1.6)	47	37.8 (35.2 - 40.5)	45	2.54	Feelings
You feeling calm and peaceful	1.4 (1.3 - 1.6)	48	31.3 (28.2 - 34.8)	48	2.64	Feelings
Your sense of control over your life	1.3 (1.2 - 1.3)	49	28.3 (26.6 - 30)	50	2.29	Meaning and Fulfillment
You having hope	1.3 (1.1 - 1.4)	50	30.3 (26.8 - 34.2)	49	3.24	Meaning and Fulfillment

You feeling well-rested	1.2 (1 - 1.3)	51	27.3 (23.9 - 31.2)	52	3.95	Health
You liking what you do every day	1.1 (1 - 1.3)	52	25.3 (22.9 - 27.9)	55	2.22	Meaning and Fulfillment
You having people you can turn to in time of need	1.1 (1 - 1.3)	53	27 (23.2 - 31.1)	53	4.98	Relationships
You reaching most of your goals over the past 12 months	1.1 (1 - 1.3)	54	27.5 (24.2 - 31.5)	51	3.69	Meaning and Fulfillment
You having many options and possibilities in your life and the freedom to choose among them	1.1 (1 - 1.2)	55	25.4 (22.3 - 28.7)	54	3.47	Meaning and Fulfillment
You not having to worry about being unemployed	1.1 (0.9 - 1.3)	56	32.4 (27.5 - 37.7)	47	5.94	Financial
You feeling active and productive every day	1.1 (1 - 1.2)	57	24.4 (22 - 27.1)	60	2.50	Meaning and Fulfillment
You having a lot of energy	1.1 (0.9 - 1.2)	58	25.2 (22.2 - 28.4)	57	3.31	Health
Women being treated fairly in your nation	1.1 (0.9 - 1.3)	59	23.1 (18.8 - 28.6)	63	12.70	Non-local Public Goods
Your life being more than just getting things done	1 (0.9 - 1.2)	60	24.7 (22 - 27.9)	59	3.01	Meaning and Fulfillment
The extent to which you feel the things you do in your life are worthwhile	1 (1 - 1.1)	61	22.6 (21.4 - 23.8)	64	1.89	Meaning and Fulfillment
Your knowledge and skills	1 (0.9 - 1.2)	62	16.4 (14.4 - 18.9)	78	4.30	Being Learned
Your sense of purpose	1 (1 - 1.1)	63	24.3 (22.7 - 26)	61	2.42	Meaning and Fulfillment
You not feeling tired all the time	1 (0.9 - 1.1)	64	25 (22 - 28.4)	58	3.95	Health
Your living environment not being spoiled by crime and violence	1 (0.8 - 1.2)	65	20.3 (17.2 - 24)	69	6.31	Local Public Goods
You getting enough sleep	1 (0.8 - 1.1)	66	23.6 (20.2 - 27.6)	62	5.30	Health
How much of the time the things you do seem meaningful	1 (0.9 - 1.1)	67	21.9 (19.6 - 24.4)	66	2.71	Meaning and Fulfillment
You being a good person	0.9 (0.9 - 1)	68	14.8 (13.4 - 16.2)	81	3.52	Prosocial Virtues
You feeling that you have enough time for the things that are most important to you	0.9 (0.9 - 1)	69	18.4 (17.3 - 19.7)	72	2.34	Time
You being able to turn your good intentions into actions	0.9 (0.8 - 1)	70	17.9 (15.7 - 20.2)	74	3.40	Prosocial Virtues
How satisfied you are with your standard of living compared to the people you spend time with	0.8 (0.8 - 1)	71	18.9 (16.7 - 21.3)	70	3.19	Status
How often you smile or laugh	0.8 (0.7 - 1)	72	18.4 (16 - 21.1)	73	3.86	Feelings
You being a winner in life	0.8 (0.7 - 1)	73	21.6 (18.9 - 24.9)	67	4.33	Status
Your spouse/partner taking on their fair share of responsibility for the household	0.8 (0.7 - 1)	74	22.2 (18.3 - 26.6)	65	5.21	Relationships
Your home being comfortable	0.8 (0.7 - 0.9)	75	15.5 (13.6 - 17.7)	80	3.87	Financial
The wisdom you have been able to accumulate	0.8 (0.7 - 0.9)	76	13.6 (12 - 15.6)	86	3.65	Being Learned
You not being lonely	0.8 (0.7 - 0.9)	77	20.4 (18.5 - 22.4)	68	3.65	Feelings
Your friends and family giving you positive energy every day	0.8 (0.7 - 0.9)	78	17.2 (15.2 - 19.6)	77	3.73	Relationships
You taking care of others instead of just yourself	0.8 (0.6 - 0.9)	79	16.2 (13.9 - 18.6)	79	4.62	Prosocial Virtues
The absence of anger in your life	0.8 (0.7 - 0.9)	80	18.7 (16.3 - 21.5)	71	4.20	Feelings
You having people in your life who see the good in you	0.7 (0.6 - 0.8)	81	13.9 (12.1 - 15.7)	83	3.60	Relationships
Your ability to climb a flight of stairs	0.7 (0.6 - 0.9)	82	13.7 (11.3 - 16.5)	84	8.69	Health
You not being bothered by having little interest or pleasure in doing things	0.7 (0.6 - 0.8)	83	17.9 (15.8 - 20.2)	75	3.14	Mental Health
You being a good friend and doing your duty by your friends	0.7 (0.6 - 0.8)	84	13 (11.2 - 14.9)	89	4.20	Prosocial Virtues
You learning or doing something interesting every day	0.6 (0.6 - 0.7)	85	13.7 (12 - 15.7)	85	3.82	Meaning and Fulfillment
You not having things stolen from you	0.6 (0.5 - 0.8)	86	11.3 (9.6 - 13.3)	94	5.75	Local Public Goods
You feeling like you have work-life balance	0.6 (0.5 - 0.7)	87	13.6 (12 - 15.4)	87	3.08	Time
You always making time for regular trips or vacations with friends and family	0.6 (0.5 - 0.7)	88	17.5 (14.9 - 20.5)	76	5.72	Time
How much fun you have when you get together with friends	0.6 (0.5 - 0.7)	89	11.2 (9.8 - 12.8)	95	3.84	Relationships
You finding meaning in the world	0.6 (0.5 - 0.7)	90	13.4 (11.8 - 15.3)	88	3.22	Meaning and Fulfillment
You being satisfied with your job	0.6 (0.5 - 0.7)	91	14.7 (12.7 - 16.8)	82	4.05	Work
You getting to use your strengths to do what you do best every day	0.5 (0.5 - 0.6)	92	11.5 (10.1 - 13)	92	3.19	Meaning and Fulfillment
You having a life outside of work	0.5 (0.5 - 0.6)	93	11.9 (10.4 - 13.7)	91	3.88	Time
How much you like your home (house or apartment)	0.5 (0.4 - 0.6)	94	11.2 (9.7 - 12.9)	96	4.41	Financial
You showing people the right way to live by your words and actions	0.5 (0.4 - 0.6)	95	9.6 (8.3 - 11.1)	101	4.53	Status
The ability of ordinary citizens to influence your national government	0.5 (0.4 - 0.6)	96	12.2 (9.7 - 14.9)	90	11.33	Non-local Public Goods
The overall quality of your experience at work	0.5 (0.4 - 0.6)	97	10.5 (9 - 12.1)	99	4.18	Work
You being able to spend a lot of time with your friends	0.5 (0.4 - 0.5)	98	11.4 (9.7 - 13.3)	93	5.11	Time
The air in your area not being polluted	0.5 (0.4 - 0.5)	99	9.8 (8 - 11.8)	100	8.23	Local Public Goods
You having the frequent opportunity to see things of beauty	0.4 (0.4 - 0.5)	100	8.2 (7.2 - 9.4)	103	3.91	Being Learned

You living in a city or area that is perfect for you	0.4 (0.4 - 0.5)	101	10.9 (9.3 - 12.6)	97	4.52	Local Public Goods
You having someone in your life who always encourages you to be healthy	0.4 (0.4 - 0.5)	102	10.9 (9.3 - 12.8)	98	5.52	Relationships
The pleasure you get from helping people	0.4 (0.3 - 0.5)	103	7.2 (6.2 - 8.4)	105	4.78	Prosocial Virtues
How much you can trust most people in your nation	0.4 (0.3 - 0.5)	104	8.9 (7.4 - 10.5)	102	6.28	Non-local Public Goods
You doing your duty	0.4 (0.3 - 0.4)	105	7 (5.9 - 8.1)	107	4.60	Prosocial Virtues
At work, you getting to use your strengths to do what you do best every day	0.3 (0.3 - 0.4)	106	7.7 (6.6 - 8.9)	104	4.04	Work
You feeling generous	0.3 (0.3 - 0.4)	107	6.1 (5.2 - 7.2)	111	4.95	Prosocial Virtues
You having an impact on the world	0.3 (0.2 - 0.3)	108	7.2 (5.9 - 8.7)	106	7.00	Meaning and Fulfillment
How much you trust the police in your nation	0.3 (0.2 - 0.3)	109	6.9 (5.4 - 8.6)	108	12.54	Non-local Public Goods
You not being discriminated against for any reason	0.3 (0.2 - 0.3)	110	5.6 (4.5 - 6.8)	112	7.80	Non-local Public Goods
How much you trust the courts in your nation	0.3 (0.2 - 0.3)	111	6.3 (5.1 - 7.7)	110	9.60	Non-local Public Goods
You having the courage to talk to people you don't already know	0.3 (0.2 - 0.3)	112	6.9 (5.7 - 8.3)	109	7.42	Prosocial Virtues
You being satisfied with your city	0.2 (0.2 - 0.3)	113	5.5 (4.7 - 6.5)	114	4.45	Local Public Goods
How easy it is for you to express who you are through your lifestyle, culture, and beliefs	0.2 (0.2 - 0.3)	114	4.2 (3.6 - 5)	116	4.49	Status
You having a leader in your life who makes you enthusiastic about the future	0.2 (0.2 - 0.2)	115	5.6 (4.5 - 6.7)	113	7.10	Status
You having others remember your accomplishments long after your death	0.2 (0.1 - 0.2)	116	5.1 (4.1 - 6.2)	115	8.54	Status
Your understanding of the society you live in	0.2 (0.1 - 0.2)	117	3 (2.5 - 3.6)	121	5.98	Being Learned
How easy it is for you to get from home to work and back	0.2 (0.1 - 0.2)	118	3.3 (2.7 - 4)	120	8.28	Time
People doing what you tell them to do	0.2 (0.1 - 0.2)	119	3.6 (2.8 - 4.6)	119	9.65	Status
Your stature in your community	0.2 (0.1 - 0.2)	120	3.8 (3.1 - 4.6)	118	6.75	Status
You being part of a grand story	0.1 (0.1 - 0.2)	121	4.1 (3.3 - 5)	117	7.63	Meaning and Fulfillment
Your freedom from government interference in your dealings with people who are citizens of other countries	0.1 (0.1 - 0.2)	122	2.7 (2.2 - 3.4)	122	10.22	Non-local Public Goods
Your cultures and traditions being honored	0.1 (0.1 - 0.1)	123	2.2 (1.7 - 2.9)	123	14.95	Status
You having good taste and being up on the latest things	0.1 (0.1 - 0.1)	124	2.1 (1.7 - 2.6)	124	8.15	Status
How much you trust the media in your nation	0.1 (0.1 - 0.1)	125	2 (1.5 - 2.7)	125	16.97	Non-local Public Goods
Your enjoyment of risky situations	0 (0 - 0.1)	126	1.2 (0.8 - 1.6)	126	21.12	Being Learned

*Notes*: Hierarchical Bayesian model estimates using Hamiltonian Monte Carlo, based on 574,282 tradeoff responses from 896 respondents who completed the first nine blocks of the survey and passed quality control. Columns: (1) Aspect names; (2) Relative marginal utility  $M_j = e^{\mu_j}$  with 95% credible intervals in parentheses; (3) Rank based on  $M_j$ ; (4) Relative marginal utility per standard deviation of the aspect's rating scale, with 95% credible intervals; (5) Rank based on  $M_j$  per SD; (6)  $e^{\sigma_j}$ , where  $\sigma_j$  is the estimated standard deviation of individual-level log marginal utilities for aspect *j*; (7) Subjective grouping of aspects. Estimates are normalized so that the average log marginal utility across aspects equals zero; thus  $M_j$  can be interpreted as the marginal utility of aspect *j* relative to the geometric mean across all aspects. Standard subjective well-being questions highlighted in blue.

A particularly interesting pattern emerges when examining the standard SWB questions (highlighted in blue in Table 2). These aspects - including *How happy you feel*, *How satisfied you are with your life*, and similar measures - consistently show both relatively low marginal utilities compared to top aspects, and remarkably low variance in our sample. While the top family-related aspects have marginal utilities 10-20 times higher than the average aspect, these standard SWB measures have marginal utilities only about 1.6-2 times the average. Intriguingly, their variance estimates range from 1.84 to 2.59, making them among the most consistently valued aspects across respondents. This pattern suggests that while people broadly agree on the relative value of these general well-being measures, they have substantially higher and more variable marginal utilities for specific aspects such as family health and financial security.

Our baseline estimates use "points" as the quantity unit, thereby relying on respondents to use points in a way that is meaningful for comparing across aspects. However, we also conduct our analysis with standard-deviation units that depend less on how respondents interpret points. Columns 4–5 of Table 2 show the results after re-scaling our estimates to examine marginal utilities in terms of one–standard-deviation changes in the level of each aspect. The new scale hardly changes the ordering: the rank correlation between the points ranking and the SD ranking is 0.992, with family, financial security, and health still occupying the top tiers. The spread in marginal utilities is so large that this rescaling leaves the qualitative picture intact.

## 3.2 Grouping Aspects

To better understand broad patterns in what people value, we group aspects into categories using two distinct approaches. First, we create subjective groupings by manually categorizing aspects based on their text descriptions alone. This yields several intuitive categories such as Family Well-being, Health, and Financial. For each category, we average the marginal utility and preference heterogeneity estimates across aspects, and report the results in Table 3.

Group Es	stimates		
Grouping (# Aspects)	$\boldsymbol{\mathcal{M}}_{j}$	LB - UB	e <sup>σ</sup> j
Family Well-being (9 aspects)	8.86	(8.05 - 9.74)	9.91
Summum Bonum (4 aspects)	3.31	(3.11 - 3.52)	3.49
Financial (17 aspects)	2.56	(2.47 - 2.67)	5.85
Mental Health (3 aspects)	1.73	(1.62 - 1.85)	3.52
Relationships (11 aspects)	1.58	(1.5 - 1.68)	6.11
Health (8 aspects)	1.55	(1.48 - 1.63)	4.94
Feelings (11 aspects)	1.41	(1.36 - 1.45)	2.94
Meaning and Fulfillment (15 aspects)	0.77	(0.74 - 0.8)	3.43
Local Public Goods (6 aspects)	0.61	(0.57 - 0.65)	5.74
Time (7 aspects)	0.59	(0.56 - 0.63)	4.44
Prosocial Virtues (8 aspects)	0.51	(0.49 - 0.54)	4.66
Work (3 aspects)	0.46	(0.42 - 0.5)	4.10
Being Learned (5 aspects)	0.30	(0.28 - 0.33)	6.76
Status (11 aspects)	0.29	(0.27 - 0.31)	6.95
Non-local Public Goods (8 aspects)	0.27	(0.25 - 0.29)	10.75

#### **Table 3 – Subjective Groups**

*Notes*: Hierarchical Bayesian model estimates using Hamiltonian Monte Carlo, based on 574,282 tradeoff responses from 896 respondents who completed the first nine blocks of the survey and passed quality control. Table shows group-level averages computed from the aspect-level estimates in Table 2. Columns: (1) Grouping name with number of aspects in parentheses; (2) Average relative marginal utility across aspects in the group  $M_j$ ; (3) 95% credible interval; (4) Average of  $e^{\sigma_j}$  across aspects in the group. Groups ordered by average relative marginal utility.

The groupings confirm a clear hierarchy of what people value most. **Family well-being** sits at the top with almost every aspect related to children's, family, or loved one's welfare commanding a far higher marginal utility than any other domain. A small "**summum bonum**" group (e.g., *The happiness of you and your family*) follows, trailed closely by **financial security**. The importance of allaying money-related worry is consistent with Locke's view that emphasizes the importance of basic property rights, and with Sen's (1999) claim that economic capability is a precondition for the freedom to live according to one's values. **Physical and mental health** form the next tier, matching Maslow's (1943) placement of health at the base of the needs pyramid, and Grossman's (2017) model of health as productive human capital. **Relationships** are ranked similarly, reinforcing their role as a key determinant of well-being. Finally, the **Feelings** (moment-to-moment affect) cluster ranks last among the above-average clusters. The finding that

people value positive daily emotions, yet are willing to trade them for security or health mirrors Kahneman and Deaton's (2010) finding that emotional well-being levels off once basic needs are met, and Benjamin et al.'s (2014) finding that individuals sometimes forgo momentary happiness to attain longer-term goals.

The second half of the distribution is more varied. **Prosocial virtues** (e.g. *You feeling generous*) and **non-local public goods** (e.g., *How much you trust the courts in your nation*) post modest marginal utilities, suggesting that altruistic concerns carry substantially less weight once respondents must trade them against family and security. **Work** aspects fall lower than one might expect given surveys such as one conducted by Pew Research Center (2021) which finds that occupation is the second-most cited source of life meaning. At the bottom sit **status** and prestige concerns, echoing Ryan and Deci's Self-Determination Theory (2020) in that extrinsic aspirations deliver weak well-being returns. Together, these contrasts indicate that when respondents confront explicit trade-offs, the importance of family, health, and financial safety crowd out extrinsic goods more than many narrative surveys may imply.

To address bias concerns with our subjective categories, we employ a second approach involving computational text analysis to create data-driven groupings. We use ChatGPT (gpt-4) to derive two-sentence descriptions of each aspect. Then, using natural language processing techniques, we convert each aspect's text description into a numerical representation (text embeddings) and apply clustering algorithms to group similar aspects together. Analogous to our subjective groupings, this procedure only utilizes the text of the aspects, and no other information. A complete description of this process is included in Appendix Section A1.

This automated approach identifies clusters that largely align with our intuitive categorization, while revealing some interesting nuances. The top three clusters mirror our subjectively labeled groups with **family health and well-being** emerging as the highest-marginal utility cluster ( $M_j = 4.17$ ), followed by **love and relationships** ( $M_j = 3.60$ ) and **financial** security ( $M_j = 2.70$ ). Notably, sleep forms its own distinct cluster with moderate marginal utility ( $M_j = 1.04$ ). At the lower end of the spectrum, we find social justice ( $M_j = 0.24$ ) and social trust ( $M_j = 0.21$ ) clusters, again reinforcing our earlier observation that aspects related to external concerns and extrinsic motivations tend to have lower marginal utilities despite their potential social desirability.

28

Group Es	stimates		
Grouping (# Aspects)	$\boldsymbol{\mathcal{M}}_{j}$	LB - UB	$e^{\sigma_j}$
Family health and well-being (17 aspects)	4.17	(3.96 - 4.4)	6.65
Love and relationships (7 aspects)	3.60	(3.33 - 3.91)	7.77
Financial security (13 aspects)	2.70	(2.58 - 2.83)	6.31
Physical health and power (4 aspects)	1.99	(1.87 - 2.11)	4.18
Peace and security (11 aspects)	1.52	(1.46 - 1.58)	3.85
Sleep (3 aspects)	1.04	(0.96 - 1.13)	4.40
Fulfillment and satisfaction (21 aspects)	0.81	(0.79 - 0.84)	3.10
Morality (10 aspects)	0.81	(0.77 - 0.86)	4.79
Growth and experience (3 aspects)	0.81	(0.75 - 0.87)	3.93
Social connection (5 aspects)	0.50	(0.47 - 0.54)	5.09
Vigor and impact (12 aspects)	0.42	(0.39 - 0.44)	6.36
Home/Community Amenities (8 aspects)	0.33	(0.31 - 0.35)	5.46
Social justice (8 aspects)	0.24	(0.22 - 0.26)	9.39
Social trust (4 aspects)	0.21	(0.18 - 0.23)	11.02

### **Table 4 - Groups from Embeddings**

*Notes*: Hierarchical Bayesian model estimates using Hamiltonian Monte Carlo, based on 574,282 tradeoff responses from 896 respondents who completed the first nine blocks of the survey and passed quality control. Table shows group-level averages computed from aspect-level estimates, where groupings were created using computational text analysis and clustering algorithms as described in Appendix A1. Columns: (1) Grouping name with number of aspects in parentheses; (2) Average relative marginal utility across aspects in the group  $M_j$ ; (3) 95% credible interval; (4) Average of  $e^{\sigma_j}$  across aspects in the group. Groups ordered by average relative marginal utility.

## 3.3 Comparing Tradeoff-Based Estimates with Direct Importance Measures

Building on the Aspect Flagging survey described in section 1.6, we analyse three specific flags that help validate our tradeoff-based approach to measuring marginal utilities. Each flag was included in a different survey wave and captures a distinct way of asking respondents about the importance they place on different aspects of well-being. We compare these direct importance measures with our theoretically-motivated tradeoff estimates to shed light on both our measurement approach and the interpretation of our results.

We analyse three distinct survey flags, each asking about aspect importance in different ways. The first asks about global importance: "How important is this aspect to you?" The second focuses on marginal importance relative to one's current circumstances: "Think about your own life as it is. Given how much you have of this aspect, how eager would you be to get more? In particular, compared to other aspects of life, how valuable would it be to have a little bit more of this aspect?" The third explicitly frames the question in terms of tradeoffs: "Imagine you could gain more of this aspect by giving up other things. How much would you be willing to give up?" These questions progressively incorporate more of the key features of our original tradeoff framework: the focus on marginal changes relative to current circumstances, and the inherent tradeoffs between different aspects of well-being.

Figure 4 presents scatter plots comparing our estimated log marginal utilities  $(\mu_j)$  with these three direct measures, each standardized as described in Section 1.6. Several patterns emerge. First, all three direct measures show substantial correlation with our tradeoff-based estimates, with correlations of 0.84 for the global importance flag, 0.87 for the marginal value flag emphasizing current circumstances, and 0.93 for the tradeoff-focused flag. This strong alignment validates our tradeoff-based approach. Aspects that people say are important are indeed the ones they prioritize when faced with explicit tradeoffs.

Second, the relationship between direct importance measures and log marginal utilities becomes stronger as the questions more closely mirror the conceptual framework underlying our tradeoff approach. The weakest correlation between flag ratings and the marginal utility estimates is with global importance, which neither emphasizes marginal changes nor makes tradeoffs salient. The correlation strengthens for marginal importance, which explicitly asks respondents to consider the value of incremental improvements given their current circumstances. The strongest relationship emerges with the tradeoff-focused question, which captures both the local nature of the marginal changes and the necessity of giving up other aspects to achieve improvements. To test if correlations differ across the flags, we conduct a pooled regression of our standardized log marginal utility estimates on the standardized flag ratings while allowing the slope to differ across flags. We find that the correlation for the global importance flag (0.84), is statistically indistinguishable from that of the marginal-value phrasing (0.87), and marginally lower than the correlation for the explicit trade-off phrasing (0.93, p  $\approx$ 0.10). Taken together, the three flags show that direct importance ratings can capture much of the information contained in tradeoffs but can come closest when the question wording itself approaches the economic notion of a marginal tradeoff.

30



## Figure 4: Comparing $\mu_i$ to Direct Importance Elicitation

*Notes*: Each panel compares log marginal utility estimates ( $\mu_j$ ) from the main hierarchical model with standardized ratings from different aspect flagging question. Left panel: "How important is this aspect to you?" (281 respondents rating 126 aspects, N = 13,056). Middle panel: "Given how much you have of this aspect, how eager would you be to get more?" (364 respondents rating 126 aspects, N = 16,018). Right panel: "Imagine you could gain more of this aspect by giving up other things. How much would you be willing to give up?" (281 respondents rating 126 aspects, N = 13,056). Flag ratings standardized within respondent as described in section 1.6. Labeled aspects represent the five furthest from the OLS regression line.

One of the most important findings from comparing "importance" flags to log marginal utility estimates would be easy to miss: the relationship between "importance" as measured by any of these three flags with log marginal utility looks quite linear, which implies that the relationship with marginal utility itself would be highly nonlinear. A common result from psychophysics is that subjective perceptions of sensations such as loudness or brightness are more or less linear in the *logarithm* of intensity as measured by the usual metrics of physics. Here we have evidence that the same principle applies to marginal utilities as well: perceptions of importance on a subjective scale are linear in the logarithm of marginal utilities quantified carefully according to economic theory.<sup>11</sup>

## 3.4 Robustness to Social Desirability Concerns

One might expect that social desirability concerns would artificially inflate the reported importance of pro-social aspects given our reliance on hypothetical tradeoffs. Yet notably, we find several prosocial virtues such as You feeling generous, and You doing your duty have low estimated marginal utilities (ranked 107 and 103 out of 126, respectively). To investigate whether social desirability influences our marginal utility estimates, we analyse data from our aspect flagging survey, where we asked respondents whether "someone would downplay how much they want this aspect because openly wanting it might make them look bad." After constructing aspect-level scores for the flag as described in section 1.6, we find a significant negative relationship (slope = -1.71, correlation = -0.39,  $R^2 = 0.148$ ) between aspects' scores and the estimated log marginal utility means. The magnitude of this relationship is substantial: an aspect scoring at the lowest end of this social desirability flag (such as Your children's health) would have a predicted log marginal utility mean 2.69 higher than an aspect scoring at the highest end (such as How high your income is compared to the income of other people around you).<sup>12</sup> This observed relationship between social desirability and marginal utilities could be interpreted in multiple ways. It might reflect genuine alignment between social norms and true preferences, or it could indicate that our tradeoff-based measurement remains susceptible to social desirability influences. Under either interpretation, this relationship explains only 15% of

<sup>&</sup>lt;sup>11</sup> A key complication is that the meaning of a "point" might change between aspects or at different levels. This could affect the linearity of "importance" flags in log marginal utility.

<sup>&</sup>lt;sup>12</sup> Figure A6 in the Appendix displays this relationship visually.

the variation in log marginal utility aspect means, indicating that while social factors matter, substantial meaningful variation in preferences across aspects remains.

To see the impact of controlling for social desirability, we look at the relationship between the log marginal utility estimates and the residuals from regressing these estimates on the social desirability scores (Figure 5). We interpret these residuals as the log marginal utility means after controlling for the impact of social desirability concerns on marginal importance as captured through our hypothetical tradeoffs. The labeled aspects are the 5 aspects furthest from the 45-degree line, indicating the aspects for which relative marginal utility changes the most after removing social desirability concerns. Aspects above the 45-degree line, such as *Your spouse/partner's health* have lower relative marginal utilities, while aspects such as *How high your income is compared to the income of other people around you* have higher relative marginal utilities after correction. Despite these shifts, the overall ranking of aspects remains largely unchanged, with the rank correlation between the uncorrected log marginal utility estimates and the residuals being 0.94.





Residual from regressing Log MU on Social Desirability

*Notes*: Comparison of log marginal utility estimates before and after controlling for social desirability concerns. Vertical axis shows residuals from regressing  $\mu_j$  on social desirability scores, where social desirability is measured using the flag "Would someone downplay how much they want this aspect because openly wanting it might make them look bad?" (364 respondents rating 126 aspects, N = 16,380). Flag ratings standardized within respondent as described in Section 1.6. Labeled aspects represent the five furthest from the OLS regression line.

# 4. Explaining differences in marginal utilities across people

# 4.1 MU vs Levels

To understand how relative marginal utilities vary across individuals, we explore heterogeneity by individual circumstance. We group individuals based on their reported levels of each aspect, dividing the sample into quintiles. Within each quintile, we average the individual estimates of log marginal utility for that aspect. Figure 6 shows this relationship for the 16 aspects that feature in Baseline and every block of the Bottomless survey.



Figure 6: Relationship between level quantiles and  $\mu_i$  for 16 core aspects

*Notes*: Individual-level log marginal utilities averaged within quintiles of self-reported aspect levels. Based on 896 respondents who completed the first nine blocks and passed quality control. For each aspect, respondents are divided into quintiles based on their self-reported level of that aspect, and mean log marginal utilities are calculated within each quintile group. Error bars represent 95% credible intervals. The 16 core aspects shown appear in the baseline survey and every block of the bottomless survey.

While many aspects display a downward-sloping relationship between levels and marginal utilities (consistent with diminishing marginal utility), this pattern is not universal nor mechanically determined by our estimation approach. To better understand what we can infer from these empirical patterns, we propose a framework of production and preferences over aspects of well-being.

Consider two aspects of well-being: "feelings of excitement" and "feelings of security." Individuals have technologies for producing these aspects and preferences over their levels. The relationship between current levels and marginal utilities depends critically on what drives variation in aspect levels across individuals. We illustrate this with two stylized scenarios.



Figure 7: Production possibilities and preferences when individuals differ only in preferences

In the first scenario (Figure 7), all individuals have identical production possibilities—they face the same concave production possibility frontier (PPF) and budget for transforming inputs into excitement and security. Variation in observed aspect levels stems entirely from heterogeneity in demand. Some individuals strongly prefer excitement over security and therefore choose points on the PPF with high excitement and low security. Others with a stronger taste for security choose the opposite.



Figure 8: Production possibilities and preferences when individuals differ only in

The second scenario (Figure 8) illustrates the opposite case. Here, all individuals have identical utility functions (represented by the same convex indifference curve), but they differ in their production technologies. Some individuals are relatively more efficient at producing excitement, while others can more easily generate feelings of security.

These two scenarios represent extremes on a spectrum. In reality, both demand heterogeneity and production technology differences likely contribute to the patterns we observe. The downward-sloping relationship observed for many aspects suggests that production differences play a dominant role—individuals with lower levels of an aspect typically value marginal improvements more highly. However, the varying slopes and occasional upwardsloping relationships indicate that demand heterogeneity also matters substantially. This framework helps interpret the heterogeneity in our estimates. When we observe strong negative relationships between aspect levels and marginal utilities (as with many health and financial security aspects), it suggests that differences in individuals' capacity to produce these aspects drive much of the observed variation. When relationships are flatter or even positive, it suggests that demand heterogeneity plays a larger role in determining aspect levels.<sup>13</sup>

## 4.2 Validation of the Supply vs. Demand Framework

To directly test our interpretive framework, we included a specific question in our aspect flagging survey designed to distinguish between supply-driven and preference-driven heterogeneity. The question asked:

For each aspect, think about **people you know who don't care much about getting more of it.** They might not care much about getting more for either of two reasons:

1. They already have plenty of it.

<sup>&</sup>lt;sup>13</sup> To tap into other elements of supply-and-demand intuition, it can be helpful to define a "supply curve" as a curve with the same variable on the horizontal axis, and the absolute value of the slope of the possibility frontier on the vertical axis. Similarly, define the "demand curve" as a curve with the same variable on the horizontal axis, and the absolute value of the slope of the slope of the tangent indifference curve on the vertical axis. Then the optimum will be at the intersection of the "supply curve" and the "demand curve." Note that in this formulation, anything that moved the individual to a higher indifference curve would shift the "demand curve" in a way that is analogous to an income effect.

2. They don't value it much. For each aspect below, of people who don't care much about getting more, what fraction have each reason?

Respondents provided their judgment on a scale where lower scores indicate "they already have plenty of it" (supply-driven indifference) dominates, while higher scores indicate "they don't value it much" (demand-driven indifference) dominates.

Our framework generates a clear prediction: aspects where supply differences dominate should show steeper negative relationships between current levels and marginal utilities, while aspects where demand differences dominate should show flatter or even positive relationships. To test this prediction, we compare each aspect's score on this supply-vs-demand flag with the empirical slope of the relationship between ratings and marginal utilities.



Figure 9: Empirical Slope vs Supply-Demand Score

*Notes*: Relationship between the supply-demand characterization of aspects and their empirical levelmarginal utility slopes. Empirical slopes from Figure 6 measure the relationship between aspect levels (quintiles) and log marginal utilities. Supply-demand scores from aspect flagging survey where respondents rated why people don't care about getting more of each aspect. Based on 364 respondents rating 126 aspects (N = 16,380), with ratings standardized within respondent as described in Section 1.6. Labeled aspects represent the 5 furthest from the OLS regression line. Correlation and slope coefficients shown with 95% confidence intervals. Figure 9 confirms this prediction. We find a statistically significant relationship between the aspect's supply-vs-demand score and the empirical slope of its ratings-log marginal utility relationship (slope = 0.86). Aspects that respondents judged as predominantly supply-driven (low scores) show steeper negative slopes, consistent with our second scenario where production technologies differ. Conversely, aspects judged as predominantly demand-driven (high scores) show flatter or more positive slopes, consistent with our first scenario where tastes differ.

This finding is in line with our interpretive framework. It demonstrates that the observed relationships between aspect levels and marginal utilities reflect meaningful economic mechanisms rather than statistical artifacts. Moreover, it suggests that respondents have intuitive understanding of the sources of heterogeneity in different aspects of well-being, and these intuitions align with the patterns we observe in our marginal utility estimates.

## 4.3 Demographic Heterogeneity

To investigate demographic differences in marginal utilities, we leverage our individual-level estimates from the full hierarchical model to examine group-level patterns. After obtaining the joint posterior distribution of all model parameters, we compute the posterior distribution of the average log marginal utility for each aspect within demographic groups by averaging the individual-level estimates across group members for each posterior draw. This approach maintains the correlation structure of our estimates while allowing us to quantify uncertainty in group differences. Figure 10 presents one example of a scatter plot of the posterior means of these group-averaged log marginal utilities, revealing the degree to which preferences align or diverge across demographic categories.



Figure 10: Scatterplot of  $\mu_i$  by women vs men

*Notes*: Comparison of group-averaged log marginal utilities between women and men. Each point represents an aspect, with coordinates showing the posterior mean of averaged individual-level log marginal utilities for each gender. Based on the same hierarchical model and 896 respondents as in Table 2. Error bars represent 95% credible intervals. Red points and labels indicate aspects where the credible intervals do not overlap between genders. The dashed diagonal line represents perfect agreement between groups. Child-related aspects use only parents' responses for both genders; spouse-related aspects use only partnered individuals' responses for both genders.

Surprisingly, we find remarkable consistency in relative marginal utilities across demographic groups (see Benjamin et al. 2014b, who finds s a similar result). Figure 9 shows that most aspects cluster tightly along the 45-degree line when comparing across men and women, indicating broad agreement about which aspects of well-being matter most. This pattern persists when comparing other demographic groups, as we show in Figure A3. This consistency suggests that despite individual heterogeneity, there exists a shared understanding of well-being priorities that transcends many demographic boundaries.

Nevertheless, we do observe some meaningful differences aligned with life circumstances and social roles. Figure 11 highlights these selective differences for parents versus non-parents. Parents place significantly higher marginal utility on aspects related to children and family relationships.<sup>14</sup> For example, *Your children's health* and *Having a strong bond with your children* show substantially higher marginal utilities among parents.



## Figure 11: Scatterplot of $\mu_i$ by non-parents vs parents

*Notes*: Comparison of group-averaged log marginal utilities between non-parents and parents. Each point represents an aspect, with coordinates showing the posterior mean of averaged individual-level log marginal utilities for each group. Based on the same hierarchical model and 896 respondents as in Table 2. Error bars represent 95% credible intervals. Red points and labels indicate aspects where the credible intervals do not overlap between groups. The dashed diagonal line represents perfect agreement between groups. For child-related aspects, non-parent values are estimated from the hierarchical model structure as described in Section 4.3; spouse-related aspects use only partnered individuals' responses for both groups.

These differences extend to relationship status more broadly, with partnered individuals placing higher marginal utility on aspects such as *Your spouse/partner's health*, *The happiness of your* 

<sup>&</sup>lt;sup>14</sup> As discussed in Section 3.1, we handle aspects such as *Your children's health* by excluding non-parent tradeoffs from the estimation. However, our hierarchical model still generates  $m_{ij}$  values for these aspects for all individuals, including non-parents, through the population-level parameters and individual random effects. The non-parent values shown in Figure 11 represent these model-implied estimates, which reflect the hierarchical structure even though they were not directly informed by non-parent tradeoffs for child-related aspects.

family, Your relationship with your spouse/partner or closest friend being stronger than ever, The quality of your spouse/partner as a mate, The quality of your romantic relationships, marriage, love life, or sex life, and Your spouse/partner taking on fair share of responsibility for the household.

Demographic differences beyond family structure are more limited. Income shows no systematic variations in marginal utilities across aspects, contrary to what one might expect if certain aspects functioned as luxury goods. Gender differences are minimal, with the only notable distinction being that women place higher marginal utility on *You not being anxious*. Education levels show no significant differences in marginal utilities across aspects. Age correlations are similarly limited, with older respondents placing higher marginal utility on *Your physical health*, reflecting the increasing salience of health concerns with age. Racial differences are minimal, with the most notable distinction being that Black respondents place higher marginal utility on *Not being discriminated against for any reason*, a pattern that likely reflects lived experiences with discrimination.

Political affiliation shows expected differences: Democrats place higher marginal utility on *Women being treated fairly*, while Republicans value *Your culture and traditions being honored* more highly. Employment status reveals that unemployed respondents place higher marginal utility on *You not being anxious*, possibly reflecting the psychological impact of joblessness. Religious differences center primarily on cultural aspects, with religious respondents placing higher marginal utility on *Your culture and traditions being honored*. We provide the figures for each of these splits in Appendix Section A4.

These patterns suggest that while there exists broad agreement on well-being priorities, life circumstances and social identity do shape marginal utilities in predictable ways. However, the modest magnitude of these differences, relative to the overall consistency across groups, suggests that demographic factors explain only a small portion of the heterogeneity in marginal utilities that we observe across individuals.

# 5. Conclusion

Philosophers and social scientists have long theorized about what matters to people. Our findings confirm some historical perspectives and challenge others. We have provided systematic evidence on the relative marginal utilities of a wide range of aspects of well-being. We have four main conclusions. First, existing theories have missed some of the aspects of well-being that we find are most highly valued, such as Your children's health, Your financial security, and The absence of worry in your life. Second and relatedly, like previous related work, we find that happiness and life satisfaction, which are often assumed to capture all of what people care about, have far lower relative marginal utilities than many other aspects of well-being. Third, while there is substantial heterogeneity in preferences across individuals, little of that heterogeneity can be identified with the sociodemographic variables we measured. Finally, heterogeneity across individuals is driven by both supply and demand: for some aspects of well-being, there is a negative relationship between the level of the aspect and its marginal utility, suggesting that supply differences are operating (though there may also be demand differences pushing somewhat in the opposite direction), while for other aspects, there is a positive relationship between the level of the aspect and its marginal utility, suggesting that demand differences are operating (though there may also be supply differences pushing somewhat in the opposite direction). Validating this interpretation, survey respondents' intuition about when supply differences should be especially important and when demand differences should be especially important predicts the correlation between levels and marginal utilities.

One important limitation of our paper is that we primarily rely on stated preferences. Although features of our design aim to account for social-desirability bias and differences across respondents in how they interpret the aspects of well-being, stated preferences could differ from "true" preferences for various reasons, including simply that people may not be able to fully imagine changes in certain aspects of well-being. By finding real-world situations in which people face tradeoffs like those we study for a small set of aspects of well-being (as in Benjamin et al. 2014b), future work with real choices could provide an independent test of our methodology. Future research should also explore how these preferences evolve over longer time horizons and across different cultural contexts. Finally, developing methodologies to translate these marginal utility estimates into practical welfare indices could provide valuable tools for policy evaluation.

42

# References

Adler, M. D., Dolan, P., & Kavetsos, G. (2017). Would you choose to be happy? Tradeoffs between happiness and the other dimensions of life in a large population survey. *Journal of Economic Behavior & Organization*, *139*, 60-73.

Aristotle. (ca. 350 BCE). *Nicomachean Ethics*. Translated and edited by Terence Irwin. 3rd ed. Indianapolis: Hackett Publishing, 2019.

Gary S. Becker, (1965). A Theory of the Allocation of Time. *The Economic Journal*, *75*(299), 493–517.

Benjamin, D.J., Cooper, K., Heffetz, O. and Kimball, M., (2023a). From happiness data to economic conclusions. *Annual Review of Economics*, 16.

**Benjamin, D. J., Debnam Guzman, J., Fleurbaey, M., Heffetz, O., & Kimball, M.** (2023b). What do happiness data mean? Theory and survey evidence. *Journal of the European Economic Association, 21*(6), 2377-2412.

Benjamin, D. J., Heffetz, O., Kimball, M. S., & Rees-Jones, A. (2012). What do you think would make you happier? What do you think you would choose?. American Economic Review, *102*(5), 2083-2110.

Benjamin, D. J., Heffetz, O., Kimball, M. S., & Rees-Jones, A. (2014a). Can marginal rates of substitution be inferred from happiness data? Evidence from residency choices. *American Economic Review*, *104*(11), 3498-3528.

Benjamin, D. J., Heffetz, O., Kimball, M. S., & Szembrot, N. (2014b). Beyond happiness and satisfaction: Toward well-being indices based on stated preference. *American economic review*, *104*(9), 2698-2735.

Bond, T. N., & Lang, K. (2019). The sad truth about happiness scales. *Journal of Political Economy*, *127*(4), 1629-1640.

**Carstensen, L. L.** (2021). Socioemotional selectivity theory: The role of perceived endings in human motivation. *The Gerontologist*, *61*(8), 1188-1196.

**Dehaene, S. (2003).** The neural basis of the Weber–Fechner law: a logarithmic mental number line. *Trends in cognitive sciences*, *7*(4), 145-147.

**Gelman, A.** (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, *1*(3), 515-534.

**Giustinelli, P., Manski, C. F., & Molinari, F.** (2022). Tail and center rounding of probabilistic expectations in the health and retirement study. *Journal of econometrics*, *231*(1), 265-281.

**Grossman, M.** (2017). *The Demand for Health: A Theoretical and Empirical Investigation*. New York: Columbia University Press.

List, J. A., Sinha, P., & Taylor, M. H. (2006). Using Choice Experiments to Value Non-Market Goods and Services: Evidence from Field Experiments. *Advances in Economic Analysis & Policy*, 6(2), Article 2.

Locke, J. (1689). *Two Treatises of Government*. Edited by Peter Laslett. Cambridge: Cambridge University Press, 1988.

Lusk, J. L., & Schroeder, T. C. (2004). Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks. *American Journal of Agricultural Economics*, *86*(2), 467-482.

Kahneman, D. & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the national academy of sciences*, *107*(38), 16489-16493.

Maslow, A. H. (1943). A theory of human motivation. Psychological Review, 50(4), 370-396.

**Pew Research Center.** (2021). *What Makes Life Meaningful? Views from 17 Advanced Economies*. Washington, DC: Pew Research Center,

https://www.pewresearch.org/global/2021/11/18/what-makes-life-meaningful-views-from-17-advanced-economies/.

**Preston, C. C., & Colman, A. M.** (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta psychologica*, *104*(1), 1-15.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, 55(1), 68-78.
Sen, A. (1999). *Development as Freedom*. New York: Alfred Knopf.