Web Appendices A-J
“Adjusting for Scale-Use Heterogeneity in Self-Reported Well-Being”

Table of Contents

A. Survey Design and Sample Details

This appendix details the survey's design and pre-registration process, outlining the methodologies and strategies implemented.

A.1. Project Pre-registration

Using the Center for Open Science’s Open Science Framework, we pre-registered our project six times as we prepared to collect the survey data and conduct the analyses for this paper. (We describe these plans as “pre-registrations” since they were created before most or all of the data were collected.) These pre-registrations can be found on our project page using this link (which is view-only, with anonymized contributor names): *https://osf.io/qk5ta/?view_only=c2c05eb8d51d4c088c363db7a26d2f15*.

- We submitted the first two pre-registrations before we collected data on a pilot version of the Baseline survey in 2020 (Registration 1, 4/8/2020; Registration 2, 7/24/2020). The pilot survey was our first test for interpersonal differences in general scale use using SWB responses from 0 (“lowest level possible”) to 100 (“highest level possible”) and calibration questions on that scale. To learn about response error, we also fielded a test-retest version of the entire pilot survey, identical to the original, about two weeks[1] later. In this paper, the pilot data are only used for the analysis in Appendix I.1, to correct two of the slope coefficients (height and weight) for response error in a regression of subjective measures on objective ones.

- We submitted the next two pre-registrations before we collected the part of the Baseline survey data used for this paper that did not include a Prescreening survey (Registration 3, 5/18/22; Registration 4, 6/10/22).
- We submitted the next pre-registration (Registration 5, 7/18/22) before our first attempt at a Prescreening survey. (See A.3.ii. for a description of the Prescreening survey as finally implemented.) We began by fielding a small sample of 18 respondents. (When we field a new survey, we typically start with a small sample to make sure there is no error in the survey code or data collection process.) When we checked the data, we found that the Prescreening survey included a question that respondents did not interpret as we intended, and we decided to cut that question from the survey.[2] Instead of pooling data between two versions of the Prescreening survey, we opted not to follow up with those 18 respondents, and we do not use their data.
- We submitted the final pre-registration (Registration 6, 7/25/22) before we collected the part of the Baseline survey data used for this paper that did include a Prescreening survey.

In this section we briefly sketch how the many pieces of our pre-registered project materials fit with the paper as actually written. At a high level, what we did largely coincides with what we planned to do. However, our specific plans for the paper evolved as we piloted the

[1] Due to a data storage glitch (our date variable was truncated), we only have timing data for a small subset ($N = 41$) of respondents that took both versions of the pilot survey. Among this subsample, the retest version of the pilot was fielded between 13.6 and 16.3 days later, with an average gap of 14.2 days.

[2] We were considering collecting additional data on Prolific, and we wanted to avoid duplicating respondents across survey platforms, so our first version of the Prescreening survey asked respondents if they had completed the Baseline survey before (as described by the MTurk HIT title and a screenshot). 64% of respondents indicated that they had taken it, even though zero of their MTurk worker ID’s appeared in our previous data. From this we concluded that respondents, many of whom do numerous MTurk surveys, likely could not remember which surveys they had done. We dropped the question from the subsequent (final) version of the Prescreening survey.

survey, developed our scale-use correction methodology (in terms of concepts, specific calibration questions, and econometric strategies), and examined preliminary results. Readers of the pre-registrations will also note that, for the most part, we described our planned analyses only conceptually. We had partially developed our econometric approach. However, we came to realize that we needed an understanding of the distributions of the hyperparameters to figure out the details of the specifications, so we moved forward with data collection. Thus, we did not specify exact regression specifications, and we did not pre-register code.

The intention of the pre-registration evolved with the project. We described our initial strategy for pre-registration in Registration 1: "The primary purpose of this registration is to define the analyses which are already planned before data collection, so we can include the N=500 sample, and increase our statistical power, when conducting them." However, we modified the Baseline survey after this "N=500" pilot data collection, so we ultimately decided not to pool data from that sample (note, however, that the results are broadly similar between our pilot and subsequent data). The final registrations focused instead on our quality control procedure. As Registrations 3 through 6 show, it took us several attempts to develop an approach that made us confident that our data quality is sufficiently high. Our final quality control procedure is described in Section A.3.i. below.

The survey design and econometric models in the paper do not have a direct counterpart in the registration materials. Earlier versions of the econometric model were registered in Registrations 1 and 2 (in 2020). The core of the theoretical framework and econometric model have changed little, but our estimation procedures have been substantially improved in the intervening years. We did not try to keep the registered econometric model updated as that work unfolded. We registered the pilot version of Baseline survey variables in Registration 1. In Registration 2 we noted adjustments to 12 of the 18 Baseline CQs to reduce top- or bottom-coding, based on pilot data. (This is the primary reason we did not end up wanting to pool with pilot data.) Knowing that we would provide screenshots and details of our survey design in the eventual paper, we did not update the survey-question registration further.

The final category of pre-registered project plans is descriptive analyses. The analysis plans fall into three categories. First, two descriptive analyses were pre-registered in substantially the same format as they appear in this paper:

1. “Suggestive evidence of general scale use” from Registration 1 (4/8/2020) evolved into Figure 4 of the current paper (“Relationships Between CQ- and SWB-Rating Means and Standard Deviations”).
2. “Translation functions between demographic groups” from Registration 1 (4/8/20) evolved into Figure 5 (“Translation Functions Across Demographic Groups”).

Two analyses in the present paper were pre-registered, but in a very different format:

3. The idea behind “Demographic correlates of general scale use” from Registration 3 (5/18/22, which was an update from Registration 1) is similar to Table 3 of the current paper (“Regression of Mean and Standard Deviation of Baseline CQs on Demographics”). However, we modified the proposed regression in several ways, including using more standard demographic questions and a simpler approach.
4. A version of the “Curvature of the translation function” analysis sketched in Registration 1 is now in Web Appendix B, with a more sophisticated statistical approach.

Finally, three other types of analyses are described in the pre-registrations but not included in the present paper. In the course of analyzing our data, we have done some analyses similar to what we proposed, but the project has evolved such that these topics are outside the scope of the present paper:

5. Registration 1 contains materials for “multidimensionality” regressions where we planned to regress a huge set of socio-behavioral variables on different SWB ratings.
6. Registration 1 proposes “Translation functions to test specific calibration questions.”
7. Registration 3 proposes to explore the “U-shape in age” using both calibration questions and SWB questions.

A.2. Design of Baseline Survey

As noted in the main text of the paper, the “Baseline” survey consists of four types of questions:

(1) SWB questions. (In the pre-registrations, we referred to the objects being rated as “aspects of well-being”; in this paper, we refer to the questions themselves as self-reported well-being (SWB) questions.)

(2) Tradeoff questions, which ask respondents to choose between two options where the levels of well-being in different dimensions are higher or lower than the previously self-reported levels. The tradeoff questions are not analyzed in this paper.

(3) Calibration questions (CQs).

(4) Demographic and behavioral questions. These fall into five general subcategories, aimed at learning about respondents' demographics, behaviors, antecedents of scale-use tendencies,[3] interpretation of specific SWB questions, and survey effort/experience. Only the demographic variables are analyzed in this paper.

A.2.i. Baseline Survey Screenshots

Examples are shown in Appendix K.

A.2.ii. SWB Questions in Baseline Survey

The 33 SWB questions are arranged in a "triple" design. A triple consists of three SWB questions, followed by six pairwise tradeoffs corresponding to those SWB questions. In the Baseline survey, there are 11 unique triples and 12 triples in total; the 11th triple is a repeat of the 2nd triple as a check for data quality. The order of the SWB questions (i.e., their assignment to triples) is randomized by respondent.

Table A.1 lists the 33 SWB questions in the Baseline survey, in alphabetical order. The 16 SWB questions with stars are also repeated in every block of the Bottomless follow-up survey (described in Section A.4); we are studying these dimensions in more depth in other work.

Table A.1. SWB Questions in Baseline Survey

Thinking about the past year, how would you rate…
How happy you feel *
How much you can trust most people in your nation
How much you enjoy your life *
How satisfied you are with your life *
The ability of ordinary citizens to influence your national government

[3] To learn about possible antecedents of general scale use, we asked respondents about their state of birth, state where they grew up, native language, and other languages spoken at home. We also asked: "Think about the standards you use to grade yourself, and standards you use to grade other people. How would you compare them? I grade myself _______ I grade others" (the options are "much tougher than", "somewhat tougher than", "the same as", "somewhat easier than", and "much easier than"). We do not analyze these data in the present paper.

The absence of anger in your life
The absence of sadness in your life *
The absence of stress in your life
The absence of worry in your life
The air in your area not being polluted
The extent to which you feel the things you do in your life are worthwhile *
The happiness of your family *
The overall well-being of you and your family *
You and your family having enough to eat
You being a good person
You being a winner in life
You being able to support your family financially *
You feeling that you have enough time for the things that are most important to you *
You having many options and possibilities in your life and the freedom to choose among them
You having people you can turn to in time of need
You not being lonely *
You not feeling anxious *
You not having to worry about being unemployed
Your cultures and traditions being honored
Your home being comfortable
Your knowledge and skills
Your living environment not being spoiled by crime and violence
Your mental health *
Your physical health *
Your physical safety and security
Your rating of your life on a ladder where the lowest rung is "worst possible life for you" and the highest rung is "best possible life for you" *
Your sense of control over your life *
Your sense of purpose *

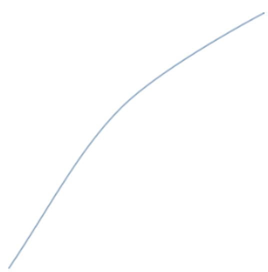
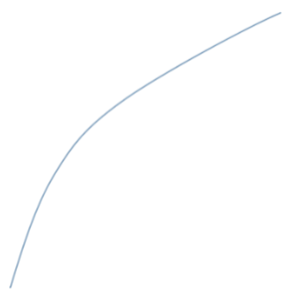
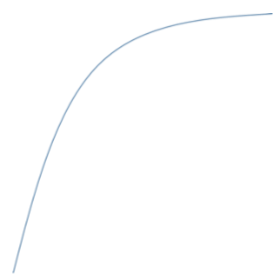
We developed the set of 33 SWB questions with several considerations in mind. The overall number of SWB questions was determined primarily by survey length constraints. For choosing specific SWB questions/dimensions, our primary goal was to try to cover dimensions of personal well-being as comprehensively as possible, informed by well-being measurement initiatives from the UK, New Zealand, Gallup, the OECD, and our own previous findings. To increase comprehensiveness, we also included four dimensions to capture the things for which people would trade off SWB measures, as found by Benjamin et al. (2012): sense of purpose, control over one's life, family happiness, and social status. Our other major consideration was choosing dimensions that had high estimated relative marginal utilities in Benjamin et al. (2014)

and/or our pilot data. We made sure to include two questions that correspond to the same dimension as a calibration question and potentially closely relate to objective measures: “Your living environment not being spoiled by crime and violence” and “The air in your area not being polluted.” Finally, we included two questions for exploratory analysis: “You and your family having enough to eat” and “You feeling that you have enough time for the things that are most important to you.”

A.2.iii. Calibration Questions in Baseline Survey

Table A.2 shows the 18 calibration questions in the Baseline survey.

Table A.2. Calibration Questions in Baseline Survey

How curved is this line?		
calibration_curve_1	calibration_curve_2	calibration_curve_3
		
How dark is this circle?		
calibration_blue_1	calibration_blue_2	calibration_blue_3
		
How big (by land area) is this region?		
calibration_region_1	calibration_region_2	calibration_region_3

If this situation described your life during the past year, how would you rate your level of
Your living environment not being spoiled by crime and violence*?*

calibration_crime_1

You have been mugged on more than one occasion, so you don't carry around much more money or valuables. The police have increased their presence in your neighborhood, which makes you feel a little safer.

calibration_crime_2

You live in an apartment complex where tenants sometimes get robbed while they are away. The complex owners recently hired a security guard. You sometimes hear people in the neighborhood fighting, but you haven't had any trouble yet.

calibration_crime_3

You live in a quiet neighborhood where there is almost no crime. Recently, you have noticed an uptick in the number of newspaper reports on home robberies, but none have been nearby.

If this situation described your life during the past year, how would you rate your level of
Your access to information*?*

calibration_info_1

An affordable internet package is not available in your area, so you do not have internet at home. You get television broadcasts, but not cable television. You have access to the internet during your breaks at work, and you read the local newspaper there.

calibration_info_2

You have a basic internet and cable television package at home. The speed of your internet connection is weak. You sometimes have difficulty streaming videos or loading webpages when multiple people in your family are using the internet.

calibration_info_3

You have internet and cable television (including premium channels) at home. You also get local access channels where several agencies feature local news. You wish you lived closer to the local library, which is several miles away.

If this situation described your life during the past year, how would you rate your level of
Your ability to remember things*?*

calibration_remember_1

You friends joke about how forgetful you are. You do often get lost, even in places you should know well. It's difficult for you to remember names, and you often lose track of things. But you usually remember to set reminders for important appointments.

calibration_remember_2

You have a pretty good memory for details of the past. However, you have to work hard to remember new things. You sometimes forget people's names or lose track of your phone, but you can usually find it right away.

calibration_remember_3

You have a good memory for the details of past events including names of people, places, streets and dates. You don't have trouble remembering things in everyday life. However, you have to write down your to-do lists and passwords to remember them.

A.2.iv. Vignette Design

Our general approach for developing the details of a vignette was to begin with a set of sub-dimensions relevant to that dimension of well-being. We considered what the highest and lowest possible ends of the spectrum for that sub-dimension would look like. For example, for the dimension ***Your living environment not being spoiled by crime and violence***, one of the relevant details or sub-dimensions is how closely you have been affected. In the worst-case extreme, you may have been affected by crime yourself; in the best-case extreme, you may not know of anyone personally who has been affected. For each SWB question, we identified two to three sub-dimensions. We wrote the three vignettes in the trio so they have monotonic changes in the sub-dimensions. This leads to a relatively clear ordering of “low,” “medium,” and “high” level vignettes. Note that our theoretical and statistical approach does *not* require monotonicity of the details in vignettes and does *not* require there to be a clear ranking of the vignettes in a trio. However, we hypothesize that such monotonicity—i.e., where every detail in one vignette is strictly higher or strictly lower than another vignette in the trio, for the dimension being rated—reduces respondents’ difficulty of forming a perception of the level.

The final feature of our vignette design attempted to reduce top- and bottom-coding (responses at 0 or 100). The high vignette in the trio always has a negative piece of information about one sub-dimension (a detail indicating non-maximality), and the low vignette in the trio always has a positive piece of information about one sub-dimension (a detail indicating non-minimality). By including these details, we are trying to guarantee that the person in the vignette cannot truly be an edge case.

Table A.3 reports the number and percentage of top- and bottom-coded responses for each calibration question in the Baseline survey. These suggest that we were reasonably successful at making top- and bottom-coding rare in CQs.

Table A.3. Top- and Bottom-Coding of Responses to Calibration Questions

Calibration Question	Number (Percentage) of Bottom-Coded Responses	Number (Percentage) of Top-Coded Responses
calibration_curve_1	4 (0.1%)	5 (0.1%)
calibration_curve_2	0 (0%)	7 (0.2%)
calibration_curve_3	3 (0.1%)	34 (1.0%)
calibration_blue_1	88 (2.5%)	3 (0.1%)
calibration_blue_2	8 (0.2%)	5 (0.1%)
calibration_blue_3	5 (0.1%)	26 (0.7%)
calibration_region_1	3 (0.1%)	1 (0.0%)
calibration_region_2	2 (0.1%)	2 (0.1%)
calibration_region_3	1 (0.0%)	13 (0.4%)
calibration_crime_1	125 (3.5%)	50 (1.4%)
calibration_crime_2	46 (1.3%)	40 (1.1%)
calibration_crime_3	7 (0.2%)	82 (2.3%)
calibration_info_1	59 (1.7%)	54 (1.5%)
calibration_info_2	17 (0.5%)	57 (1.6%)
calibration_info_3	11 (0.3%)	137 (3.9%)
calibration_remember_1	20 (0.6%)	22 (0.6%)
calibration_remember_2	6 (0.2%)	23 (0.6%)
calibration_remember_3	4 (0.1%)	50 (1.4%)

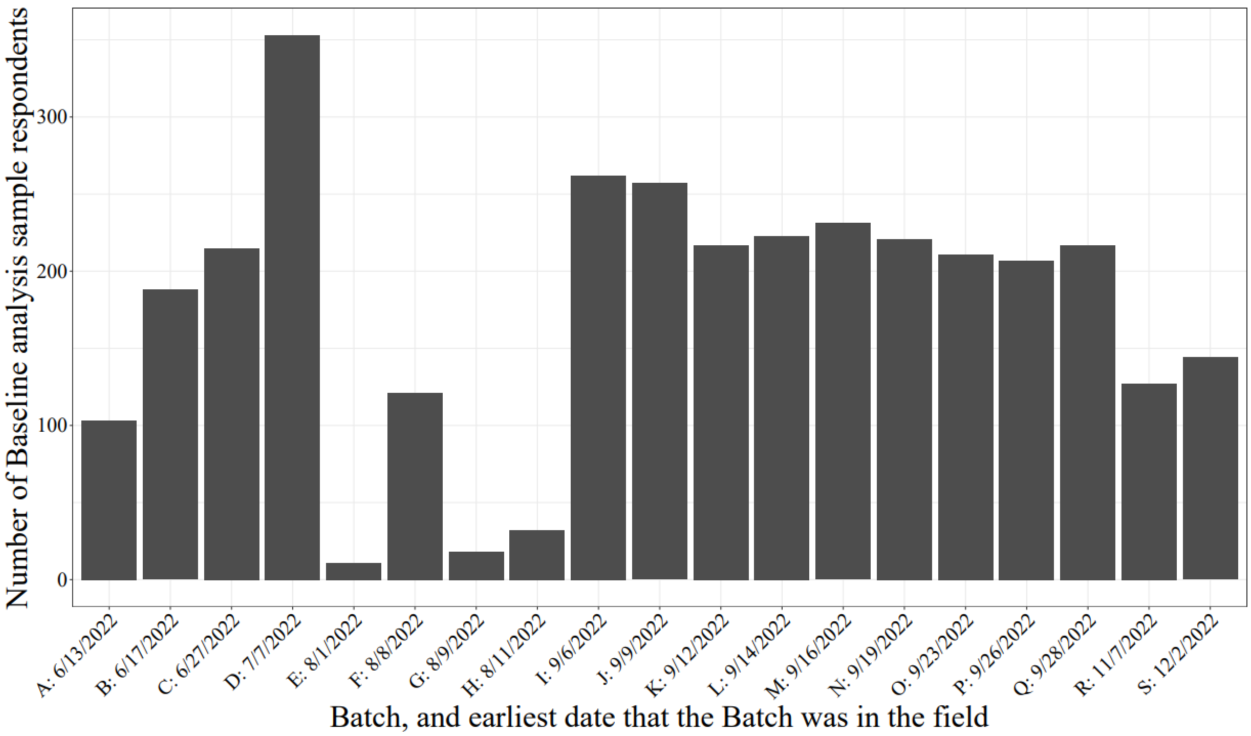
Notes: Bottom-coded responses refer to responses of 0, while top-coded responses refer to responses of 100. The total number of responses for each calibration question was 3,558.

A.3. Baseline Survey Sample

As described in Section I.C. of the paper, we fielded the Baseline survey on Amazon's Mechanical Turk (MTurk) platform between June 13 and December 7, 2022. This section describes the fielding in more detail. We proceeded in 19 batches, which we call Batch A through Batch S. Figure A.1 shows the number of respondents in our sample for analysis (N =

3,358) from each batch, by start date of batch. Changes in our quality control methodology underlie the respondent volume pattern seen in Figure A.1; we ramped up fielding quickly, then adjusted our recruitment method and fielded at a slower pace. These issues are discussed in the next section.

Figure A.1. Respondents in Baseline Sample for Analysis, by Start Date of Batch



A.3.i. Quality Control Procedure

This subsection describes the "quality control" (QC) criteria for excluding respondents from survey invitations (for Baseline or follow-up surveys) and excluding respondents from our primary analysis sample. (We also excluded individuals from invitations if they opted out of being contacted for follow-up surveys.) The goal of the criteria, overall, is to exclude respondents whom we suspect of being outside the U.S., being a bot (non-human respondent), and/or putting forth extremely low effort. This procedure is especially important given the concerns that many researchers have raised about the quality of MTurk samples (e.g., Kennedy et al., 2020; Peer et al., 2022).

<u>Baseline QC (Exclusion criterion for analysis sample and invitation to follow-up surveys)</u>

Our Baseline QC criterion was established in Registration 4 (6/10/2022), before the first fielding of the “Baseline” sample used for analysis in the paper. It was the only QC criterion used for Batches A through D. The exclusion criterion is based on three factors: IP address, textbox responses, and survey completion time. We weight the factors using points, as described below. If a respondent has three or more total points, they are excluded from invitations to follow-up surveys (including Bottomless) and from the primary analysis sample.

I. 3 points: Baseline survey was completed in less than 10 minutes. This based on (a) the fact that the fastest we can physically complete the survey is approximately 10 minutes and (b) 10 minutes is approximately the bottom percentile (fastest survey takers) of Baseline respondents in the pilot data who would not have otherwise been excluded based on the other criteria.
II. 2 points: IP Address originated from a virtual private server or from a non-US server. This is checked using an IP-checking service, DB-IP. We use an R package, rgeolocate, to access this service.
III. 1 point: IP Address appeared more than once in the dataset.
IV. 2 or 3 points: Poor responses to two textbox questions. The questions were:
    i. When you rated: “how satisfied you are with your life”, what was your thought process?
    ii. Think of a ladder where the lowest rung is ‘worst possible life for you’ and the highest rung is ‘best possible life for you’. Describe the life you imagine on the middle rung of the ladder.

    For both questions, the prompt in the textbox is: “Insert one or more sentences. Please use complete sentences.”

Free-response textbox answers were scored by a team of three or four undergraduate research assistants, hereafter raters, with worse answers getting more points. (We started with four raters, but one found they did not have enough time for the project and quit.) For the QC criterion, we use the largest point value at least half of the raters give a participant. For example, if there were three raters and the rating values were (0, 0, 3), the points used would be a 0. If there were four raters and the rating values were (0, 0, 2, 3), the points used for exclusion would be a 2. This rule is robust to the number of raters changing.

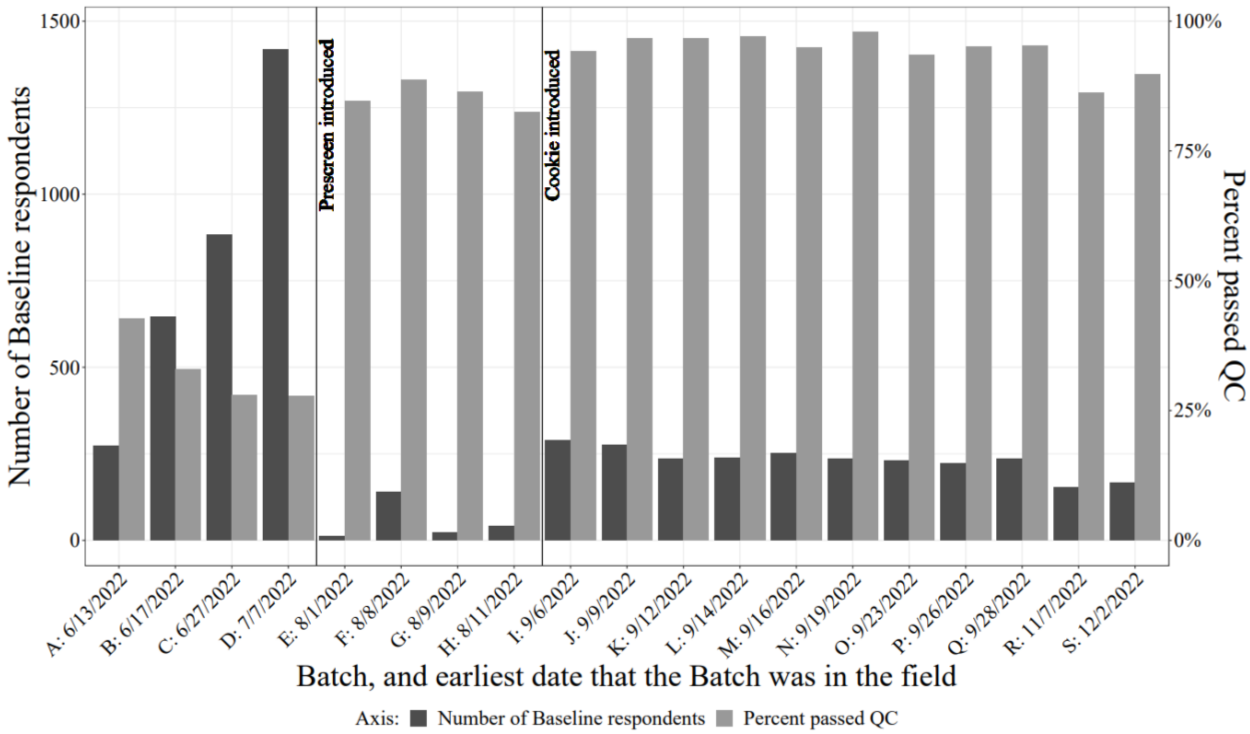
To develop the textbox rating procedure, we began by rating 597 text boxes ourselves from the pilot sample. We gave the raters written instructions for rating textbox responses, including examples of graded responses. They rated the pilot sample text boxes based on these instructions. Then, we provided feedback and modified the instructions based on the raters' points assignments, to bring their ratings closer to our ratings. Finally, the raters completed the textbox scoring for the non-pilot survey sample. To make sure the raters evaluated only the textbox responses, not other quality metrics or scale use tendencies, the raters did not work on any analysis tasks nor have access to any other data when they rated the textboxes.

Quality Control Results and Addition of Prescreening Survey

As we scaled up our recruiting in Batches A through D, we found that the percentage of respondents who passed QC declined. (By "passed QC" we mean those who were not excluded from follow-up studies and the analysis sample, based on the Baseline QC exclusion criterion described above.) This was especially troubling because the percentage passing QC in Batch A was already lower than it had been in the pilot. To avoid paying so many subjects for surveys with unusable data, we introduced a "Prescreening" survey, described further below. After testing the Prescreening survey and QC procedure with a few small batches (E-H), we introduced our final tool for quality control: using Qualtrics' "prevent multiple submissions" feature on the Prescreening survey. With this feature enabled on a Qualtrics survey, Qualtrics places a cookie on the respondent's browser when they submit the survey, and Qualtrics does not permit the respondent to take that survey again. Figure A.2 shows the total number of Baseline respondents and the percentage of those respondents who passed the Baseline QC, by fielding Batch. It also notes the points at which the Prescreening survey and cookie were introduced.

In Figure A.2, the effect of the Prescreening survey on our percentage of respondents passing Baseline QC is a clear increase, from Batch D to E. The Baseline QC pass rate also increased when the cookie was introduced, from Batch H to I. The results of the QC procedure on our survey data are discussed further in section A.3.iii.

Figure A.2. Quality Control Results by Batch



A.3.ii. Prescreening Survey Details

The Prescreening Survey and related QC procedure are described in Registration 6 (7/25/2022). The Prescreening Survey has 11 questions in total. We administered the survey using Qualtrics. The outline of the survey is as follows: (a) Consent form; (b) Abbreviated version of the Baseline instructions; (c) One “triple” (three SWB questions followed by six tradeoff questions) featuring three dimensions of well-being chosen randomly from the 20 Baseline SWB dimensions that are not “public” or “double negative” dimensions (to make the survey instructions as simple as possible); (d) Two free-response (textbox) questions about what the respondent was thinking about while answering survey questions.

Prescreening QC (Exclusion criteria for invitation to Baseline)

The Prescreening QC criterion is based on IP address and textbox responses in the Prescreening survey. If a respondent has three or more total points, they are not invited to the Baseline survey. (Note that our payment of subjects was not tied to these criteria. Respondents were compensated $0.75 for completing the Prescreening survey, regardless of whether or not they passed our Prescreening QC.)

I. 2 points: IP Address originated from a virtual private server or from a non-US server. This is checked using an IP-checking service, DB-IP. We use an R package, rgeolocate, to access this service.

II. 1 point: IP Address appeared more than once in the dataset.

III. 2 or 3 points: Poor responses to the two free-response questions in the Prescreening survey. These questions are:

   i. In the rating questions, 0 was the "Lowest level possible" and 100 was the "Highest level possible." How would you describe how you interpreted the word 'possible'?

   ii. When you made the decisions, what were you thinking about? As a reminder, here is an example decision. [Screen includes a cropped screenshot of a tradeoff question.]

   For both questions, the prompt in the textbox is: "Insert one or more sentences here. Please use complete sentences."

The Prescreening textboxes were scored by the same team of three or four undergraduate research assistants who scored the Baseline QC textboxes.

A.3.iii. Results of Quality Control Procedure

Table A.4 reports respondent-level quality control variables by fielding group and QC results. Overall, the results reassure us that our quality control procedure led to a sample of respondents who are real humans and putting forth at least a minimum level of effort. Each variable (row) in the table is reported so that a higher number corresponds to a higher level of response quality.

The first three columns are three respondent groups from before the Prescreening survey was implemented (Batches A-D): all who completed Baseline, those who failed Baseline QC, and those who passed Baseline QC. Looking down the rows for these columns, we see that the respondents who passed QC (column 3) have consistently higher quality than those who failed QC (column 2). This relationship must hold by definition for the variables that are directly related to the QC criterion. However, the other response-quality variables show a similar pattern.

Note that the state comparison quality check—whether the respondent's state derived from their IP address matches the state we derived from their ZIP code—has a relatively low percentage of respondents passing the check in all columns. However, we did not expect these percentages to be as high as those for other quality checks. There would be some mismatches even if respondents behaved perfectly, for three reasons: IP address is an approximate geographic variable; ZIP codes do not map uniquely to states; and respondents could be traveling away from home while taking the survey. Still, we see the same patterns of the quality control procedure improving that variable as it improves the others.

The final row is based on our tradeoff data. Here, we report the percentage of respondents who never fail an "opposite-direction trap," i.e., they never choose a decrease in one dimension of SWB over an increase in another dimension of SWB, conditional on the respondent facing at least one opposite direction trap. The Baseline survey is designed so that tradeoff questions only show opposite changes in SWB dimensions with 2% probability. Before the Prescreening survey was implemented, the percentage of Baseline respondents who passed all the opposite-direction traps they faced was roughly twice as high in the passed-QC sample (23.5%, Column 3) than in the failed-QC sample (11.9%, Column 2). After the Prescreening survey was introduced, that percentage rose to 38.7%, for all Baseline respondents (Column 6). Because opposite-direction traps are rare, note that the sample sizes corresponding to the final row for Columns 7 and 8 are very small (N=1 and N=4, respectively). Therefore, the percentage of respondents who never failed an opposite-direction trap is basically unchanged between Column 6 (all who took Baseline after Prescreen) and Column 9 (the subset of those who took Prescreen after the cookie was introduced and passed Baseline QC).

We now turn back to general comparisons across columns. Comparing Column 1 to Column 6, we see the intended effect of introducing the Prescreening survey: the quality level for respondents completing Baseline is better on every metric. The improvement in the Baseline textbox grade is especially striking, with only 25.5% receiving the highest grade before the Prescreening QC criterion was introduced and 90.5% receiving the highest grade after it was introduced.

Column 7 shows the response-quality variables for the group of respondents who passed Prescreening QC but went on to fail Baseline QC. The quality measures in Column 7 are worse than those for all the other groups who passed the Prescreening survey (Columns 6, 8, and 9), with the exception of two variables with no difference (Unique IP address and Prescreening completion time greater than or equal to 3 minutes). This suggests that the Prescreening QC criterion had some "type 2 error" in the sense of inaccurately allowing some low-quality respondents to proceed to Baseline. This also shows that there was value added by the second (Baseline) round of the QC procedure.

Table A.4. QC Variables by Respondent Group

	1	2	3	4	5	6	7	8	9
Respondents who…	**All who took Baseline before Prescreen**	**Failed Baseline QC before Prescreen**	**Passed Baseline QC before Prescreen**	**All Prescreen**	**Failed Prescreen QC**	**All who took Baseline after Prescreen**	**Failed Baseline QC after Prescreen**	**Passed Baseline QC after Prescreen, before cookie**	**Passed Baseline QC after Prescreen, after cookie**
Sample size	3,218	2,250	968	10,781	5,846	2,752	162	221	2,369
Variables related to QC criterion									
Unique IP Address	86.7%	82.5%	96.5%	98.1%	96.9%	99.6%	100.0%	100.0%	99.6%
Not using a virtual private server	90.7%	87.6%	98.0%	95.7%	93.3%	98.5%	95.1%	97.3%	98.9%
Best possible textbox grade, Prescreening survey				41.2%	0.0%	90.6%	65.4%	85.1%	92.8%
Best possible textbox grade, Baseline	25.5%	0.4%	83.9%			90.5%	1.2%	91.0%	96.5%
Baseline completion time >= 10 minutes	96.9%	95.5%	100.0%			99.9%	98.8%	100.0%	100.0%
Other response quality variables									
Prescreening completion time >= 3 minutes				83.0%	74.9%	93.4%	93.2%	93.2%	93.4%
State derived from IP Address matches state derived from zip code	20.2%	6.1%	56.4%			78.6%	50.4%	65.5%	81.8%
CQ responses have correct ranking of high and low stimuli: region size trio	73.7%	64.1%	95.9%			97.6%	88.3%	96.8%	98.4%
CQ responses have correct ranking of high and low stimuli: darkness of circle trio	74.2%	64.6%	96.4%			97.2%	87.0%	97.3%	97.8%
CQ responses have correct ranking of high and low stimuli: curvy line trio	68.4%	58.4%	91.7%			94.5%	84.6%	93.7%	95.2%
CQ responses have correct ranking of high and low stimuli: 'your access to information'	56.1%	47.6%	75.6%			84.8%	70.4%	80.5%	86.2%
CQ responses have correct ranking of high and low stimuli: 'your ability to remember'	56.3%	47.2%	77.5%			81.3%	68.5%	79.2%	82.4%
CQ responses have correct ranking of high and low stimuli: 'your living area not being spoiled by crime and violence'	53.8%	46.1%	71.7%			78.8%	69.8%	73.3%	79.9%
Test-retest correlation of repeated SWB ratings within Baseline survey (3 per person)	0.73	0.57	0.83			0.84	0.74	0.86	0.84
Never failed an "opposite direction trap" in tradeoff questions (choosing a decrease in one direction of SWB instead of an increase in another dimension of SWB)	15.3%	11.9%	23.5%			38.7%	0.0%	50.0%	38.6%

We can also evaluate the success of our quality control procedure by comparing three of our quality metrics to similar metrics from Douglas et al. (2023), who compare respondent quality between five online survey platforms: MTurk, Prolific, CloudResearch, SONA, and Qualtrics. They construct a variable that indicates a high-quality respondent based on 11 measures: five attention checks, unique worker ID, unique IP address, unique geolocation, meaningful or blank open response, sufficient time taken, and self-reported high data quality. Here we will just discuss their MTurk and Prolific results.

Douglas et al. find that 97.2% of respondents have unique IP addresses in their MTurk sample and 98.8% have unique IP addresses on Prolific. As seen in Table 4, our percentage of Baseline respondents with a unique IP address increases from 86.7% without any QC procedure (Column 1) to 96.5% with the Baseline QC procedure (Column 3) and 99.6% with the Prescreening survey and cookie (Column 9). For Douglas et al.'s open response question, a graduate student RA and an undergrad RA both graded the answer to "Do you have any additional comments for us?" A worker's textbox is classified as meaningful if (a) both graders rate the answer as meaningful, or (b) the worker leaves the textbox blank. Otherwise the textbox is classified as meaningless. On MTurk, 82.2% of their respondents have a blank or meaningful response. On Prolific, it is 99.2%. We believe that a high-quality response is more difficult for our textbox questions, since we ask about question interpretation and we do not accept blank responses (a blank response is assigned the worst score of 3 points). Nevertheless, looking at the percentage of respondents with the best possible textbox grade on Baseline, we see an increase from 25.5% before any QC procedure (Column 1) to 83.9% with the Baseline QC (Column 3) and 96.5% after the Prescreening survey and cookie (Column 9).

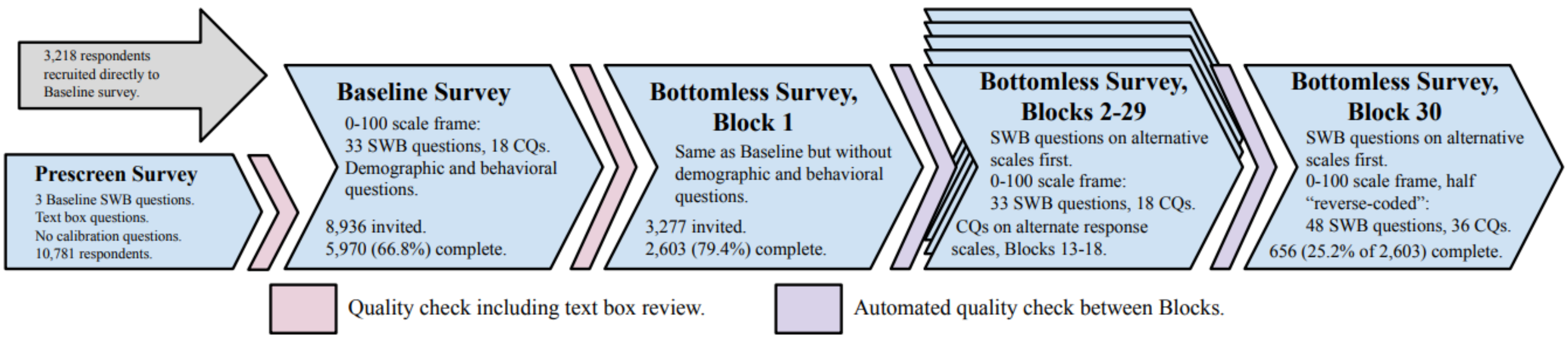
Finally, Douglas et al. asked participants the same survey question twice, 21 questions apart in the survey, to assess the test-retest correlation of participant responses. They find a correlation of 0.54 for MTurk and 0.87 for Prolific. Recall that in our Baseline survey, the 11th triple is a repeat of the 2nd triple as a check for data quality; this means that each respondent rates three SWB questions twice within the same survey. We find that the test-retest correlation is 0.73 before any QC procedure (Column 1), 0.83 after the Baseline QC (Column 3), and 0.84 after the Prescreening QC and cookie (Column 9). Given these three comparisons (though they are admittedly rough), we conclude that our survey sample's quality is closer to Douglas et al.'s Prolific sample quality than their MTurk quality.

A.4. Bottomless Survey Description

A.4.i. Design

We invited qualified respondents—those who completed Baseline, passed quality control, and were willing to be contacted for follow-up surveys—to several follow-up surveys. This paper utilizes a subset of data from one of the follow-up surveys that we refer to as the "Bottomless" survey because it is very long. The Bottomless survey is arranged in 30 sequential "blocks," each of which corresponds to a single HIT on MTurk. Figure A.3 shows the progression of respondents through the Prescreening, Baseline, and Bottomless survey, with an overview of the survey content.

Figure A.3 Overview of surveys and recruitment flow



Block 1 of the Bottomless survey is identical to the Baseline survey, except it has abbreviated and slightly different instructions (described further below) and no demographic/behavioral or exit questions at the end. Blocks 2-29 begin either with one or more SWB questions on "alternative scales" (with response options other than the 0-100 scale from the Baseline survey) or (at the start of Blocks 8 and 9) questions about remembered or expected SWB at ages 50 and 75; the SWB questions come from the European Social Survey (ESS), General Social Survey (GSS), Health and Retirement Study (HRS), the global life satisfaction question in Kapteyn, Smith, and van Soest (2009) (KSV), World Values Survey (WVS), Gallup World Poll, and UK Office of National Statistics (ONS). The details of these questions, including their block location in the Bottomless survey, can be found in Appendix K.4. In Figures 7 and 8 of the main text, we compare the mean or standard deviation of these SWB ratings on alternative scales to the mean or standard deviation, respectively, of CQ ratings on the 0-100 scale. We describe the 388 CQs used in those analyses (and others) in A.4.iii. below.

In Blocks 2-29, after the survey screen with SWB questions on alternative scales, each block contains CQs, SWB questions, and tradeoff questions that have the same structure as the Baseline survey but with different question details. The 33 SWB questions partially vary by block; the 16 SWB questions that are starred in Table A.1 are repeated in every block, while the other 17 questions are varied. The new SWB questions in each block are listed in Appendix K2. The 18 CQs (6 trios) in each block are fixed across respondents, but we randomize the order of the trios within block (first randomizing whether vignettes or visuals come first, and then randomizing trios within CQ category) and also randomize the order of CQs within trio. We describe the CQs in more detail below.

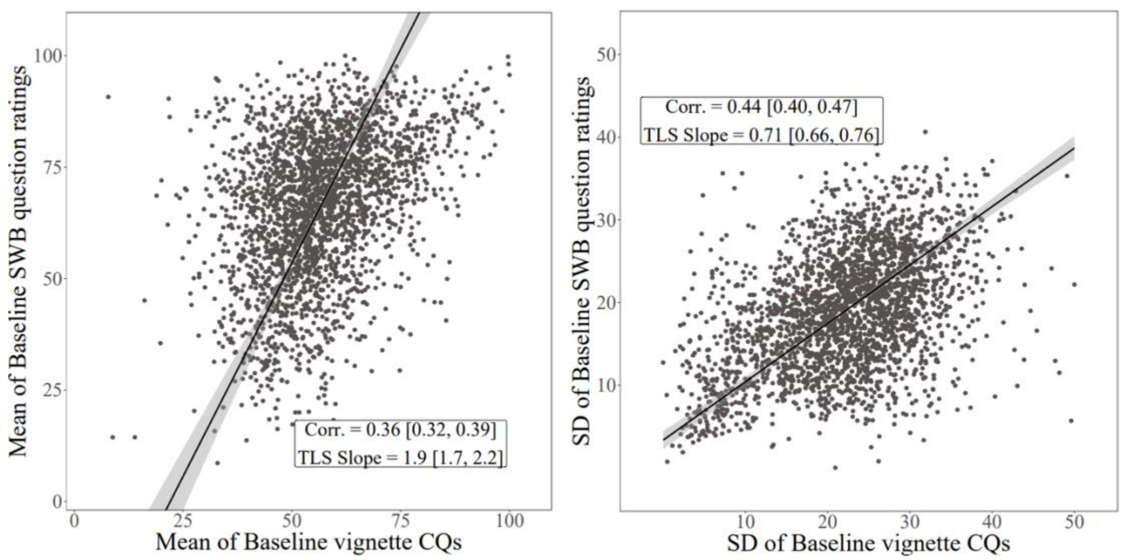
Block 30 begins with the five Satisfaction With Life Scale (SWLS) questions from Diener et al. (1985). The rest of the block includes 48 SWB questions and 36 CQs (and no tradeoff questions) to explore to which extent "reverse-coding" of SWB questions may help researchers deal with general-scale-use heterogeneity (for example, rating "How anxious you feel" instead of "You not feeling anxious"). Data from Block 30 are not analyzed in this paper. In particular, we did not include the CQs in our analyses because we hypothesize that scale use may be different for reverse-coded questions or immediately after completing a reverse-coded question.

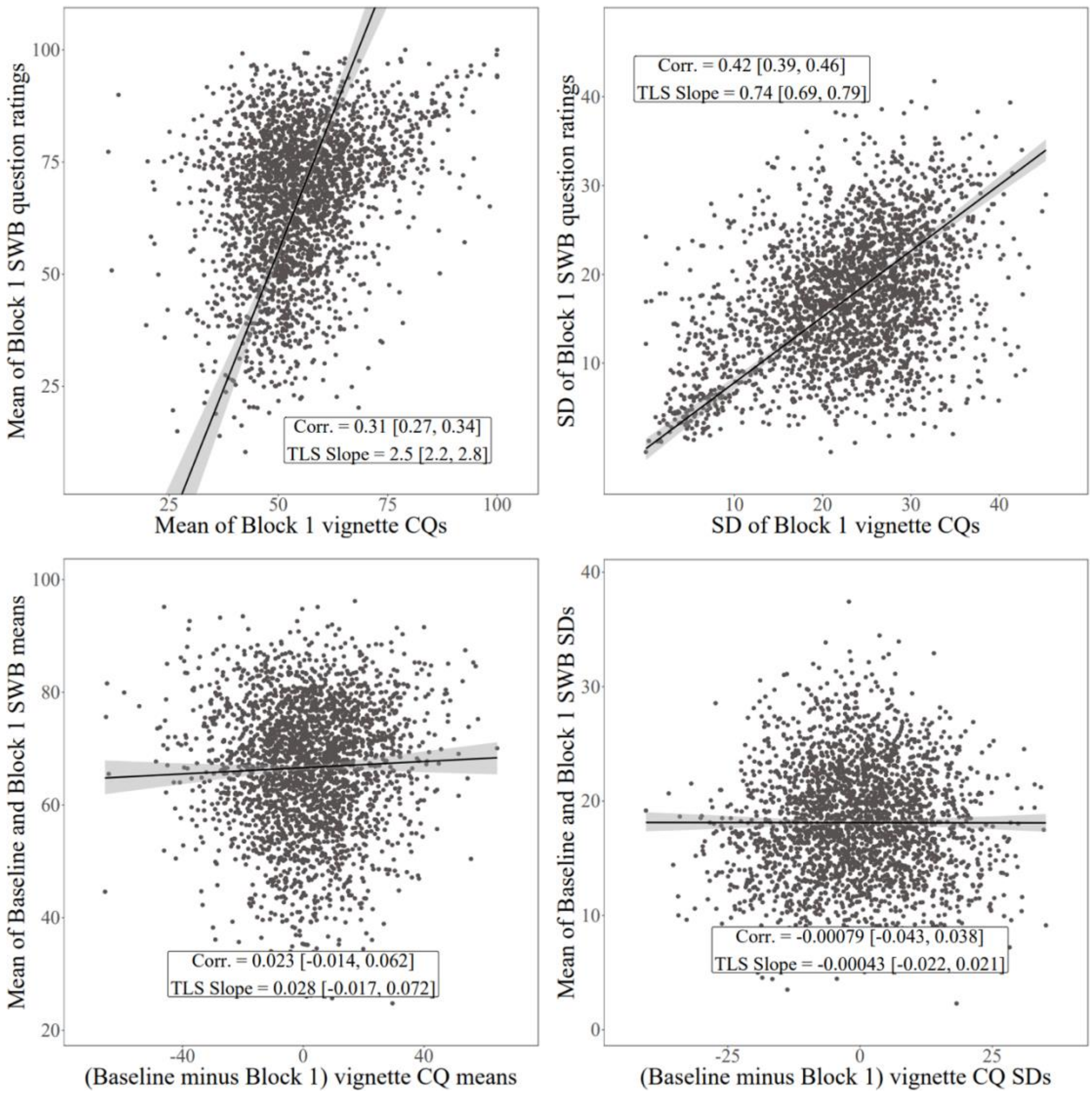
Screenshots of all 558 CQs in the Bottomless survey (18 per block in Blocks 1 through 29 plus 36 in Block 30) can be found in Appendix K2. Here, we briefly describe two features of vignettes in the Bottomless survey that differ from those in Baseline: survey instructions and randomization of the vignette subject's gender and age.

First, as the note in Figure 3 of the main text explains, the instructions at the top of each vignette CQ screen in the Baseline survey ask respondents to "Imagine everything in your life is the same as it is now, except for the details described in each situation below." We included this instruction because we anticipated that respondents would find it difficult to accurately imagine a situation very far from their own life. However, this instruction could induce violations of Assumption 2 of our statistical model: if respondents are "filling in the blanks" about the vignette person's life with their own life details, then the underlying state ω_c that is being rated will not be the same across individuals. In the Bottomless survey, we instruct respondents with a different preamble at the top of the screen: "In this set of questions, you will rate situations in other people's lives. Try to use the scale as you would if you were rating that situation for yourself. (But remember, 'you' or 'your' refers to the other person.)" We call this instruction framing "second-person-other."

In Figure A.4, we test for an effect of the instruction wording on general scale use, using the Figure 4 analysis from the main text. The first row replicates Figure 4, limited to the sample of respondents who did the first block of the Bottomless survey; the second row uses data from Block 1 of Bottomless (with the second-person-other instructions); the third row uses the difference of Baseline and Bottomless Block 1 ratings, as the x-axis variable. While the point estimates suggest slightly higher correlations for the CQ ratings in Baseline for both mean and SD, neither correlation is statistically distinguishable between the CQ ratings in Baseline and Bottomless Block 1. Thus, the Baseline instructions appear to induce little or no additional bias relative to the Bottomless Block 1 instructions. Since we anticipate that instructing participants to "fill in the blanks" with their own life details generates a bias relative to Assumption 2, we conclude that this bias is similarly present even when participants are not explicitly instructed to do so.

Figure A.4: The effect of vignette instruction wording on general scale use





The second variation from Baseline is that, for 38 of the vignette dimensions studied in this paper, we have at least one trio where the gender of the person described by the vignette is randomized at the trio level: male, female, or second-person-other. A summary of the instructions, vignette text, and question prompt for each type of vignette is shown in Table A.5. Our strategy for choosing male and female names was the following: (i) use the most common names, as reported by the U.S. Social Security Administration, from the birth year corresponding to the average age of MTurk respondents in our pilot data (age 37, birth year 1983: https://www.ssa.gov/oact/babynames/state/top5_1983.htm); (ii) select a subset of names which seemed as much as possible, a priori, to have racial/ethnic neutrality and unambiguous gender connotation; (iii) omit names which were already used in vignettes from other sources (six names from our pilot CQs (Juan, Peter, Sasha, Rachel, James, and Taylor), and 24 names from Kapteyn et al. (2009) vignettes). We randomly assigned the names to CQs. Table A.6 shows an example of how the text for a given vignette would be modified for male and female versions.

Table A.5: Instructions and prompts for vignette CQs

Survey location and vignette type	**Preamble (top of screen)**	**Vignette text**	**Prompt**
Baseline survey, “self” vignettes	In this set of questions, you will rate situations that are different from the situation in your life. Imagine everything in your life is the same as it is now, except for the details described in each situation below.	[Vignette_self]	If this situation described your life during the past year, how would you rate your level of [***dimension***]?
Bottomless survey, “second-person-other” vignettes	In this set of questions, you will rate situations in other people's lives. Try to use the scale as you would if you were rating that situation for yourself. (But remember, “you” or “your” refers to the other person.)	[Vignette_self]	If this situation described your life during the past year, how would you rate your level of [***dimension***]?
Bottomless survey, male vignettes		[Name_male] is [AGE] years old. [Vignette_male]	Thinking about the past year, how would you rate the level of [***dimension***] in [Name_male]’s life?
Bottomless survey, female vignettes		[Name_female] is [AGE] years old. [Vignette_female]	Thinking about the past year, how would you rate the level of [***dimension***] in [Name_female]’s life?

Table A.6: Example of vignette text and prompt for all vignette types in Bottomless survey

Vignette type	**Vignette text and prompt for the dimension *Your ability to remember things* (Low level)**
Second-person-other	Your friends joke about how forgetful you are. You do often get lost, even in places you should know well. It’s difficult for you to remember names, and you often lose track of things. But you usually remember to set reminders for important appointments. If this situation described your life during the past year, how would you rate your level of ***Your ability to remember things***?
Male	Shaun is [AGE] years old. His friends joke about how forgetful he is. He does often get lost, even in places he should know well. It’s difficult for him to remember names, and he often loses track of things. But he usually remembers to set reminders for important appointments. Thinking about the past year, how would you rate the level of ***Your ability to remember things*** in Shaun’s life?

Female	Barbara is [AGE] years old. Her friends joke about how forgetful she is. She does often get lost, even in places she should know well. It's difficult for her to remember names, and she often loses track of things. But she usually remembers to set reminders for important appointments. Thinking about the past year, how would you rate the level of ***Your ability to remember things*** in Barbara's life?

Finally, as Tables A.5 and A.6 show, for almost all vignettes with a gender randomization, we also randomize the vignette subject's age. The age is randomized at the CQ level, drawn from integers 22 to 80, inclusive, with uniform probability. The only vignettes with gender randomized but no age randomization are the 12 vignettes we drew from Kapteyn et al. (2009); we tried to match their questions as closely as possible, and they did not randomize the vignette person's age.[4] There are no CQs with an age randomization but no gender randomization. In this paper, all of our results with Bottomless survey CQs are pooled across vignette types and ages.

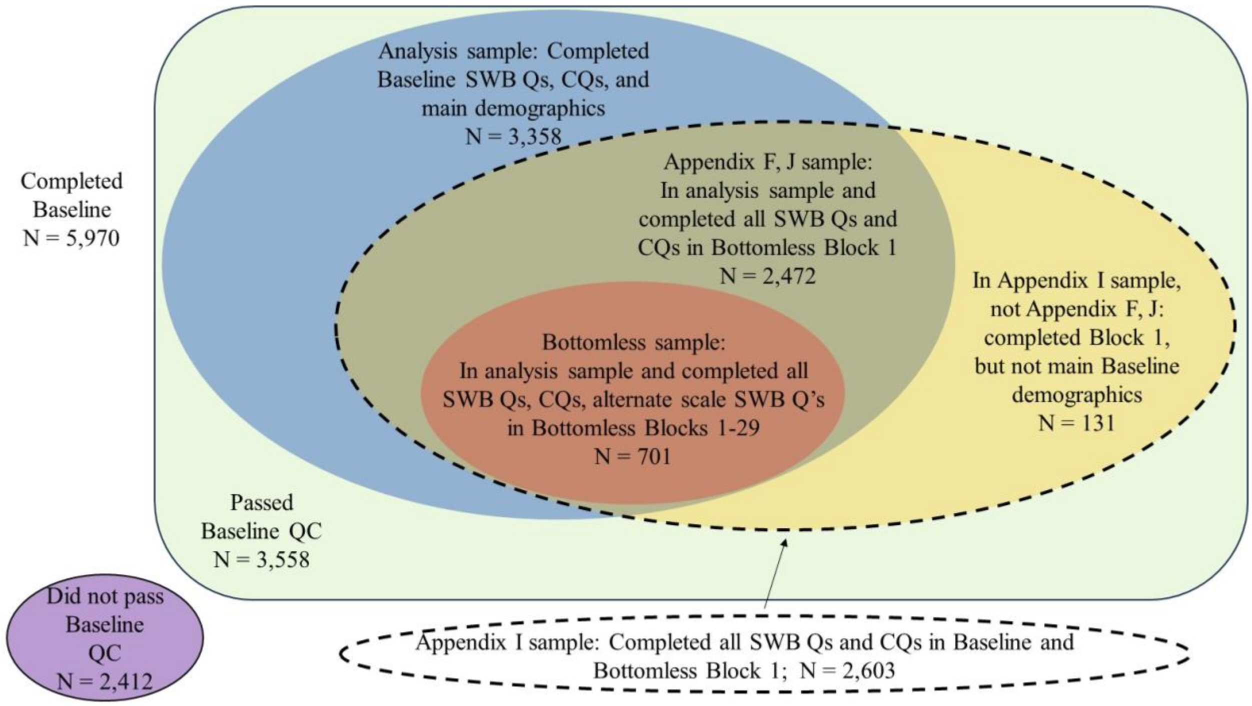
A.4.ii. Bottomless Survey Fielding

Using Amazon's MTurk platform, we collected data on the Bottomless survey between September 20, 2022, and January 3, 2023. The median length of time between completing Baseline and the first block of Bottomless survey was about 7 weeks (46 days), with a range from 6 days to 164 days. Respondents were compensated $1.50 per Bottomless block completed, with a bonus of $15 if all 30 blocks were completed. The median completion time for Block 1 was 18 minutes; Blocks 2-29, 13 minutes; and Block 30, 13 minutes. Of the 3,277 respondents who passed Baseline quality control and were invited to follow-up surveys, N = 2,603 (79.4 percent) completed the first block of Bottomless. (Note that in Appendix F, which is described in section VII.C. of the main text, and Appendix J, we limit the Block 1 sample to the 2,472 respondents who also completed all the main demographic variables in the Baseline survey. We do not exclude respondents based on completion of demographic variables in the other analyses of Bottomless data, which do not require demographic variables (Figures 7 and 8 and Appendices B, H, and I)). Of the respondents who completed Block 1, N = 701 (26.9 percent)

[4] We also include five vignettes about political efficacy, based on King et al. (2004), which include a name but no age, and the name is not randomized. In order to use the same trio design for all CQs, we wrote a sixth vignette about the political efficacy dimension, with the same format as the others. As seen in Table A.7 below, the wording of our dimension for political efficacy is, "You having a say in getting the government to address issues that interest you."

completed the survey through Block 29; this is the sample analyzed in Figures 7 and 8 and Appendices B, H and I. For completeness, we note that $N = 656$ (25.2 percent) completed the survey through all 30 blocks, taking an average of 7.2 hours to complete them. Figure A.5 illustrates the survey samples used for the paper and appendices.

Figure A.5: Survey Samples



A.4.iii. Calibration Questions from Bottomless Survey

The set of 388 CQs used in Figures 7 and 8 of the main text, and in analyses in Appendix B and Appendix H, consists of all the CQs from Blocks 1-29, with the following exceptions. In Blocks 13-18, we asked the Baseline CQs on alternative scales; we exclude these because this paper only analyzes CQs on the 0-100 scale. We exclude the CQs from Block 19, which are visual CQs intended for exploratory analysis and as “placebos” (with intentionally extreme stimuli). We exclude a trio of visual CQs about the cuteness of a dog, in Block 12, because we concluded a priori that it was not likely to satisfy the assumption that the true states are the same; there may be systematic differences in perceived cuteness by “dog people” vs. “cat people” that are correlated with many other variables. Finally, we excluded five CQs that had technical problems during fielding.[5]

The 388 CQs that we analyze from the Bottomless survey cover 60 dimensions in total; see Table A.7 for the list and categorization. The visual category has ten dimensions; the category of non-local public goods has eight dimensions; there are three dimensions of local public goods; the remaining 39 dimensions in the vignette category relate directly to personal well-being. In section VII.B of the main text (Relative Importance of General-Scale-Use Heterogeneity) and its companion Appendix H, we restrict our analysis further to 42 dimensions of CQs, excluding visual dimensions and non-local public goods. In Table A.7, these 42 dimensions are the ones in the vignette and local public good categories. The number of CQs for a given dimension ranges from 3 (just one trio) to 27 (nine trios for the vignette dimension, "How satisfied you are with your life"). See Appendix K.3 screenshots for detailed wording of vignettes and visual CQ images.

Table A.7: 60 CQ Dimensions from Bottomless Survey

Dimension	Category of Dimension	Number of CQs in Dimension	Number of trios with gender/age randomization	Number of unique trios*
Your access to information	Vignette	15	1	4
Your ability to remember things	Vignette	9	1	2
Your ability to breathe in and out easily	Vignette	5	1	2
You and your family having enough to eat	Vignette	2	1	1
The quality and quantity of green spaces in your area	Local public good	6	0	2
You having people you can turn to in time of need	Vignette	3	0	1
Your ability to hear	Vignette	6	2	2
Your knowledge and skills	Vignette	3	1	1
Your physical health	Vignette	9	2	2
You having many options and possibilities in your life and the freedom to choose among them	Vignette	3	1	1
You being able to rise to the challenges you face	Vignette	6	2	2

[5] For part of our sample, the survey text did not appear correctly for the "you not feeling anxious" trio in Block 24. And for part of the sample, the survey text did not appear correctly for two other individual CQs: the "low" option from the "having enough to eat" trio (Block 11) and the "low" option from the "ability to breathe" trio (Block 20). We exclude these five CQs from analysis.

Your ability to see	Vignette	6	2	2
Your sense of control over your life	Vignette	9	2	2
Your sense of purpose	Vignette	9	2	2
You being able to sleep well at night	Vignette	6	2	2
Your ability to walk several blocks	Vignette	6	2	2
Your living environment not being spoiled by crime and violence	Local public good	6	0	2
The air in your area not being polluted	Local public good	3	0	1
The absence of anger in your life	Vignette	9	3	3
You not feeling anxious	Vignette	18	5	6
You not being worried about money	Vignette	6	2	2
You not being trapped in physical pain	Vignette	6	2	2
The absence of sadness in your life	Vignette	9	2	2
The absence of stress in your life	Vignette	9	2	3
The absence of worry in your life	Vignette	3	1	1
You not being lonely	Vignette	9	2	2
You not having to worry about being unemployed	Vignette	3	1	1
How much you enjoy your life	Vignette	9	2	2
The happiness of your family	Vignette	9	2	2
You being able to support your family financially	Vignette	9	2	2
You being a good person	Vignette	9	2	2
How happy you feel	Vignette	9	2	2
Your health	Vignette	6	2	2
Your home being comfortable	Vignette	3	1	1
Your rating of your life on a ladder where the lowest rung is worst possible life for you and the highest rung is best possible life for you	Vignette	9	2	2
Your mental health	Vignette	15	4	4
Your physical safety and security	Vignette	3	1	1
How satisfied you are with your life	Vignette	27	Gender: 8, Age: 4**	8
You being a winner in life	Vignette	3	1	1
You feeling that you have enough time for the things that are most important to you	Vignette	9	2	2
The overall wellbeing of you and your family	Vignette	3	1	1
The extent to which you feel the things you do in your life are worthwhile	Vignette	9	2	2
Curved	Visual	6	0	2
Dark	Visual	6	0	2
Big (region size)	Visual	6	0	2
Confident	Visual	3	0	1
Big (country size)	Visual	6	0	2
Large (diamond culet)	Visual	3	0	1
Complex	Visual	3	0	1

Sharp	Visual	3	0	1
Symmetric	Visual	3	0	1
Absence of pain (visual)	Visual	3	0	1
The ability of ordinary citizens to influence your national government	Non-local public good	3	0	1
How much you can trust most people in your nation	Non-local public good	3	0	1
Your cultures and traditions being honored	Non-local public good	3	0	1
Freedom of the press in your nation	Non-local public good	3	0	1
The condition of the natural environment	Non-local public good	3	0	1
You having a say in getting the government to address issues that interest you	Non-local public good	6	0	2
The leaders of your state government not being corrupt	Non-local public good	6	0	2
How little violence there is in the world	Non-local public good	3	0	1

* For some trios, we asked the second-person-other vignette type in one block and asked the male/female vignette type (randomized) in another block. When this happens, we only count one "unique" trio. However, we still count this as 6 CQs for the Number of CQs in Dimension column (given that we are pooling across types in our analyses).
** As noted in section A.4.i, there are 4 trios (12 CQs) for the dimension, "How satisfied you are with your life," which have a randomization of gender but not age, since these vignettes come from Kapteyn et al. (2009). For all other dimensions, a CQ either has both age and gender randomization or neither.

Appendix A References

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

Benjamin, Daniel J., Ori Heffetz, Miles S. Kimball, and Nichole Szembrot. 2014. "Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference." *American Economic Review* 104 (9): 2698–2735.

Diener, Ed. Robert A. Emmons, Randy J. Larsen, and Sharon Griffin. 1985. "The Satisfaction With Life Scale." *Journal of Personality Assessment* 49 (1): 71-75.

Douglas, Benjamin D., Patrick J. Ewell, and Markus Brauer. 2023. "Data Quality in Online Human-Subjects Research: Comparisons Between MTurk, Prolific, CloudResearch, Qualtrics, and SONA." *PLoS ONE* 18 (3): e0279720.

Kapteyn, Arie, James P. Smith, and Arthur van Soest. 2009. "Life Satisfaction." *IZA Discussion Paper* No. 4015.

Kennedy, Ryan, Scott Clifford, Tyler Burleigh, Philip D. Waggoner, Ryan Jewell, and Nicholas J. G. Winter. 2020. "The Shape of and Solutions to the MTurk Quality Crisis." *Political Science Research and Methods* 8 (4): 614-629.

King, Gary, Christopher J.L. Murray, Joshua A. Salomon, and Ajay Tandon. 2004. "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research." *The American Political Science Review* 98 (1): 191–207.

Peer, Eyal, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer. 2022. "Data Quality of Platforms and Panels for Online Behavioral Research." *Behavior Research Methods* 54 (4): 1643-1662.

B. Evidence on the Linearity of the Translation Function

Our translation function (equation 4) assumes that the observed CQ rating r_{ic} is linear in the common-scale CQ rating w_c. To gauge the plausibility of the linearity assumption, we extend the translation function with a quadratic term:[6]

$$r_{ic} - \gamma = \alpha_i + \beta_i(w_c - \gamma) + \delta_i(w_c - \gamma)^2 + \beta_i\epsilon_{ic} + \eta_{ic}.$$

As in the linear translation function, we assume $\alpha_i \sim \mathcal{N}(0, \sigma_\alpha^2)$, $\beta_i \sim \mathcal{N}\left(1, \sigma_\beta^2\right)$, $\epsilon_{ic} \sim \mathcal{N}(0, \sigma_\epsilon^2)$, and $\eta_{ic} \sim \mathcal{N}\left(0, \sigma_{\eta_i}^2\right)$, all mutually independent. To be consistent with w_c's definition as the population mean of respondents' rating of CQ c, we assume that δ_i has mean zero in the population. In Section B.2 (but not B.1), we additionally assume $\delta_i \sim \mathcal{N}\left(0, \sigma_\delta^2\right)$, independent of everything else. The closeness of our estimate of σ_δ to its theoretical lower bound, zero, serves as a test of the existence and importance of a quadratic term in the translation function and hence a gauge of the curvature of the translation function.

B.1. OLS Evidence

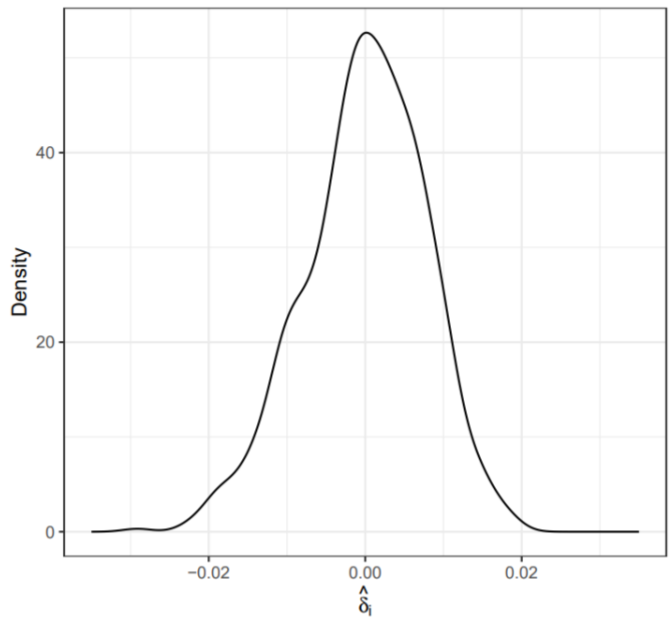
We can obtain an estimate of σ_δ by the standard deviation of estimates of δ_i's; this estimator is biased upward by estimation error, but this bias can be adjusted for—and will be small if the number of CQs is large. The δ_i's can be estimated by a respondent-level OLS regression of r_{ic} on w_c and w_c^2 with a constant term,[7] where w_c is estimated by the mean rating ($\sum_i r_{ic} / I$). The coefficient on w_c^2 estimates δ_i.

In a subsample from our Bottomless survey where 701 respondents completed more than 388 analysis CQs,[8] the estimated δ_i's have a mean of -5.57×10^{-18} and a standard deviation of

[6] We do not also include squared error terms because they would have non-normal distributions, substantially complicating the MLE.
[7] There is no need to subtract γ from w_c in the regression because the terms related to γ will be subsumed into the constant term and the w_c term.
[8] See Appendix A.4 for the process of arriving at the 388 analysis CQs.

0.0077; after adjusting for estimation errors, the standard deviation becomes 0.0071. A nonparametric density plot (Figure B.1) of the estimated δ_i's suggests that the distributional assumption of normality does not hold exactly but may be a reasonable approximation.

Figure B.1: Density plot of $\hat{\boldsymbol{\delta}}_{\boldsymbol{i}}$



Note: The depicted density plot is constructed using a Gaussian kernel estimator (with bandwidth selected by Silverman's 'rule of thumb' procedure, as implemented by R's bw.nrd0 function).

B.2. MLE Evidence

To formally and more efficiently estimate σ_δ, we utilize maximum likelihood estimation with a likelihood function constructed from the following sub-likelihood function, evaluated on the data of each respondent's CQ responses ($\mathbf{r}_{i\mathcal{C}} \equiv \{r_{ic}\}_{c=1}^{C}$):

$$l_i\left(\boldsymbol{r}_{i\mathcal{C}} \mid \alpha_i, \beta_i, \delta_i, \gamma, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \sigma_{\eta_i}\right)$$
$$= \phi\left(\left[\boldsymbol{r}_{i\mathcal{C}} - (\alpha_i + \gamma)\mathbf{1}_C - \beta_i(\boldsymbol{w}_{\mathcal{C}} - \gamma\mathbf{1}_C) - \delta_i(\boldsymbol{w}_{\mathcal{C}} - \gamma\mathbf{1}_C)^2\right]\left(\beta_i^2\sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-\frac{1}{2}}\right)\left(\beta_i^2\sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-\frac{C}{2}},$$

where ϕ is the multivariate standard-normal density function, $\mathbf{1}_C$ is the C-dimensional vector full of ones, and $\boldsymbol{w}_{\mathcal{C}} \equiv [w_1, \ldots, w_C]'$.

Assuming independence across respondents, the log-likelihood function for CQ ratings of all respondents ($\boldsymbol{r}_{\cdot\mathcal{C}} \equiv \{\boldsymbol{r}_{i\mathcal{C}}\}_{i=1}^{I}$) is thus

$$\begin{aligned}
&\ln\mathcal{L}\left(\boldsymbol{r}_{\cdot\mathcal{C}} \mid \sigma_\alpha, \sigma_\beta, \sigma_\delta, \gamma, \boldsymbol{w}_\mathcal{C}, \sigma_\epsilon, \mu_{\ln\sigma_\eta}, \sigma_{\ln\sigma_\eta}\right) \\
= \; &\sum_i \ln \iiiint l_i\left(\boldsymbol{r}_{i\mathcal{C}} \mid \alpha_i, \beta_i, \delta_i, \gamma, \boldsymbol{w}_\mathcal{C}, \sigma_\epsilon, \sigma_{\eta_i}\right) f(\alpha_i) f(\beta_i) f(\delta_i) f\left(\sigma_{\eta_i}\right) d\alpha_i d\beta_i d\delta_i d\sigma_{\eta_i},
\end{aligned}$$

where f denotes a probability density function. More explicitly,

$$\begin{aligned}
&\ln\mathcal{L}\left(\boldsymbol{r}_{\cdot\mathcal{C}} \mid \sigma_\alpha, \sigma_\beta, \sigma_\delta, \gamma, \boldsymbol{w}_\mathcal{C}, \sigma_\epsilon, \sigma_{\eta_i}\right) \\
= \; &\sum_{i=1}^{N} \ln\Bigg[\int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty (2\pi)^{-\frac{C}{2}} \left(\beta_i^2\sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-\frac{C}{2}} \\
&\cdot \exp\left(-\frac{1}{2}[\boldsymbol{r}_{i\mathcal{C}} - (\alpha_i + \gamma)\mathbf{1}_C - \beta_i(\boldsymbol{w}_\mathcal{C} - \gamma\mathbf{1}_C) - \delta_i(\boldsymbol{w}_\mathcal{C} - \gamma\mathbf{1}_C)^2]'\left(\beta_i^2\sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-1}\right. \\
&\left.\cdot [\boldsymbol{r}_{i\mathcal{C}} - (\alpha_i + \gamma)\mathbf{1}_C - \beta_i(\boldsymbol{w}_\mathcal{C} - \gamma\mathbf{1}_C) - \delta_i(\boldsymbol{w}_\mathcal{C} - \gamma\mathbf{1}_C)^2]\right) \frac{1}{\sqrt{2\pi}\sigma_\alpha} \exp\left(-\frac{\alpha_i^2}{2\sigma_\alpha^2}\right) \\
&\cdot \frac{1}{\sqrt{2\pi}\sigma_\beta} \exp\left(-\frac{(\beta_i - 1)^2}{2\sigma_\beta^2}\right) \frac{1}{\sqrt{2\pi}\sigma_\delta} \exp\left(-\frac{\delta_i^2}{2\sigma_\delta^2}\right) \frac{1}{\sqrt{2\pi}\sigma_{\eta_i}\sigma_{\ln\sigma_\eta}} \exp\left(-\frac{\left(\ln\sigma_{\eta_i} - \mu_{\ln\sigma_\eta}\right)^2}{2\sigma_{\ln\sigma_\eta}^2}\right) d\alpha_i d\beta_i d\delta_i d\sigma_{\eta_i}\Bigg].
\end{aligned}$$

As in the estimation of our main specification, we use hierarchical modeling to deal with the computational burden of the multidimensional numerical integration.

In the subsample mentioned above ($I = 701$), the MLE estimate of σ_δ is 0.0068 with a standard error of 0.00057.[9] Quantitatively, the point estimate implies that, when w_c is 10 points higher, the quadratic term contributes less than 1.36 points to the raw rating for 95% of respondents. The MLE estimate of σ_δ based on our Baseline data (which has far fewer CQs but many more respondents) is 0.000556 and not statistically distinguishable from zero (SE = 0.00116). This estimate suggests that the quadratic coefficient in the translation function is even smaller.

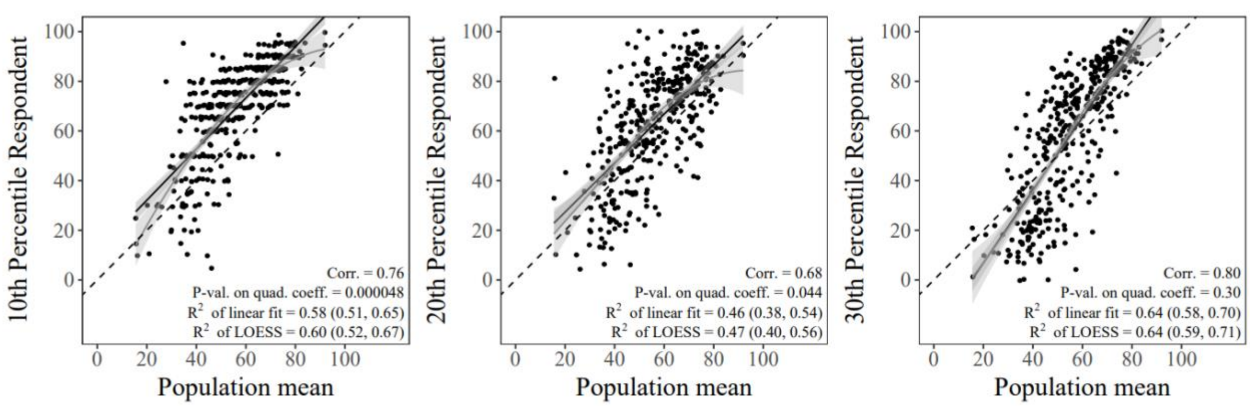
[9] The MLE estimate could be biased away from zero (making it an upper-bound estimate) due to the positivity constraint, although this bias is likely small because the estimate is quite precise.

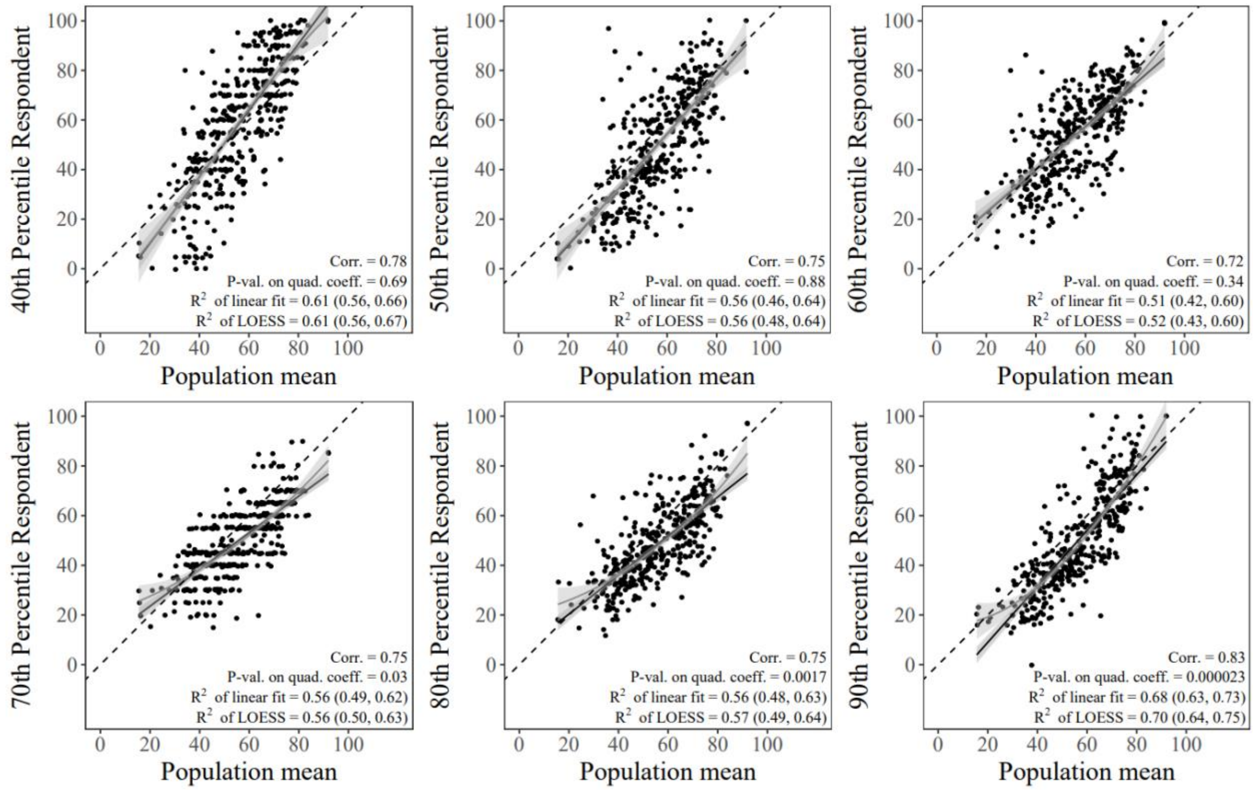
B.3. Graphical Evidence

To provide visual evidence for the assumption of linearity in the translation functions, we plot respondent-level 388 Bottomless CQ ratings (y-axis) against the sample means of these CQs. To keep the figure manageable, we randomly draw nine respondents from the 1st through 9th deciles of the estimated δ_i.

We use both an OLS line and a nonlinear LOESS curve to fit the points in each plot. The proximity of the nonlinear LOESS curve to the OLS line as shown in Figure B.2 makes clear how well linearity approximates the translation function. In every plot shown, including those for respondents with the most concave translation functions (shown in the top-left plot, corresponding to the 1st decile) and the most convex translation functions (shown in the bottom-right plot, corresponding to the 9th decile), there is considerable overlap between the LOESS curves and the OLS lines. Moreover, the correlations between individual ratings and the population means remain high even for relatively more nonlinear data (in fact, the correlations for the plots of the individuals from deciles 1 and 9 turn out to be higher than the correlations from deciles 2, 5, 6, 7, and 8). Furthermore, for all the individuals shown, the R^2 of the linear fit is nearly equal to (and well within the 95% confidence interval of) the R^2 value of the LOESS regression.

Figure B.2: Respondent-Level Translation Functions





Notes: The sample is 701 respondents who completed all relevant CQs. Each point on a graph is one of 388 Bottomless survey CQs (see Web Appendix A.4 for details on the CQs excluded for this analysis). Each CQ is rated on a 0-100 scale. *x*-axis: population mean rating of each CQ. *y*-axis: CQ rating for each of the nine respondents at the 10^{th}, 20^{th}, …, and 90^{th} percentiles of the population indicated in *y*-axis title, where the population is ordered by a respondent's squared-term coefficient in a regression of that respondent's CQ ratings on the population means of the CQ ratings and the squared population means. Dashed line: 45-degree line. Solid black line: OLS regression; dark gray region: 95% CI. Gray curve: LOESS regression of respondent's CQ ratings on population means; light gray region: 95% CI. Correlations are reported in each plot.

C. Biases From Not Adjusting for Scale-Use Heterogeneity

This appendix contains derivations for the (asymptotic) biases resulting from not adjusting for scale-use differences. We will focus on the biases for estimating the common-scale SWB's first and second moments of interest. We will also discuss the biases arising from the additional assumptions of our MOM estimators, such as the independence between the stretcher and the common-scale SWB.

C.1. Bias in Estimating $E(w_{is})$

When ignoring scale use—or when correcting for scale use but ignoring the dependence between β_i and w_{is}—the estimator is $E(r_{is})$ (for a derivation, see Appendix E.1). Under our translation function (equation 4),

$$
\begin{aligned}
E(r_{is}) &= E(\alpha_i + (1-\beta_i)\gamma + \beta_i(w_{is} + \epsilon_{is}) + \eta_{is}) \\
&= E(\beta_i w_{is}) \\
&= Cov(\beta_i, w_{is}) + E(w_{is}),
\end{aligned}
$$

where the last equality holds because $E(\beta_i) = 1$. The bias is thus

$$
E(r_{is}) - E(w_{is}) = Cov(\beta_i, w_{is}).
$$

If β_i and w_{is} are uncorrelated, not correcting for scale-use differences does not bias the estimation for $E(w_{is})$. When such a correlation exists, however, the estimator that ignores scale use and the MOM estimator for $E(w_{is})$ are biased, but the semi-parametric estimator and the comprehensive MLE estimator can correct the bias.

C.2. Bias in Estimating $Cov(x_i, w_{is})$

C.2.1. When Ignoring Scale Use

The estimator ignoring scale use is $Cov(x_i, r_{is})$. Because

$$
\begin{aligned}
Cov(x_i, r_{is}) &= Cov(x_i, \alpha_i + (1-\beta_i)\gamma + \beta_i(w_{is} + \epsilon_{is}) + \eta_{is}) \\
&= Cov(x_i, \alpha_i) - \gamma Cov(x_i, \beta_i) + Cov(x_i, \beta_i w_{is}) \\
&= Cov(x_i, \alpha_i) - \gamma Cov(x_i, \beta_i) \\
&\quad + E(\beta_i) Cov(x_i, w_{is}) + E(w_{is}) Cov(x_i, \beta_i) \\
&\quad + E\big[\big(\beta_i - E(\beta_i)\big)\big(w_{is} - E(w_{is})\big)\big(x_i - E(x_i)\big)\big] \\
&= Cov(x_i, \alpha_i) + [E(w_{is}) - \gamma] Cov(x_i, \beta_i) + Cov(x_i, w_{is}) \\
&\quad + E\big[\big(\beta_i - E(\beta_i)\big)\big(w_{is} - E(w_{is})\big)\big(x_i - E(x_i)\big)\big],
\end{aligned}
$$

the bias is

$$
\begin{aligned}
Cov(x_i, r_{is}) - Cov(x_i, w_{is}) &= Cov(x_i, \alpha_i) + [E(w_{is}) - \gamma] Cov(x_i, \beta_i) \\
&\quad + E\big[\big(\beta_i - E(\beta_i)\big)\big(w_{is} - E(w_{is})\big)\big(x_i - E(x_i)\big)\big].
\end{aligned}
$$

The co-skewness term relates to the asymmetry of the joint distribution of (β_i, w_{is}, x_i). It can be nonzero even when x_i is independent of both w_{is} and β_i separately.

C.2.2. When Using the MOM Estimator for $\boldsymbol{Cov(x_i, w_{is})}$

The MOM estimator is $Cov(x_i, r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})$, where $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}$ is the MMB, a weighted average of respondent i's CQ responses (for a derivation, see Appendix E.2). The bias for this estimator is

$$\begin{aligned} Cov(x_i, r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}) - Cov(x_i, w_{is}) \quad &= (E(w_{is}) - \boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{w}_{\mathcal{C}})E\left(\beta_i\big(x_i - E(x_i)\big)\right) \\ &+E\big[\big(\beta_i - E(\beta_i)\big)\big(w_{is} - E(w_{is})\big)\big(x_i - E(x_i)\big)\big]. \end{aligned}$$

The first bias term vanishes if β_i is independent of x_i or if the MMB matches $E(w_{is})$, i.e., $\boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{w}_{\mathcal{C}} = E(w_{is})$. The second bias term is the same co-skewness term in the last subsection.

The MMB's effect in (partially) correcting for the bias in estimating $Cov(x_i, w_{is})$ is two-fold. First, because it is some weighted average of respondent i's CQ responses, it cancels the bias induced by $Cov(x_i, \alpha_i)$. Second, when β_i is correlated with x_i, the MMB cancels the first bias term above because it matches $E(w_{is})$.

Note that the bias does not depend on whether or not β_i and w_{is} are independent. All the corrections are done by the MMB. The MOM estimator, however, does not deal with the co-skewness bias—a bias that only the semi-parametric estimator corrects for.

C.3. **Bias in Estimating $\boldsymbol{Cov(w_{is}, w_{is'})}$**

Utilizing the fact that $Var(w_{is})$ is a special case of $Cov(w_{is}, w_{is'})$, where $s = s'$, we only show the bias in estimating $Cov(w_{is}, w_{is'})$.

C.3.1. When Ignoring Scale Use

The estimator is $Cov(r_{is}, r_{is'})$ when scale use is ignored. Note that

$$
\begin{aligned}
&Cov(r_{is}, r_{is'}) \\
&= Cov(\alpha_i + \gamma + \beta_i(w_{is} - \gamma + \epsilon_{is}) + \eta_{is}, \alpha_i + \gamma + \beta_i(w_{is'} - \gamma + \epsilon_{is'}) + \eta_{is'}) \\
&= \sigma_\alpha^2 + Cov\big(\beta_i(w_{is} - \gamma), \beta_i(w_{is'} - \gamma)\big) + 1(s = s')\big[\sigma_\varepsilon^2(Var(\beta_i) + 1) + E\big(\sigma_{\eta_i}^2\big)\big] \\
&= \sigma_\alpha^2 + 1(s = s')\big[\sigma_\varepsilon^2(Var(\beta_i) + 1) + E\big(\sigma_{\eta_i}^2\big)\big] \\
&+E(\beta_i^2)Cov(w_{is}, w_{is'}) + Var(\beta_i)[E(w_{is}) - \gamma][E(w_{is'}) - \gamma] \\
&+E\left[\left(\beta_i^2 - E(\beta_i^2)\right)\left(w_{is} - E(w_{is})\right)\left(w_{is'} - E(w_{is'})\right)\right] - Cov(\beta_i, w_{is})Cov(\beta_i, w_{is'}) \\
&+[E(w_{is'}) - \gamma][Cov(\beta_i^2, w_{is}) - Cov(\beta_i, w_{is})] + [E(w_{is}) - \gamma][Cov(\beta_i^2, w_{is'}) - Cov(\beta_i, w_{is'})].
\end{aligned}
$$

The bias is thus

$$
\begin{aligned}
&Cov(r_{is}, r_{is'}) - Cov(w_{is}, w_{is'}) \\
&= \sigma_\alpha^2 + 1(s = s')\big[\sigma_\varepsilon^2(Var(\beta_i) + 1) + E\big(\sigma_{\eta_i}^2\big)\big] \\
&+Var(\beta_i)Cov(w_{is}, w_{is'}) + Var(\beta_i)[E(w_{is}) - \gamma][E(w_{is'}) - \gamma] \\
&+E\left[\left(\beta_i^2 - E(\beta_i^2)\right)\left(w_{is} - E(w_{is})\right)\left(w_{is'} - E(w_{is'})\right)\right] - Cov(\beta_i, w_{is})Cov(\beta_i, w_{is'}) \\
&+[E(w_{is'}) - \gamma][Cov(\beta_i^2, w_{is}) - Cov(\beta_i, w_{is})] + [E(w_{is}) - \gamma][Cov(\beta_i^2, w_{is'}) - Cov(\beta_i, w_{is'})].
\end{aligned}
$$

Note that the bias induced by the stretcher depends on cross-question moments, some of which are centered at the means, such as in the term $Var(\beta_i)Cov(w_{is}, w_{is'})$, and some of which are centered at γ, such as in the term $Var(\beta_i)[E(w_{is}) - \gamma][E(w_{is'}) - \gamma]$.

C.3.2. When Ignoring $\boldsymbol{\beta_i}$-$(\boldsymbol{w_{is}}, \boldsymbol{w_{is'}})$ Dependence in Correcting for Scale Use

Assuming β_i-$(w_{is}, w_{is'})$ independence, the MOM estimator for $Cov(w_{is}, w_{is'})$ is (for a derivation, see Appendix E.3):

$$
\frac{1}{E(\beta_i^2)}Cov\big(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}, r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\big) - [1(s = s') + \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}]\sigma_\epsilon^2 - \frac{E\big(\sigma_{\eta_i}^2\big)}{E(\beta_i^2)}[1(s = s') + \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}].
$$

The bias for this estimator is thus

$$
\begin{aligned}
&\frac{1}{E(\beta_i^2)} Cov\left(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{r}_{i\mathcal{C}}, r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{r}_{i\mathcal{C}}\right) - Cov(w_{is}, w_{is'}) \\
&-[1(s = s') + \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\theta}_{s'\mathcal{C}}]\sigma_\epsilon^2 - \frac{E\left(\sigma_{\eta_i}^2\right)}{E(\beta_i^2)}[1(s = s') + \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\theta}_{s'\mathcal{C}}] \\
&= \frac{1}{E(\beta_i^2)} Cov\left(\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}), \beta_i\left(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\right)\right) - Cov(w_{is}, w_{is'}) \\
&= \frac{1}{E(\beta_i^2)} \{E(\beta_i^2) Cov(w_{is}, w_{is'}) + Var(\beta_i)[E(w_{is}) - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}]\left[E(w_{is'}) - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\right] \\
&+ E\left[\left(\beta_i^2 - E(\beta_i^2)\right)\left(w_{is} - E(w_{is})\right)\left(w_{is'} - E(w_{is'})\right)\right] - Cov(\beta_i, w_{is}) Cov(\beta_i, w_{is'}) \\
&+ \left[E(w_{is'}) - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\right][Cov(\beta_i^2, w_{is}) - Cov(\beta_i, w_{is})] \\
&+ [E(w_{is}) - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}][Cov(\beta_i^2, w_{is'}) - Cov(\beta_i, w_{is'})]\} - Cov(w_{is}, w_{is'}) \\
&= \frac{1}{E(\beta_i^2)} \{Var(\beta_i)[E(w_{is}) - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}]\left[E(w_{is'}) - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\right] \\
&+ E\left[\left(\beta_i^2 - E(\beta_i^2)\right)\left(w_{is} - E(w_{is})\right)\left(w_{is'} - E(w_{is'})\right)\right] - Cov(\beta_i, w_{is}) Cov(\beta_i, w_{is'}) \\
&+ \left[E(w_{is'}) - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\right][Cov(\beta_i^2, w_{is}) - Cov(\beta_i, w_{is})] \\
&+ [E(w_{is}) - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}][Cov(\beta_i^2, w_{is'}) - Cov(\beta_i, w_{is'})]\}.
\end{aligned}
$$

When β_i and $(w_{is}, w_{is'})$ are independent, the bias vanishes if the MMB matches either $E(w_{is})$ or $E(w_{is'})$. When β_i and $(w_{is}, w_{is'})$ are dependent, the estimator is biased (by the second and third terms) even if the MMBs match both $E(w_{is})$ and $E(w_{is'})$. The semi-parametric estimator or the comprehensive MLE estimator is needed to fully correct the bias.

D. Semi-parametric Estimators

This appendix details our semi-parametric estimators: we specify the parametric model for α_i, β_i, ϵ_{ic}, and η_{ic} from Section III.D, and we approximate with finite-order polynomials the unknown relationships between functions of w_{is} relevant to the SWB moments of interest and β_i. Our procedure uses the $\hat{\beta}_{i,OLS}$ estimates, together with the estimated distribution of β_i, to correct for noise in the $\hat{\beta}_{i,OLS}$ estimates. It then uses the noise-corrected $\hat{\beta}_{i,OLS}$'s to estimate the relationship between the function of w_{is} and β_i; this aggregate relationship can be estimated adequately despite the uncertainty in the individual-level $\hat{\beta}_{i,OLS}$'s. Below we summarize the steps of our estimator, followed by full details.

Step 1: Eliminating the shifter

Removing bias due to the shifter is straightforward. While subtracting just a single CQ response, $r_{is} - r_{ic}$, would eliminate the shifter (and center), we instead achieve the same goal by subtracting the mean of CQ responses in order to minimize the variance of the added errors:

$$r_{is} - \bar{r}_{i\mathcal{C}} = \beta_i(w_{is} - \bar{w}_{\mathcal{C}} + \epsilon_{is} - \bar{\epsilon}_{i\mathcal{C}}) + \eta_{is} - \bar{\eta}_{i\mathcal{C}}, \quad \text{(D.1)}$$

where $\bar{r}_{i\mathcal{C}} \equiv \frac{1}{C}\sum_{c=1}^{C} r_{ic}$, $\bar{\epsilon}_{i\mathcal{C}} \equiv \frac{1}{C}\sum_{c=1}^{C} \epsilon_{ic}$, and $\bar{\eta}_{i\mathcal{C}} \equiv \frac{1}{C}\sum_{c=1}^{C} \eta_{ic}$. We refer to $\bar{r}_{i\mathcal{C}}$ as the "benchmark."

The remaining steps of the procedure adjust for bias from the stretcher (see Appendix C for details on the bias).

Step 2: Calculate consistent estimates for $\boldsymbol{E(\beta_i^k)}$.

We plug in our estimate of σ_β from the MLE in Section IV into the normal distribution from Section III.D to obtain an estimated distribution of β_i. We then calculate however many moments of the distribution of β_i we will need for subsequent steps: $\hat{E}(\beta_i)$ (which is equal to 1 by construction), $\hat{E}(\beta_i^2)$, $\hat{E}(\beta_i^3)$, etc.

Step 3: Calculate consistent estimates for $\boldsymbol{E(\beta_i^k|\hat{\beta}_{i,OLS})}$.

We derive the conditional distribution of $\hat{\beta}_{i,OLS}|\beta_i$, which is a linear function of the CQ errors (see Section D.1 below). Therefore, $\hat{\beta}_{i,OLS}|\beta_i$ is normally distributed with mean zero and variance that is a known linear function of σ_ϵ^2 and $\sigma_{\eta_i}^2$. We then numerically apply Bayes' rule using the estimated distribution of β_i to obtain an estimated distribution of $\beta_i|\hat{\beta}_{i,OLS}$. From this, we calculate consistent estimates for $E(\beta_i|\hat{\beta}_{i,OLS})$, $E(\beta_i^2|\hat{\beta}_{i,OLS})$, $E(\beta_i^3|\hat{\beta}_{i,OLS})$, etc.

The steps up until here are computationally intensive, but we highlight that (i) they rely only on the CQ responses, and (ii) they need only be performed once, and then the estimated moments of β_i and $\beta_i|\hat{\beta}_{i,OLS}$ can be stored with the dataset for use with whatever future research questions may come up with the data. In contrast, steps 4-6 (see below) are computationally quick.

Intermezzo: Model the correlation between common-scale SWB and the stretcher.

The subsequent steps of the econometric procedure depend on the SWB moment of interest. Before proceeding with those steps, we formulate a framework within which we can describe the remainder of the econometric procedure in general for all four SWB moments we study, as well as other SWB moments we do not study in this paper.

For each of the four SWB moments of interest, we define a related quantity called ψ_i, given in the table below. It involves $w_{is} - \overline{w}_{\mathcal{C}}$ (defined in step 1), and its mean differs from the SWB moment only by estimable quantities that are not functions of the SWB moment. Steps 4-5 below aim to estimate the mean $E(\psi_i)$, while step 6 adjusts the resulting estimate to obtain an estimate of the SWB moment. The reason we focus on ψ_i is that we can learn about it using its reported-SWB analog, which we call Ψ_i, also defined in the table below. Since Ψ_i involves $r_{is} - \bar{r}_{i\mathcal{C}}$, it does not depend on the shifter (or the center), and the only scale-use heterogeneity we need to deal with is the stretcher.

Table D.1. SWB Moments and Corresponding $\boldsymbol{\psi_i}$, $\boldsymbol{\Psi_i}$, and $\boldsymbol{E\left(\Psi_i|\beta_i, \hat{\beta}_{i,OLS}\right)}$

SWB Moment	ψ_i	Ψ_i	$E\left(\Psi_i\|\beta_i, \hat{\beta}_{i,OLS}\right)$
$E(w_{is})$	$w_{is} - \overline{w}_{\mathcal{C}}$	$r_{is} - \bar{r}_{i\mathcal{C}}$	$\sum_{k=0}^{K} A_{ks}\beta_i^{k+1}$
$Cov(x_i, w_{is})$	$(w_{is} - \overline{w}_{\mathcal{C}})x_i$	$(r_{is} - \bar{r}_{i\mathcal{C}})x_{is}$	$\sum_{k=0}^{K} A_{ks}\beta_i^{k+1}$
$Cov(w_{is}, w_{is'})$	$(w_{is} - \overline{w}_{\mathcal{C}})(w_{is'} - \overline{w}_{\mathcal{C}})$	$(r_{is} - \bar{r}_{i\mathcal{C}})(r_{is'} - \bar{r}_{i\mathcal{C}})$	$\sum_{k=0}^{K} A_{ks}\beta_i^{k+2} + V_1$
$Var(w_{is})$	$(w_{is} - \overline{w}_{\mathcal{C}})^2$	$(r_{is} - \bar{r}_{i\mathcal{C}})^2$	$\sum_{k=0}^{K} A_{ks}\beta_i^{k+2} + V_1 + V_2$

To model the relationship between ψ_i and β_i, we approximate $E(\psi_i|\beta_i)$ as a polynomial in β_i:

(D.2) $$E\left(\psi_i|\beta_i\right) = A_{0s} + A_{1s}\beta_i + A_{2s}\beta_i^2 + \cdots + A_{Ks}\beta_i^K = \sum_{k=0}^{K} A_{ks}\beta_i^k,$$

where $A_{0s}, A_{1s}, A_{2s}, \ldots, A_{Ks}$ are unknown parameters.

In what follows, we will use the law of iterated expectations to recover $E(\psi_i)$ as a function of the A_{ks}'s and the moments of β_i: $E(\psi_i) = E[E(\psi_i|\beta_i)]$. We estimated the moments of β_i in step 2, and we will estimate the A_{ks}'s in step 4. In step 5, plugging these in will give us an estimate of $E(\psi_i)$.

We cannot use equation (D.2) to estimate the A_{ks}'s because both w_{is} and β_i are unobserved, but we can instead work with $E\left(\Psi_i|\hat{\beta}_{i,OLS}\right)$, which depends only on observables. To derive that conditional expectation, we begin by considering $E\left(\Psi_i|\beta_i, \hat{\beta}_{i,OLS}\right)$, and we claim that conditioning on β_i, $\hat{\beta}_{i,OLS}$ adds no information about $\bar{r}_{i\mathcal{C}}$:

$$E\left(\Psi_i|\beta_i, \hat{\beta}_{i,OLS}\right) = E(\Psi_i|\beta_i). \tag{D.3}$$

There are two approaches to justify equation (D.3). The first approach, while relatively straightforward, necessitates dividing the CQs into two sets, and separately calculating the benchmark $\bar{r}_{i\mathcal{C}}$ from step 1 and the $\hat{\beta}_{i,OLS}$'s using these distinct sets. The second approach involves more derivations (elaborated in Section D.4 below), but provides the advantage of using all CQs for calculating both $\bar{r}_{i\mathcal{C}}$ and the $\hat{\beta}_{i,OLS}$'s. We adopt the second approach because it offers more efficient utilization of our data.

Equation (D.3) is useful because its right-hand side can be calculated using equation (D.2). For the first two moments of SWB in the table, it is equal to $E(\psi_i|\beta_i)$ multiplied by β_i. For example, when the SWB moment is the mean,

$$E(\Psi_i|\beta_i) = E(r_{is} - \bar{r}_{i\mathcal{C}}|\beta_i) = E[\beta_i(w_{is} - \bar{w}_{\mathcal{C}})|\beta_i] = \beta_i E[(w_{is} - \bar{w}_{\mathcal{C}})|\beta_i] = \beta_i E(\psi_i|\beta_i).$$

For the other two SWB moments in the table, $E(\Psi_i|\beta_i)$ is equal to $E(\psi_i|\beta_i)$ multiplied by β_i^2, plus additional term(s). The table above shows the resulting expression for $E(\Psi_i|\beta_i) = E\left(\Psi_i|\beta_i, \hat{\beta}_{i,OLS}\right)$ for each of the four SWB moments of interest after substituting for $E(\psi_i|\beta_i)$ using equation (D.2). The additional terms for the last two moments of SWB are related to variances:

$$V_1 = \frac{1}{C}\left[\sigma_\epsilon^2 E(\beta_i^2) + E\left(\sigma_{\eta_i}^2\right)\right]$$

$$V_2 = \sigma_\epsilon^2 E(\beta_i^2) + E\left(\sigma_{\eta_i}^2\right).$$

To obtain $E(\Psi_i|\hat{\beta}_{i,OLS})$, we apply the law of iterated expectations: $E(\Psi_i|\hat{\beta}_{i,OLS}) = E[E(\Psi_i|\beta_i, \hat{\beta}_{i,OLS})|\hat{\beta}_{i,OLS}]$. For example, when the SWB moment of interest is the mean,

(D.4) $$E(\Psi_i|\hat{\beta}_{i,OLS}) = E\left(\sum_{k=0}^{K} A_{ks}\beta_i^{k+1}\,|\hat{\beta}_{i,OLS}\right) = \sum_{k=0}^{K} A_{ks}E(\beta_i^{k+1}|\hat{\beta}_{i,OLS}).$$

In general, for any SWB moment of interest, $E(\Psi_i|\hat{\beta}_{i,OLS})$ will equal the expression in the table above for $E(\Psi_i|\beta_i, \hat{\beta}_{i,OLS})$, but with each β_i term replaced by its expectation conditional on $\hat{\beta}_{i,OLS}$.

Step 4: Estimate A_{ks}.

For the first two moments of SWB in the table above, based on equation (D.4) we can consistently estimate the A_{ks}'s by the OLS regression of Ψ_i on $\hat{E}(\beta_i|\hat{\beta}_{i,OLS})$, $\hat{E}(\beta_i^2|\hat{\beta}_{i,OLS})$, etc., from step 3, without a constant term.[10] For example, when the SWB moment of interest is the mean, we run a regression of $r_{is} - \bar{r}_{i\mathcal{C}}$ on the estimated moments of $\beta_i|\hat{\beta}_{i,OLS}$ without a constant term.

For the covariance between two SWB questions, we need to construct an estimate of $E(V_1|\hat{\beta}_{i,OLS})$, which we do by plugging in our estimates of σ_ϵ^2, $E\left(\sigma_{\eta_i}^2\right)$, and $E(\beta_i^2|\hat{\beta}_{i,OLS})$. We then run the same regression as above but with the dependent variable being Ψ_i minus the estimate of $E(V_1|\hat{\beta}_{i,OLS})$. Note that σ_ϵ^2 and $E\left(\sigma_{\eta_i}^2\right)$ are estimated from the data on CQs, rather than the SWB questions. That is because the response-error variances for the SWB questions are not identified, as discussed in the paper in the context of the comprehensive MLE estimator for the variance of SWB and the covariance between two SWB questions.

[10] While both OLS and weighted least squares (WLS) generate consistent estimates, one might expect WLS to be more efficient because, from equation (D.1), the error in $r_{is} - \bar{r}_{i\mathcal{C}}$ is $\eta_{is} - \bar{\eta}_{i\mathcal{C}} + \beta_i(\epsilon_{is} - \bar{\epsilon}_{i\mathcal{C}})$, which depends on β_i and is therefore heteroskedastic. However, in simulations, we find that OLS gives more precise estimates, potentially because the weights in WLS are not robustly estimated when β_i is small.

For the variance of SWB, we proceed similarly, but we need to also subtract an estimate of $E\left(V_2|\hat{\beta}_{i,OLS}\right)$ before running the regression.

Step 5: Plug in $\hat{\boldsymbol{A}}_{\boldsymbol{ks}}$ and $\hat{\boldsymbol{E}}\left(\boldsymbol{\beta}_{\boldsymbol{i}}^{\boldsymbol{k}}\right)$.

Applying the law of iterated expectations to $E(\psi_i|\beta_i)$ and using equation (D.2) yields the SWB moment of interest, $E(\psi_i) = E[E(\psi_i|\beta_i)]$, as a function of the A_{ks}'s and moments of β_i. For example, for mean SWB,

$$E(\psi_i) = E(w_{is} - \bar{w}_{\mathcal{C}}) = E[E(w_{is} - \bar{w}_{\mathcal{C}}|\beta_i)] = \sum_{k=0}^{K} A_{ks} E\left(\beta_i^k\right).$$

Substituting the estimated A_{ks}'s from step 4 and estimated moments of β_i from step 2 into the above expression gives a consistent estimate of $E(\psi_i)$.

Step 6: Estimate the SWB moment by adjusting the estimate of $\boldsymbol{E}(\boldsymbol{\psi}_{\boldsymbol{i}})$.

Each of the four SWB moments can be decomposed into $E(\psi_i)$ and a function of estimable quantities:

$$E(w_{is}) = E(w_{is} - \bar{w}_{\mathcal{C}}) + \bar{w}_{\mathcal{C}}$$
$$Cov(x_i, w_{is}) = E[(w_{is} - \bar{w}_{\mathcal{C}})x_i] + [\bar{w}_{\mathcal{C}} - E(w_{is})]E(x_i)$$
$$Cov(w_{is}, w_{is'}) = E[(w_{is} - \bar{w}_{\mathcal{C}})(w_{is'} - \bar{w}_{\mathcal{C}})] - [E(w_{is}) - \bar{w}_{\mathcal{C}}][E(w_{is'}) - \bar{w}_{\mathcal{C}}]$$
$$Var(w_{is}) = E[(w_{is} - \bar{w}_{\mathcal{C}})^2] - [E(w_{is}) - \bar{w}_{\mathcal{C}}]^2.$$

We estimate the SWB moments by substituting our estimates of $E(\psi_i)$ for the first right-hand-side term in each equation above, $\bar{r}_{i\mathcal{C}}$ for $\bar{w}_{\mathcal{C}}$, the sample mean of x_i for $E(x_i)$, and the estimated SWB mean from the first equation above for $E(w_{is})$ in the other equations.

D.1. Estimating $\boldsymbol{E}\left(\boldsymbol{\beta}_{\boldsymbol{i}}^{\boldsymbol{k}}|\hat{\boldsymbol{\beta}}_{\boldsymbol{i,OLS}}\right)$

To estimate $E\left(\beta_i^k|\hat{\beta}_{i,OLS}\right)$, as described by step 3 above, we need the density of $\beta_i|\hat{\beta}_{i,OLS}$, which, by Bayes' rule, is

$$f(\beta_i|\hat{\beta}_{i,OLS}) = \frac{f(\hat{\beta}_{i,OLS}|\beta_i)f(\beta_i)}{\int f(\beta_i,\hat{\beta}_{i,OLS})d\beta_i}.$$

We estimate $f(\beta_i)$ as part of our CQ-only MLE. To find $f(\hat{\beta}_{i,OLS}|\beta_i)$, we note that $\hat{\beta}_{i,OLS}$ estimates β_i from the individual-level OLS-regression of r_{ic} on $\bar{r}_c$:

$$r_{ic} = a_i + \beta_i \bar{r}_c + u_{ic},$$

where a_i is the coefficient for the constant term, $\bar{r}_c = \frac{1}{I}\sum_{i=1}^{I} r_{ic}$, and u_{ic} is the regression error. Thus, the exact expression for $\hat{\beta}_{i,OLS}$ is

$$\begin{aligned}
\hat{\beta}_{i,OLS} \quad &= \frac{\sum_{c=1}^{C}\left(r_{ic} - \frac{1}{C}\sum_{c=1}^{C} r_{ic}\right)\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right)}{\sum_{c=1}^{C}\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right)^2} \\
&= \frac{\frac{1}{C}\sum_{c=1}^{C} r_{ic}\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right)}{\frac{1}{C}\sum_{c=1}^{C}\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right)^2} \\
&\equiv \mathbf{z}'\boldsymbol{r}_{i\mathcal{C}},
\end{aligned}$$

where $\mathbf{z} \equiv \mathbf{m}_{\mathcal{C}}/v$, $\mathbf{m}_{\mathcal{C}} \equiv \{m_c\}_{c=1}^{C} \equiv \left\{\frac{1}{C}\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right)\right\}_{c=1}^{C}$, and $v \equiv \frac{1}{C}\sum_{c=1}^{C}\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right)^2$.

To simplify the expression for $\hat{\beta}_{i,OLS}$, note that $\mathbf{z}'\mathbf{1}_C = 0$, where $\mathbf{1}_C$ is the C-dimensional vector full of ones, and $\mathbf{z}'\boldsymbol{w}_{\mathcal{C}} \approx 1$. The second expression holds because

$$\begin{aligned}
\mathbf{m}'_{\mathcal{C}}\boldsymbol{w}_{\mathcal{C}} - v \quad &= \sum_{c=1}^{C} w_c \frac{1}{C}\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right) - \frac{1}{C}\sum_{c=1}^{C}\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right)^2 \\
&= \frac{1}{C}\sum_{c=1}^{C}\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right)\left(w_c - \bar{r}_c + \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right) \\
&= \frac{1}{C}\sum_{c=1}^{C}\left(\bar{r}_c - \frac{1}{C}\sum_{c=1}^{C}\bar{r}_c\right)(w_c - \bar{r}_c) \\
&\approx 0,
\end{aligned}$$

where the approximation holds because $\bar{r}_c = w_c + \mathcal{O}\left(I^{-1/2}\right)$ and becomes exact as $I \to \infty$.

Plugging $\mathbf{z}'\mathbf{1}_C = 0$ and $\mathbf{z}'\boldsymbol{w}_{\mathcal{C}} \approx 1$ into the expression of $\hat{\beta}_{i,OLS}$, we get

$$\begin{aligned}\hat{\beta}_{i,OLS} &= \mathbf{z}'\boldsymbol{r}_{i\mathcal{C}} \\ &= \mathbf{z}'\left[\left(\alpha_i + \beta_i(1-\gamma)\right)\mathbf{1}_C + \beta_i(\boldsymbol{w}_{\mathcal{C}} + \boldsymbol{\epsilon}_{i\mathcal{C}}) + \boldsymbol{\eta}_{i\mathcal{C}}\right] \\ &\approx \beta_i + \beta_i\mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}.\end{aligned}$$

This shows that $\hat{\beta}_{i,OLS}$ is (approximately) the individual's true stretcher β_i plus a linear combination of the individual's C-dimensional vectors of CQ response errors $(\boldsymbol{\epsilon}_{i\mathcal{C}}, \boldsymbol{\eta}_{i\mathcal{C}})$. The approximation should be quite accurate when the number of respondents is not small.

Since ϵ_{ic} and η_{ic} are independently distributed normals, $\hat{\beta}_{i,OLS}|\beta_i$ is (approximately) normally distributed with density

$$f\left(\hat{\beta}_{i,OLS}|\beta_i\right) = \sqrt{\frac{Cv}{2\pi\left(\beta_i^2\sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)}}\, e^{-\frac{Cv\left(\hat{\beta}_{i,OLS}-\beta_i\right)^2}{2\;\beta_i^2\sigma_\epsilon^2+\sigma_{\eta_i}^2}}.$$

To avoid the numerical integration in the denominator of $f\left(\beta_i|\hat{\beta}_{i,OLS}\right)$ and the numerical integration for calculating the expectation $E\left(\beta_i^k|\hat{\beta}_{i,OLS}\right)$, we use Hamiltonian Monte Carlo with hierarchical modeling to draw samples for $\beta_i|\hat{\beta}_{i,OLS}$ and then raise the sample to the k-th power to estimate $E\left(\beta_i^k|\hat{\beta}_{i,OLS}\right)$. These moments vary across individuals because $\hat{\beta}_{i,OLS}$ varies across individuals.

D.2. Optimal Benchmarking

We use the term *benchmarking* to mean subtracting a linear combination of CQ ratings from each respondent's SWB rating. Step 1 above is benchmarking with uniform weights. Here we show that this choice of the weights is optimal.

We perform benchmarking first and foremost to eliminate the shifter α_i in the SWB rating r_{is} and the (stretcher-scaled) center $(1-\beta_i)\gamma$. Subtracting even a single CQ achieves the goal:

$$r_{is} - r_{ic} = \beta_i(w_{is} - w_c + \epsilon_{is} - \epsilon_{ic}) + \eta_{is} - \eta_{ic}.$$

The differencing introduces additional noise because the differenced quantity, $r_{is} - r_{ic}$, includes response errors from both SWB and CQ ratings, whereas r_{is} includes only SWB ratings' errors. To minimize the noise, we subtract a weighted average of CQ ratings—i.e., a benchmark, as defined above—instead of a single CQ rating. The weights in the benchmark have to sum to unity so that the cancellation of the shifter and the center is retained. The mean-matched benchmark discussed in the paper is a special type of benchmark that also sets the mean of the benchmark to some target value for removing certain biases (see Appendix E.2.1 for details). The weights are then chosen to minimize the noise introduced from subtracting the benchmark.

To achieve this objective, we choose $\boldsymbol{\theta}_{s\mathcal{C}}$ to minimize the noise variance of the resulting benchmarked SWB rating, $Var(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})$, where $\boldsymbol{\theta}_{s\mathcal{C}}$ denotes a C-dimensional vector of weights for benchmarking SWB rating s using CQs in set $\mathcal{C}$, subject to the constraint: $\boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{1}_C = 1$.

Since the variance of the benchmarked SWB rating is

$$\begin{aligned}
&Var(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}) \\
&= Var[\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}) + \beta_i(\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}) + \eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}] \\
&= Var[\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})] + Var[\beta_i(\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}})] + Var(\eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}) \\
&= Var[\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})] + Var(\beta_i\epsilon_{is}) + Var(\eta_{is}) + \left[\sigma_\epsilon^2 E(\beta_i^2) + E\left(\sigma_{\eta_i}^2\right)\right]\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s\mathcal{C}},
\end{aligned}$$

only the last term is relevant for the minimization problem.

Minimizing the last term subject to the sum-to-unity constraint gives the optimal benchmarking weights:

$$\boldsymbol{\theta}^*_{s\mathcal{C}} = \frac{\mathbf{1}_C}{C},$$

which shows that the benchmarking in step 1 above is optimal.

D.3. Derivation of Semi-parametric Estimators

This subsection details the derivation of the semi-parametric estimators for the common-scale SWB's moments listed in Table D.1.

D.3.1. Semi-parametric Estimator for $E(\boldsymbol{w}_{is})$

Since

$$
\begin{aligned}
E\left(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}} | \hat{\beta}_i\right) &= E\left[E\left(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}} | \beta_i, \hat{\beta}_{i,OLS}\right) | \hat{\beta}_{i,OLS}\right] \\
&= E\left[E(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}} | \beta_i) | \hat{\beta}_{i,OLS}\right] \\
&= E\left[\beta_i E(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}} | \beta_i) | \hat{\beta}_{i,OLS}\right] \\
&\approx \sum_{k=0}^{K_{Ew}} A_{ks}\, E\left(\beta_i^{k+1} | \hat{\beta}_{i,OLS}\right),
\end{aligned}
$$

where the 2nd equality follows from equation (D.3) and the last approximates $E(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}} | \beta_i)$ with $\sum_{k=0}^{K_{Ew}} A_{ks}\, \beta_i^k$, the A_{ks}'s can be estimated by regressing (without a constant term) the benchmarked rating $(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})$ on $E\left(\beta_i^{k+1} | \hat{\beta}_{i,OLS}\right)$ for $k = 0, \cdots, K_{Ew}$. In practice, we choose $K_{Ew} = 1$ and provide sufficient conditions for the satisfaction of the equation (D.3) in Section D.4 of this appendix.

With the estimated A_{ks}'s, we can estimate $E(w_{is})$ by

$$
\hat{E}(w_{is}) = \sum_{k=0}^{K_{Ew}} \hat{A}_{ks}\, \hat{E}\left(\beta_i^k\right) + \boldsymbol{\theta}'_{s\mathcal{C}}\hat{\boldsymbol{w}}_{\mathcal{C}},
$$

where $E\left(\beta_i^k\right)$'s can be estimated based on the output of the CQ-only MLE and $\hat{\boldsymbol{w}}_{\mathcal{C}} = \frac{1}{I}\sum_{i=1}^{I} \boldsymbol{r}_{i\mathcal{C}}$.

D.3.2. Semi-parametric Estimator for $Cov(\boldsymbol{x}_i, \boldsymbol{w}_{is})$

Note that

$$
\begin{aligned}
E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})x_i | \hat{\beta}_{i,OLS}\right) &= E\left[E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})x_i | \beta_i, \hat{\beta}_{i,OLS}\right) | \hat{\beta}_{i,OLS}\right] \\
&= E\left[E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})x_i | \beta_i\right) | \hat{\beta}_{i,OLS}\right] \\
&= E\left[\beta_i E\left((w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})x_i | \beta_i\right) | \hat{\beta}_{i,OLS}\right] \\
&\approx \sum_{k=0}^{K_{Covwx}} A_{ksx}\, E\left(\beta_i^{k+1} | \hat{\beta}_{i,OLS}\right),
\end{aligned}
$$

where the 2nd equality uses a modified version of equation (D.3) and the last approximates $E\big((w_{is}-\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})x_i|\beta_i\big)$ with $\sum_{k=0}^{K_{Covwx}} A_{ksx}\,\beta_i^k$. This suggests that we can estimate the A_{ksx}'s by regressing (without a constant term) $(r_{is}-\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})x_i$ on $E\big(\beta_i^{k+1}|\hat{\beta}_{i,OLS}\big)$ for $k=0,\dots,K_{Covwx}$. We set K_{Covwx} to 1 in practice and provide sufficient conditions for the satisfaction of the version of equation (D.3) in Section D.4 below.

With the estimated A_{ksx}'s we can estimate $Cov(x_i,w_{is})$ by

$$\widehat{Cov}(w_{is},x_i)=\sum_{k=0}^{K_{Covwx}}\hat{A}_{ksx}\,\hat{E}\big(\beta_i^k\big)+\big[\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}-\hat{E}(w_{is})\big]\hat{E}(x_i),$$

where $\hat{E}(w_{is})$ is estimated using the estimator in the last subsection and $\hat{E}(x_i)=\frac{1}{I}\sum_{i=1}^{I}x_i$.

D.3.3. Semi-parametric Estimator for $Cov(w_{is},w_{is'})$

As $Var(w_{is})=Cov(w_{is},w_{is})$, we only derive the semi-parametric estimator for $Cov(w_{is},w_{is'})$ while allowing for $s'=s$ below.

Because

$$\begin{aligned}
&E\left((r_{is}-\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})\big(r_{is'}-\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\big)|\hat{\beta}_{i,OLS}\right)\\
&=E\left[E\left((r_{is}-\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})\big(r_{is'}-\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\big)|\beta_i,\hat{\beta}_i\right)|\hat{\beta}_{i,OLS}\right]\\
&=E\left[E\left((r_{is}-\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})\big(r_{is'}-\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\big)|\beta_i\right)|\hat{\beta}_{i,OLS}\right]\\
&=E\left(\beta_i^2E\left((w_{is}-\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})\big(w_{is'}-\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}\big)|\beta_i\right)|\hat{\beta}_{i,OLS}\right)\\
&+[1(s=s')+\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}]\sigma_\epsilon^2E\big(\beta_i^2|\hat{\beta}_{i,OLS}\big)+[1(s=s')+\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}]E\big(\sigma_{\eta_i}^2\big)\\
&\approx\sum_{k=0}^{K_{Covww}}A_{kss'}\,E\big(\beta_i^{k+2}|\hat{\beta}_{i,OLS}\big)\\
&+1(s=s')\big[\sigma_\epsilon^2E\big(\beta_i^2|\hat{\beta}_{i,OLS}\big)+E\big(\sigma_{\eta_i}^2\big)\big]+\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}\big[\sigma_\epsilon^2E\big(\beta_i^2|\hat{\beta}_{i,OLS}\big)+E\big(\sigma_{\eta_i}^2\big)\big],
\end{aligned}$$

where the 2nd equality uses a modified version of equation (D.3) and the last approximates $E\left((w_{is}-\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})\big(w_{is'}-\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}\big)|\beta_i\right)$ with $\sum_{k=0}^{K_{Covww}}A_{kss'}\,\beta_i^k$, the $A_{kss'}$'s can be estimated by regressing (without a constant term)

$$(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})\left(r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\right) - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}\left[\sigma_\epsilon^2 E\left(\beta_i^2|\hat{\beta}_{i,OLS}\right) + E\left(\sigma_{\eta_i}^2\right)\right]$$
$$- 1(s = s')\left[\sigma_\epsilon^2 E\left(\beta_i^2|\hat{\beta}_{i,OLS}\right) + E\left(\sigma_{\eta_i}^2\right)\right]$$

on $E\left(\beta_i^{k+2}|\hat{\beta}_{i,OLS}\right)$ for $k = 0, \dots, K_{Covww}$. We set $K_{Covww} = 2$ in practice and provide sufficient conditions for the satisfaction of the version of equation (D.3) in Section D.4 of this appendix.

With the estimated $A_{kss'}$'s we can estimate $Cov(w_{is}, w_{is'})$ by

$$\widehat{Cov}(w_{is}, w_{is'}) = \sum_{k=0}^{K_{Covww}} \hat{A}_{kss'}\,\hat{E}\left(\beta_i^k\right) - \left[\hat{E}(w_{is}) - \boldsymbol{\theta}'_{s\mathcal{C}}\widehat{\boldsymbol{w}}_{\mathcal{C}}\right]\left[\hat{E}(w_{is'}) - \boldsymbol{\theta}'_{s'\mathcal{C}}\widehat{\boldsymbol{w}}_{\mathcal{C}}\right].$$

D.4. Sufficient Conditions for Equation (D.3)

To satisfy equation (D.3), as discussed above, one can use disjoint sets of CQs for calculating $\hat{\beta}_{i,OLS}$ and the benchmark. Here we provide sufficient conditions for various versions of equation (D.3) where benchmarking and calculation of $\hat{\beta}_{i,OLS}$ use the same set of CQs.

The discussion shows that, among other things, uniform benchmarking used by step 1 satisfies these conditions. To briefly see why, recall from Section D.1 above, $\left(\hat{\beta}_{i,OLS} - \beta_i\right)|\beta_i$ is (i) a weighted average of response errors where the weights (**z**) sum to zero, and (ii) normally distributed. Moreover, $Cov\left(\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}, \hat{\beta}_{i,OLS} - \beta_i|\beta_i\right)$ is proportional to the dot product of $\boldsymbol{\theta}_{s\mathcal{C}}$ and **z**. The covariance thus equals zero when we use equal weights: $\boldsymbol{\theta}'_{s\mathcal{C}} = \frac{1}{C}\mathbf{1}_{\boldsymbol{C}}$. And because, conditional on β_i, $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}$ and $\hat{\beta}_{i,OLS} - \beta_i$ are jointly normally distributed, zero covariance implies independence. The conditional independence between these two quantities further implies equation (D.3): the conditional mean-independence of $\hat{\beta}_{i,OLS}$ and certain functions of $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}$.

D.4.1. For Semi-parametric Estimator of $\boldsymbol{E(w_{is})}$

The semi-parametric estimator for $E(w_{is})$ assumes equation (D.3), the conditional (on β_i) mean independence of $\hat{\beta}_{i,OLS}$ and $(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})$. A sufficient condition for this assumption is that $\hat{\beta}_{i,OLS}$ is conditionally independent of $(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})$. As shown in Section D.1 of this appendix, $\left(\hat{\beta}_{i,OLS} - \beta_i\right)$ is a linear combination of CQ response errors. The conditional independence is therefore equivalent to the conditional independence between $\left(\hat{\beta}_{i,OLS} - \beta_i\right)$ and $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}$ (since the SWB rating r_{is} is independent of the CQ response errors). To achieve the conditional independence, we need

$$
\begin{aligned}
& Cov\left(\hat{\beta}_{i,OLS} - \beta_i, \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}} | \beta_i\right) \\
= \;& Cov(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s\mathcal{C}}(\alpha_i \mathbf{1}_C + (1-\beta_i)\gamma\mathbf{1}_C + \beta_i(\boldsymbol{w}_{\mathcal{C}} + \boldsymbol{\epsilon}_{i\mathcal{C}}) + \boldsymbol{\eta}_{i\mathcal{C}}) | \beta_i) \\
= \;& Cov(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s\mathcal{C}}(\beta_i\boldsymbol{\epsilon}_{i\mathcal{C}} + \boldsymbol{\eta}_{i\mathcal{C}}) | \beta_i) \\
= \;& \beta_i^2 Cov(\mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}} | \beta_i) + Cov(\mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}} | \beta_i) \\
= \;& \beta_i^2 \sigma_\epsilon^2 \boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{z} + E\left(\sigma_{\eta_i}^2\right)\boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{z} \\
= \;& 0.
\end{aligned}
$$

Because $\sigma_\epsilon^2 \boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{z}$ and $E\left(\sigma_{\eta_i}^2\right)\boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{z}$ do not vary with i, the only way for the last equality to hold for all i is to have

$$\boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{z} = 0,$$

which is true as $\boldsymbol{\theta}_{s\mathcal{C}}$ contains uniform weights (and $\mathbf{1}'\mathbf{z} = 0$).

D.4.2. For Semi-parametric Estimator of $Cov(x_i, w_{is})$

The semi-parametric estimator for $Cov(x_i, w_{is})$ assumes a version of equation (D.3): the conditional (on β_i) mean independence of $\hat{\beta}_{i,OLS}$ and $(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})x_i$:

$$E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})x_i | \beta_i, \hat{\beta}_{i,OLS}\right) = E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})x_i | \beta_i\right).$$

Since

$$
\begin{aligned}
& E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})x_i | \beta_i, \hat{\beta}_{i,OLS}\right) \\
&= E\left([\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}) + \beta_i(\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}) + \eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}]x_i | \beta_i, \hat{\beta}_{i,OLS}\right) \\
&= \beta_i E\left((w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})x_i | \beta_i, \hat{\beta}_{i,OLS}\right) + \beta_i E\left((\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}})x_i | \beta_i, \hat{\beta}_{i,OLS}\right) + E\left((\eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}})x_i | \beta_i, \hat{\beta}_{i,OLS}\right),
\end{aligned}
$$

and

$$\begin{aligned}
&E\big((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})x_i|\beta_i\big) \\
&= E([\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}) + \beta_i(\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}) + \eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}]x_i|\beta_i) \\
&= \beta_i E\big((w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})x_i|\beta_i\big) + \beta_i E\big((\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}})x_i|\beta_i\big) + E\big((\eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}})x_i|\beta_i\big),
\end{aligned}$$

to satisfy the version of equation (D.3), one set of sufficient conditions is the following:

(1) $E\left((\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}})x_i|\beta_i, \hat{\beta}_{i,OLS}\right) = E\big((\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}})x_i|\beta_i\big)$,

(2) $E\left((\eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}})x_i|\beta_i, \hat{\beta}_{i,OLS}\right) = E\big((\eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}})x_i|\beta_i\big)$, and

(3) $E\left((w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})x_i|\beta_i, \hat{\beta}_{i,OLS}\right) = E\big((w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})x_i|\beta_i\big)$.

To see that conditions (1) and (2) are satisfied, note that the independence between the mean-zero response errors and (x_i, β_i) implies that the right-hand sides of both equal zero. Additionally, the conditional (on β_i) independence between $\hat{\beta}_{i,OLS}$ and $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}$ implies that the left-hand side of (1) equals zero, while the conditional independence between $\hat{\beta}_{i,OLS}$ and $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}$ implies that the left-hand side of (2) equals zero.

To satisfy condition (3), we note that it is equivalent to

$$E\big(w_{is}x_i|\beta_i, \hat{\beta}_{i,OLS}\big) = E(w_{is}x_i|\beta_i),$$

which holds because the response errors are independent from (w_{is}, x_i).

D.4.3. For Semi-parametric Estimator of $Cov(\boldsymbol{w_{is}}, \boldsymbol{w_{is'}})$

The semi-parametric estimator for $Cov(w_{is}, w_{is'})$ assumes a version of equation (D.3): the conditional (on β_i) mean independence of $\hat{\beta}_{i,OLS}$ and $(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})(r_{is'} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})$:

$$E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})\big(r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\big)|\beta_i, \hat{\beta}_{i,OLS}\right) = E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})\big(r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\big)|\beta_i\right).$$

Because

$$
\begin{aligned}
& (r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{r}_{i\mathcal{C}})\big(r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{r}_{i\mathcal{C}}\big) \\
= \quad & [\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}) + \beta_i(\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}) + (\eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}})] \\
& \cdot \big[\beta_i\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\big) + \beta_i\big(\epsilon_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}\big) + \big(\eta_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}\big)\big] \\
= \quad & \beta_i^2(w_{i\mathcal{C}} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}})\big(w_{i\mathcal{C}'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\big) + \beta_i^2(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}})\big(\epsilon_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}\big) \\
& + \beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}})\big(\eta_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}\big) + \beta_i^2\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\big)(\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}) \\
& + \beta_i^2(\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}})\big(\epsilon_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}\big) + \beta_i(\epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}})\big(\eta_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}\big) \\
& + \beta_i\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\big)(\eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}) + \beta_i\big(\epsilon_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}\big)(\eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}) \\
& + (\eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}})\big(\eta_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}\big),
\end{aligned}
$$

we have

$$
\begin{aligned}
& E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{r}_{i\mathcal{C}})\big(r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{r}_{i\mathcal{C}}\big)|\beta_i, \hat{\beta}_{i,OLS}\right) \\
= \quad & \beta_i^2 E\left((w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}})\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\big)|\beta_i\right) + \beta_i^2 E(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}|\beta_i) E\big(-\boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) \\
& + \beta_i E(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}|\beta_i) E\big(-\boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) + \beta_i^2 E\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}|\beta_i\big) E\big(-\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) \\
& + \beta_i^2\big[1(s = s')\sigma_\epsilon^2 + E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big)\big] + \beta_i E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) \\
& + \beta_i E\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}|\beta_i\big) E\big(-\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) + \beta_i E\big(\boldsymbol{\theta}'_{\boldsymbol{s'}\boldsymbol{\mathcal{C}}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) \\
& + 1(s = s') E\big(\sigma_{\eta_i}^2\big) + E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{\boldsymbol{i}\boldsymbol{\mathcal{C}}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) \\
= \quad & \beta_i^2 E\left((w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}})\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\big)|\beta_i\right) + \beta_i E(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}|\beta_i) E\big(-\beta_i \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) \\
& + \beta_i E\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}|\beta_i\big) E\big(-\beta_i \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) \\
& + \beta_i^2\big[1(s = s')\sigma_\epsilon^2 + E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big)\big] + \beta_i E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) \\
& + \beta_i E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big) + 1(s = s') E\big(\sigma_{\eta_i}^2\big) + E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{\boldsymbol{i}\boldsymbol{\mathcal{C}}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}\big),
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
& E\left((r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{r}_{i\mathcal{C}})\big(r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{r}_{i\mathcal{C}}\big)|\beta_i\right) \\
= \quad & \beta_i^2 E\left((w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}})\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\big)|\beta_i\right) + \beta_i E(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}|\beta_i) E\big(-\beta_i \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i\big) \\
& + \beta_i E\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}|\beta_i\big) E(-\beta_i \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i) + \beta_i^2\big[1(s = s')\sigma_\epsilon^2 + E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}|\beta_i\big)\big] \\
& + \beta_i E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i\big) + \beta_i E\big(\boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i\big) + 1(s = s') E\big(\sigma_{\eta_i}^2\big) + E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i\big) \\
= \quad & \beta_i^2 E\left((w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{w}_{\mathcal{C}})\big(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{w}_{\mathcal{C}}\big)|\beta_i\right) + \beta_i^2\big[1(s = s')\sigma_\epsilon^2 + E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\epsilon}_{i\mathcal{C}}|\beta_i\big)\big] \\
& + 1(s = s') E\big(\sigma_{\eta_i}^2\big) + E\big(\boldsymbol{\theta}'_{s\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}} \boldsymbol{\theta}'_{s'\mathcal{C}} \boldsymbol{\eta}_{i\mathcal{C}}|\beta_i\big).
\end{aligned}
$$

For the version of equation (D.3) to hold, one set of sufficient conditions is thus

(1) $\beta_i E(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}|\beta_i)E(-\beta_i\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}) = 0,$

(2) $\beta_i E(w_{i\mathcal{C}'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}|\beta_i)E(-\beta_i\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}) = 0,$

(3) $E(\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}) = E(\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}|\beta_i),$

(4) $E(\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}) = E(\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i),$

(5) $E(\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}) = 0$, and

(6) $E(\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}) = 0.$

We now show that uniform benchmarking satisfies all the six conditions for all i. Since

$$\begin{aligned} & Cov(\beta_i\mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, -\beta_i\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i) \\ = \; & -\beta_i^2 Cov(\mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}) - Cov(\mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}) \\ = \; & -\beta_i^2\sigma_\epsilon^2\boldsymbol{\theta}'_{s'\mathcal{C}}\mathbf{z} - E(\sigma_{\eta_i}^2)\boldsymbol{\theta}'_{s'\mathcal{C}}\mathbf{z} \\ = \; & 0 \end{aligned}$$

under uniform benchmarking, it follows that

$$E(-\beta_i\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}) = E(-\beta_i\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i) = 0,$$

which implies that condition (1) is satisfied. Similarly,

$$\begin{aligned} & Cov(\beta_i\mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, -\beta_i\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i) \\ = \; & -\beta_i^2 Cov(\mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}) - Cov(\mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}) \\ = \; & -\beta_i^2\sigma_\epsilon^2\boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{z} - E(\sigma_{\eta_i}^2)\boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{z} \\ = \; & 0 \end{aligned}$$

guarantees

$$E(-\beta_i\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_{i,OLS}) = E(-\beta_i\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i) = 0,$$

implying the satisfaction of condition (2).

Since $(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}})$ and $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}$ are jointly normal when conditional on β_i and

$$Cov(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}|\beta_i) = \beta_i \sigma_\epsilon^2 \boldsymbol{\theta}'_{s\mathcal{C}}\mathbf{z} = 0$$

under uniform benchmarking, $(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}})$ and $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}$ are conditionally independent. Since $\boldsymbol{\theta}_{s\mathcal{C}} = \boldsymbol{\theta}_{s'\mathcal{C}}$, $(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}})$ and $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}$ are conditionally independent, leading to the satisfaction of condition (3). In exactly the same way, one can show that condition (4) is satisfied under uniform benchmarking.

Because, conditional on β_i, $(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}})$, $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}$, and $\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}$ are multivariate normal, the above zero covariance between $(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}})$ and $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}$ and a similar zero covariance between $(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}})$ and $\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}$, together with the independence between the two types of response errors, imply that $(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}})$, $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}$, and $\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}$ are mutually independent conditional on β_i. The conditional mutual independence implies

$$Cov\left(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i\right) = 0$$

and

$$Cov\left(\beta_i \mathbf{z}'\boldsymbol{\epsilon}_{i\mathcal{C}} + \mathbf{z}'\boldsymbol{\eta}_{i\mathcal{C}}, \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i\right) = 0.$$

It thus follows

$$E\left(\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_i\right) = E\left(\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i\right) = 0$$

and

$$E\left(\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i, \hat{\beta}_i\right) = E\left(\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}|\beta_i\right) = 0.$$

In words, conditions (5) and (6) are satisfied.

E. Method-of-Moments Estimators

This appendix contains the derivation of all of our method-of-moments estimators. With the help of additional assumptions, these estimators have analytic expressions and are thus computationally lighter compared to our semi-parametric and comprehensive MLE estimators. As highlighted in the paper, we only recommend the method-of-moments estimator for $Cov(x_i, w_{is})$; in all other cases, the method-of-moments estimator requires assuming that the stretcher β_i and the common-scale SWB w_{is} are independent, and as Appendix C shows, the estimator will be biased if instead they are correlated.

E.1. Method-of-Moments Estimator for $\boldsymbol{E(w_{is})}$

The key assumption is independence between β_i and w_{is}, which gives us

$$E(r_{is}) = Cov(\beta_i, w_{is}) + E(w_{is}) = E(w_{is}).$$

The estimator is thus

$$\frac{1}{I}\sum_{i=1}^{I} r_{is}.$$

E.2. Method-of-Moments Estimator for $\boldsymbol{Cov(x_i, w_{is})}$

To estimate $Cov(x_i, w_{is})$, the key assumption is zero co-skewness between the stretcher, the common-scale SWB, and the demographic x_i:

$$E\big[\big(\beta_i - E(\beta_i)\big)\big(w_{is} - E(w_{is})\big)\big(x_i - E(x_i)\big)\big] = 0.$$

Since the MMB $\hat{r}_i(h)$ is the variance-minimizing weighted average of CQ responses, $\boldsymbol{\theta}'_{sC}\boldsymbol{r}_{iC}$, that has mean h (see Section E.2.1 below),

$$
\begin{aligned}
Cov(x_i, r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}) \quad &= E\big[(\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}} + \epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}) + \eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}})\big(x_i - E(x_i)\big)\big] \\
&= E\big[\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}})\big(x_i - E(x_i)\big)\big] \\
&= E\big[\beta_i\big(w_{is} - E(w_{is})\big)\big(x_i - E(x_i)\big)\big] + [E(w_{is}) - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}]E\big[\beta_i\big(x_i - E(x_i)\big)\big] \\
&= Cov(x_i, w_{is}) + E\big[\big(\beta_i - E(\beta_i)\big)\big(w_{is} - E(w_{is})\big)\big(x_i - E(x_i)\big)\big] \\
&+[E(w_{is}) - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}]E\big[\beta_i\big(x_i - E(x_i)\big)\big] \\
&= Cov(x_i, w_{is}),
\end{aligned}
$$

where the last equality holds because of the zero-co-skewness assumption and mean-matched benchmarking ($h = E(w_{is})$).

Thus, the estimator is

$$
\frac{1}{I}\sum_{i=1}^{I}\left(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}} - \frac{1}{I}\sum_{i=1}^{I}(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})\right)\left(x_i - \frac{1}{I}\sum_{i=1}^{I}x_i\right).
$$

E.2.1. Mean-Matched Benchmark (MMB) As Mean-Matched Minimum Variance Weighted Average of CQ Ratings

To see that MMB is the mean-matched variance-minimizing weighted average of CQ ratings, we first derive its explicit expression as a weighted average of CQ ratings. Then we show that the weights minimize the variance of the benchmarked ratings while matching the mean h.

Appendix D.1 shows that

$$
\begin{aligned}
\hat{\beta}_{i,OLS} \quad &= \mathbf{z}'\boldsymbol{r}_{i\mathcal{C}} \\
&= \frac{C\widehat{\boldsymbol{w}}'_{\mathcal{C}} - \widehat{\boldsymbol{w}}'_{\mathcal{C}}\mathbf{1}_C\mathbf{1}'_C}{C\widehat{\boldsymbol{w}}'_{\mathcal{C}}\widehat{\boldsymbol{w}}_{\mathcal{C}} - \widehat{\boldsymbol{w}}'_{\mathcal{C}}\mathbf{1}_C\mathbf{1}'_C\widehat{\boldsymbol{w}}_{\mathcal{C}}}\boldsymbol{r}_{i\mathcal{C}}
\end{aligned}
$$

and

$$
\hat{a}_{i,OLS} = \frac{1}{C}\mathbf{1}'_C\big(\boldsymbol{r}_{i\mathcal{C}} - \hat{\beta}_{i,OLS}\widehat{\boldsymbol{w}}_{\mathcal{C}}\big).
$$

This implies that

$$\begin{aligned}
\hat{r}_i(h) &\equiv \hat{a}_{i,OLS} + \hat{\beta}_{i,OLS} h \\
&= \frac{1}{C}\mathbf{1}_C' \boldsymbol{r}_{i\mathcal{C}} + \left(h - \frac{1}{C}\mathbf{1}_C' \widehat{\boldsymbol{w}}_{\mathcal{C}}\right)\hat{\beta}_{i,OLS} \\
&= \left[\frac{1}{C}\mathbf{1}_C' + \left(h - \frac{1}{C}\mathbf{1}_C' \widehat{\boldsymbol{w}}_{\mathcal{C}}\right)\frac{C\widehat{\boldsymbol{w}}_{\mathcal{C}}' - \widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C\mathbf{1}_C'}{C\widehat{\boldsymbol{w}}_{\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} - \widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C\mathbf{1}_C'\widehat{\boldsymbol{w}}_{\mathcal{C}}}\right]\boldsymbol{r}_{i\mathcal{C}} \\
&= \left[\frac{\widehat{\boldsymbol{w}}_{\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} - h\widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C}{C\widehat{\boldsymbol{w}}_{\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} - \widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C\mathbf{1}_C'\widehat{\boldsymbol{w}}_{\mathcal{C}}}\mathbf{1}_C' + \frac{Ch - \mathbf{1}_C'\widehat{\boldsymbol{w}}_{\mathcal{C}}}{C\widehat{\boldsymbol{w}}_{\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} - \widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C\mathbf{1}_C'\widehat{\boldsymbol{w}}_{\mathcal{C}}}\widehat{\boldsymbol{w}}_{\mathcal{C}}'\right]\boldsymbol{r}_{i\mathcal{C}},
\end{aligned}$$

which shows that the MMB is a weighted average of $\boldsymbol{r}_{i\mathcal{C}}$.

To see that the weights are variance-minimizing and mean-matching, consider the following problem that minimizes the variance of the benchmarked SWB rating:

$$\min_{\boldsymbol{\theta}_{s\mathcal{C}}} Var(r_{is} - \boldsymbol{\theta}_{s\mathcal{C}}'\boldsymbol{r}_{i\mathcal{C}})$$
$$\text{s.t. } \boldsymbol{\theta}_{s\mathcal{C}}'\mathbf{1}_{\boldsymbol{C}} = 1 \text{ and } \boldsymbol{\theta}_{s\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} = h.$$

(In Appendix D.2, only the first constraint was present, and the objective function was only the noise component of the variance. In that case, uniform benchmarking was optimal.) Due to the second constraint, the weighting vector $\boldsymbol{\theta}_{s\mathcal{C}}$ varies across SWB questions, since the target mean h varies across the SWB questions.

The Lagrangian for the constrained minimization problem is

$$\Lambda = \frac{1}{2}\left(\sigma_\epsilon^2 E(\beta_i^2) + E\left(\sigma_{\eta_i}^2\right)\right)\boldsymbol{\theta}_{s\mathcal{C}}'\boldsymbol{\theta}_{s\mathcal{C}} + \lambda(1 - \boldsymbol{\theta}_{s\mathcal{C}}'\mathbf{1}_C) + \kappa(h - \boldsymbol{\theta}_{s\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}}),$$

where λ and κ are Lagrange multipliers. Matrix calculus leads to the first-order condition:

$$\left(\sigma_\epsilon^2 E(\beta_i^2) + E\left(\sigma_{\eta_i}^2\right)\right)\boldsymbol{\theta}_{s\mathcal{C}} - \lambda\mathbf{1}_C - \kappa\widehat{\boldsymbol{w}}_{\mathcal{C}} = \mathbf{0}_C.$$

Solving for $\boldsymbol{\theta}_{s\mathcal{C}}$ gives

$$\boldsymbol{\theta}_{s\mathcal{C}}^{*} = \left(\sigma_{\epsilon}^{2} E(\beta_i^2) + E\left(\sigma_{\eta_i}^2\right)\right)^{-1} (\lambda \mathbf{1}_C + \kappa \widehat{\boldsymbol{w}}_{\mathcal{C}}).$$

Inserting $\boldsymbol{\theta}_{s\mathcal{C}}^{*}$ into the two constraints yields a system of two simultaneous equations in the Lagrange multipliers λ and κ,

$$\left(\sigma_{\epsilon}^{2} E(\beta_i^2) + E\left(\sigma_{\eta_i}^2\right)\right)^{-1} \begin{bmatrix} (\lambda \mathbf{1}_C' + \kappa \widehat{\boldsymbol{w}}_{\mathcal{C}}')\mathbf{1}_C \\ (\lambda \mathbf{1}_C' + \kappa \widehat{\boldsymbol{w}}_{\mathcal{C}}')\widehat{\boldsymbol{w}}_{\mathcal{C}} \end{bmatrix} = \begin{bmatrix} 1 \\ h \end{bmatrix},$$

which can be simplified to

$$\begin{bmatrix} \lambda \\ \kappa \end{bmatrix} = \frac{\sigma_{\epsilon}^{2} E(\beta_i^2) + E\left(\sigma_{\eta_i}^2\right)}{C\widehat{\boldsymbol{w}}_{\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} - \widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C\mathbf{1}_C'\widehat{\boldsymbol{w}}_{\mathcal{C}}} \begin{bmatrix} \widehat{\boldsymbol{w}}_{\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} - h\widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C \\ Ch - \mathbf{1}_C'\widehat{\boldsymbol{w}}_{\mathcal{C}} \end{bmatrix}.$$

Substituting the expressions for the Lagrange multipliers into the first-order condition gives the optimal weights:

$$\boldsymbol{\theta}_{s\mathcal{C}}^{*} = \frac{\widehat{\boldsymbol{w}}_{\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} - h\widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C}{C\widehat{\boldsymbol{w}}_{\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} - \widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C\mathbf{1}_C'\widehat{\boldsymbol{w}}_{\mathcal{C}}} \mathbf{1}_C + \frac{Ch - \mathbf{1}_C'\widehat{\boldsymbol{w}}_{\mathcal{C}}}{C\widehat{\boldsymbol{w}}_{\mathcal{C}}'\widehat{\boldsymbol{w}}_{\mathcal{C}} - \widehat{\boldsymbol{w}}_{\mathcal{C}}'\mathbf{1}_C\mathbf{1}_C'\widehat{\boldsymbol{w}}_{\mathcal{C}}} \widehat{\boldsymbol{w}}_{\mathcal{C}},$$

which are exactly the MMB's weights.

E.3. Method-of-Moments Estimator for $\boldsymbol{Cov(w_{is}, w_{is'})}$

Because $Var(w_{is})$ is a special case of $Cov(w_{is}, w_{is'})$ where $s' = s$, we only show the method-of-moments estimator for $Cov(w_{is}, w_{is'})$. The key assumption is independence between β_i and w_{is}.

Note that

$$
\begin{aligned}
& Cov\left(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}, r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\right) \\
= \quad & Cov(\beta_i(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}} + \epsilon_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}) + \eta_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}, \\
& \beta_i\left(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{w}_{\mathcal{C}} + \epsilon_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\epsilon}_{i\mathcal{C}}\right) + \eta_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{\eta}_{i\mathcal{C}}) \\
= \quad & E\left[\beta_i^2 Cov\left(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}, w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{w}_{\mathcal{C}} | \beta_i\right)\right] + Cov\left(\beta_i E(w_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}|\beta_i), \beta_i E\left(w_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}|\beta_i\right)\right) \\
& +E(\beta_i^2)[1(s = s') + \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}]\sigma_\epsilon^2 + [1(s = s') + \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}]E\left(\sigma_{\eta_i}^2\right) \\
= \quad & E(\beta_i^2)Cov(w_{is}, w_{is'}) + [E(w_{is}) - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}]\left[E(w_{is'}) - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}\right][E(\beta_i^2) - 1] \\
& +[1(s = s') + \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}]\left[E(\beta_i^2)\sigma_\epsilon^2 + E\left(\sigma_{\eta_i}^2\right)\right],
\end{aligned}
$$

where the last equality follows from the independence between β_i and w_{is}.

With $\boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}$ matching $E\left(w_{ij}\right)$ or $\boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{w}_{\mathcal{C}}$ matching $E(w_{is'})$, we have

$$
\begin{aligned}
Cov(w_{is}, w_{is'}) = \quad & \frac{1}{E(\beta_i^2)} Cov\left(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}, r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\right) \\
& -[1(s = s') + \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}]\left[\sigma_\epsilon^2 + \frac{E\left(\sigma_{\eta_i}^2\right)}{E(\beta_i^2)}\right].
\end{aligned}
$$

It then follows that the estimator is

$$
\begin{aligned}
& \frac{1}{\hat{E}(\beta_i^2)} \frac{1}{I} \sum_{i=1}^{I} \left(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}} - \frac{1}{I}\sum_{i=1}^{I}(r_{is} - \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}})\right)\left(r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}} - \frac{1}{I}\sum_{i=1}^{I}\left(r_{is'} - \boldsymbol{\theta}'_{s'\mathcal{C}}\boldsymbol{r}_{i\mathcal{C}}\right)\right) \\
& -[1(s = s') + \boldsymbol{\theta}'_{s\mathcal{C}}\boldsymbol{\theta}_{s'\mathcal{C}}]\left[\hat{\sigma}_\epsilon^2 + \frac{\hat{E}\left(\sigma_{\eta_i}^2\right)}{\hat{E}(\beta_i^2)}\right],
\end{aligned}
$$

where the CQ-only MLE provides the estimates $\hat{E}(\beta_i^2)$, $\hat{\sigma}_\epsilon^2$, and $\hat{E}\left(\sigma_{\eta_i}^2\right)$.

F. Maximum-Likelihood Estimators

This appendix provides additional information on the MLEs discussed in the paper.

F.1. CQ-Only MLE

This subsection details the likelihood function and implementation for the MLE that estimates the scale-use hyperparameters using CQ ratings. Given the distributional assumptions in Section IV and the translation function (equation 4), the sub-likelihood function for respondent i's CQ ratings ($\boldsymbol{r}_{i\mathcal{C}}$) is

$$
\begin{aligned}
& l_i\left(\boldsymbol{r}_{i\mathcal{C}} \mid \alpha_i, \beta_i, \gamma, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \sigma_{\eta_i}\right) \\
= \quad & \left(\beta_i^2 \sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-\frac{C}{2}} \phi\left(\left(\beta_i^2 \sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-\frac{1}{2}} [\boldsymbol{r}_{i\mathcal{C}} - (\alpha_i + \gamma)\mathbf{1}_C - \beta_i(\boldsymbol{w}_{\mathcal{C}} - \gamma \mathbf{1}_C)]\right).
\end{aligned}
$$

where $\mathbf{1}_C$ denotes the C-dimensional vector full of ones, $\boldsymbol{w}_{\mathcal{C}} \equiv [w_1, \ldots, w_C]'$, and ϕ the multivariate standard-normal probability density function.

Assuming independence across respondents, the log-likelihood function for CQ ratings from all respondents ($\boldsymbol{r}_{\cdot\mathcal{C}}$) is thus

$$
\begin{aligned}
& \ln \mathcal{L}\left(\boldsymbol{r}_{\cdot\mathcal{C}} | \sigma_\alpha, \sigma_\beta, \gamma, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \mu_{\ln\sigma_\eta}, \sigma_{\ln\sigma_\eta}\right) \\
= \quad & \sum_{i=1}^{I} \ln \iiint l_i\left(\boldsymbol{r}_{iC} \mid \alpha_i, \beta_i, \gamma, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \sigma_{\eta_i}\right) f(\alpha_i) f(\beta_i) f\left(\sigma_{\eta_i}\right) d\alpha_i d\beta_i d\sigma_{\eta_i},
\end{aligned}
$$

where f denotes a probability density function.

The three-dimensional integration renders the MLE too slow to be practically useful. To deal with this computational issue, we use hierarchical modeling, in which the log-likelihood function is written

$$
\begin{aligned}
& \ln \mathcal{L}\left(\boldsymbol{r}_{\cdot\mathcal{C}} | \{\alpha_i\}, \{\beta_i\}, \gamma, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \{\sigma_{\eta_i}\}\right) \\
= \quad & \sum_{i=1}^{I} \ln l_i\left(\boldsymbol{r}_{i\mathcal{C}} \mid \alpha_i, \beta_i, \gamma, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \sigma_{\eta_i}\right),
\end{aligned}
$$

and α_i, β_i, and σ_{η_i} are drawn from $\mathcal{N}(0, \sigma_\alpha^2)$, $\mathcal{N}\left(1, \sigma_\beta\right)$, and $\ln \mathcal{N}\left(\mu_{\ln\sigma_\eta}, \sigma_{\ln\sigma_\eta}^2\right)$ independently. Because α_i, β_i, and σ_{η_i} are at the individual level while $\boldsymbol{r}_{i\mathcal{C}}$ is at the individual-CQ level, the model can be viewed as a hierarchical model. The hierarchical modeling approach thus should give consistent and efficient estimates of the parameters of interest. Le Cam (1953) shows that such Bayesian point estimators are consistent and efficient estimators in the frequentist sense.

The MLE takes as an input $\boldsymbol{w}_{\mathcal{C}}$ that is estimated by $1/I \sum_{i=1}^{I} \boldsymbol{r}_{i\mathcal{C}}$, which simulations suggest to be a very good approximation with a few hundred respondents. We use Stan's no-U-turn implementation of Hamiltonian Monte Carlo to sample from the hierarchical model. We specify 4 chains, each with 1,000 draws of warm-up and 1,000 draws of sampling, resulting in a total of 4,000 draws for each parameter of interest. We report the posterior modes for the hyperparameters of interest, exploiting the equivalence between maximizing the likelihood function and estimating the posterior mode under uniform priors. We bootstrap standard errors through 100 repetitions of sampling with replacement.

F.2. Comprehensive MLE

This subsection details the likelihood function for the comprehensive MLE estimator in Section V.C, which uses both SWB and CQ ratings.

The sub-likelihood function for respondent i's SWB ratings is

$$\begin{aligned} & l_{i\mathcal{S}}\left(\boldsymbol{r}_{i\mathcal{S}} | a_i, \beta_i, \boldsymbol{w}_{i\mathcal{S}}, \sigma_\epsilon, \sigma_{\eta_i}\right) \\ = \quad & \left(\beta_i^2 \sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-\frac{S}{2}} \phi\left(\left(\beta_i^2 \sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-\frac{1}{2}} (\boldsymbol{r}_{i\mathcal{S}} - a_i \mathbf{1}_S - \beta_i \boldsymbol{w}_{i\mathcal{S}})\right), \end{aligned}$$

where $\mathcal{S}$ is the set of S SWB questions, $\mathbf{1}_S$ denotes the S-dimensional vector full of ones, and $\boldsymbol{w}_{i\mathcal{S}} \equiv [w_{i1}, \dots, w_{iS}]'$.

Similarly, the sub-likelihood function for the same respondent's CQ ratings is

$$\begin{aligned} & l_{i\mathcal{C}}\left(\boldsymbol{r}_{i\mathcal{C}} | a_i, \beta_i, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \sigma_{\eta_i}\right) \\ = \quad & \left(\beta_i^2 \sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-\frac{C}{2}} \phi\left(\left(\beta_i^2 \sigma_\epsilon^2 + \sigma_{\eta_i}^2\right)^{-\frac{1}{2}} (\boldsymbol{r}_{i\mathcal{C}} - a_i \mathbf{1}_C - \beta_i \boldsymbol{w}_{\mathcal{C}})\right). \end{aligned}$$

The joint distribution of $(a_i, \beta_i, \boldsymbol{w}_{i\mathcal{S}}{}')$ in equation (15),

$$\begin{bmatrix} \alpha_i \\ \beta_i \\ w_{i1} \\ \vdots \\ w_{i\mathcal{S}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 1 \\ \boldsymbol{x}_i'\boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{x}_i'\boldsymbol{b}_S \end{bmatrix}, \begin{bmatrix} \sigma_\alpha^2 & \sigma_{\alpha,\beta} & \sigma_{\alpha,w_1} & \cdots & \sigma_{\alpha,w_S} \\ \sigma_{\alpha,\beta} & \sigma_\beta^2 & \sigma_{\beta,w_1} & \cdots & \sigma_{\beta,w_S} \\ \sigma_{\alpha,w_1} & \sigma_{\beta,w_1} & \sigma_{w_1}^2 & \cdots & \sigma_{w_1,w_S} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{\alpha,w_S} & \sigma_{\beta,w_S} & \sigma_{w_1,w_S} & \cdots & \sigma_{w_S}^2 \end{bmatrix} \right)$$
$$\equiv \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_S}, \boldsymbol{\Omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_S}\right),$$

implies their density:

$$f(\alpha_i, \beta_i, \boldsymbol{w}_{i\mathcal{S}}) = \left|\boldsymbol{\Omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_S}\right|^{-\frac{1}{2}} \phi\left(\boldsymbol{\Omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_S}^{-\frac{1}{2}}\left([\alpha_{\mathrm{i}}, \beta_i, \boldsymbol{w}_{i\mathcal{S}}']' - \boldsymbol{\mu}_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_S}\right)\right).$$

Under independence across respondents, the log-likelihood function for both types of ratings from all respondents $(\boldsymbol{r}_{\cdot\mathcal{S}}, \boldsymbol{r}_{\cdot\mathcal{C}})$ is thus

$$\begin{aligned} &\ln \mathcal{L}\left(\boldsymbol{r}_{\cdot\mathcal{S}}, \boldsymbol{r}_{\cdot\mathcal{C}} | \boldsymbol{\mu}_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_S}, \boldsymbol{\Omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_S}, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \mu_{\ln\sigma_\eta}, \sigma_{\ln\sigma_\eta}\right) \\ = &\sum_{i=1}^{I} \ln \int \cdots \int l_{i\mathcal{S}}\left(\boldsymbol{r}_{i\mathcal{S}} | \alpha_i, \beta_i, \boldsymbol{w}_{i\mathcal{S}}, \sigma_\epsilon, \sigma_{\eta_i}\right) l_{i\mathcal{C}}\left(\boldsymbol{r}_{i\mathcal{C}} | \alpha_i, \beta_i, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \sigma_{\eta_i}\right) f(\alpha_i, \beta_i, \boldsymbol{w}_{i\mathcal{S}}) f\left(\sigma_{\eta_i}\right) d\alpha_i d\beta_i d\boldsymbol{w}_{i\mathcal{S}} d\sigma_{\eta_i}. \end{aligned}$$

The high $(S + 3)$-dimensional integration renders direct estimation of this MLE infeasible. To deal with this, we again use hierarchical modeling where the log-likelihood function becomes

$$\begin{aligned} &\ln \mathcal{L}\left(\boldsymbol{r}_{\cdot\mathcal{S}}, \boldsymbol{r}_{\cdot\mathcal{C}} | \{\alpha_i\}, \{\beta_i\}, \boldsymbol{w}_{\cdot\mathcal{S}}, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \{\sigma_{\eta_i}\}\right) \\ = &\sum_{i=1}^{I} \left[\ln l_{i\mathcal{S}}\left(\boldsymbol{r}_{i\mathcal{S}} | \alpha_i, \beta_i, \boldsymbol{w}_{i\mathcal{S}}, \sigma_\epsilon, \sigma_{\eta_i}\right) + \ln l_{i\mathcal{C}}\left(\boldsymbol{r}_{i\mathcal{C}} | \alpha_i, \beta_i, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, \sigma_{\eta_i}\right)\right], \end{aligned}$$

and $(\alpha_{\mathrm{i}}, \beta_i, \boldsymbol{w}_{i\mathcal{S}}')$ and σ_{η_i} will be drawn from $\mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_S}, \boldsymbol{\Omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\mathbf{w}_S}\right)$ and $\ln \mathcal{N}\left(\mu_{\ln\sigma_\eta}, \sigma_{\ln\sigma_\eta}^2\right)$, respectively.

The comprehensive MLE is computationally the most intensive among our estimators. To improve its stability, we fix the two hyperparameters corresponding to the standard deviations of the gross shifter and stretcher (σ_α and σ_β) to their estimates from the CQ-only MLE.

F.3. Allowing for Top- and Bottom-Coding

About 5.0% of the SWB responses in our Baseline data are exactly 100, and roughly 0.9% are exactly 0. Our main analyses treated these boundary responses as truly equal to 100 and 0. In this subsection, we instead accommodate the possibility that these responses might have been even more extreme if the scale were not constrained to the 0-100 interval.

To accommodate top-coding and bottom-coding, we extend our comprehensive MLE with a straightforward Tobit-style adjustment, making the estimator's computation even heavier. The estimator treats a boundary response as the truncated value of a true, unobserved latent response. Tables F.1-3 report that the estimates of the hyperparameters, $E(w_{is})$, and demographic regression are basically unchanged by accommodating top- and bottom-coding. Only the estimates of $Var(w_{is})$ are noticeably affected—the estimates and their standard errors increased, which is consistent with adjustments for top- and bottom-coding—however, the magnitude of these differences remains relatively minor.

Table F.1. MLE Estimates of Hyperparameters with and without Top- and Bottom-Coding Adjustment

Hyperparameter	Without adjustment	With adjustment
σ_α	7.69	7.82
	(0.12)	(0.14)
γ	60.22	60.27
	(1.16)	(1.08)
σ_β	0.29	0.29
	(0.01)	(0.01)
σ_ϵ	6.38	7.38
	(0.32)	(0.27)
$\mu_{\ln\sigma_\eta}$	2.52	2.50
	(0.02)	(0.02)
$\sigma_{\ln\sigma_\eta}$	0.34	0.39
	(0.01)	(0.01)

Notes: Sample includes 3,358 Baseline respondents who passed QC. Standard errors in parentheses. Estimates from the comprehensive MLE estimator.

Table F.2. Estimates of $E(w_{is})$ with and without Top- and Bottom-Coding Adjustment

SWB	Without adjustment	With adjustment
Life Satisfaction	67.69 (0.40)	68.16 (0.43)
Happiness	68.16 (0.39)	68.58 (0.39)
Worthwhileness	69.99 (0.40)	70.46 (0.41)
No Anxiety	55.27 (0.53)	55.34 (0.53)

Notes: Sample includes 3,358 Baseline respondents who passed QC. Standard errors in parentheses. Scale-use correction through the comprehensive MLE estimator with and without adjustment for top- and bottom-coding.

Table F.3. Life Satisfaction and "No Anxiety" Regressions with and without Top- and Bottom-Coding Adjustment

	Life Satisfaction		No Anxiety	
Demographics	Without adjustment	With adjustment	Without adjustment	With adjustment
Demeaned age/10	1.6†††	1.5†††	3.9†††	4.2†††
	(0.4)	(0.4)	(0.4)	(0.5)
(Demeaned age)^2/100	1.6†††	1.7†††	0.9†††	1.0†††
	(0.2)	(0.2)	(0.3)	(0.3)
Log(HH income)	5.3†††	5.3†††	2.7†††	3.0†††
	(0.6)	(0.6)	(0.8)	(0.9)
Unemployed	-7.0†††	-8.0†††	-5.6†††	-6.1†††
	(1.3)	(1.3)	(1.8)	(1.7)
Employed part-time	-2.3	-3.0††	-4.6†††	-5.3†††
	(1.2)	(1.3)	(1.5)	(1.5)
Out of labor force/other	-4.5†††	-5.4†††	-4.3††	-4.2††
	(1.3)	(1.3)	(1.7)	(1.7)
Married, not separated	8.6†††	8.7†††	3.1††	2.9†
	(1.0)	(1.0)	(1.3)	(1.3)
Ever divorced	2.1	2.4	0.5	-0.3
	(1.2)	(1.3)	(1.7)	(1.7)
Have ≥1 child	3.8†††	4.2†††	1.7	1.5
	(0.9)	(1.0)	(1.2)	(1.3)
Log(HH size)	-2.4††	-2.6††	-2.2	-2.3
	(1.0)	(1.0)	(1.1)	(1.2)
College grad	1.5	1.7	2.3†	2.5†
	(0.9)	(0.9)	(1.0)	(1.1)
Male	0.1	-0.1	5.7†††	5.9†††
	(0.9)	(0.9)	(1.0)	(1.1)

	Life Satisfaction		No Anxiety	
Religious attendance (0 to 5, "Never" to "More than once a week")	1.5†††	1.6†††	1.0†††	1.0†††
	(0.2)	(0.2)	(0.3)	(0.3)
Asian	0.4	0.6	3.2	3.2
	(1.8)	(1.9)	(1.9)	(1.9)
Obs.	3,358	3,358	3,358	3,358

Notes: Sample includes Baseline respondents who passed QC. Standard errors in parentheses. Scale-use adjustment through the MOM estimator. Daggers signal false-discovery-rate (FDR) significance levels using the Benjamini-Hochberg procedure applied to the 29 *p*-values in each column separately (variables included in FDR correction also include additional race and employment status indicators, as well as indicators for region, day of week, political party, obesity, and population density; "Other" categories in race and employment status are excluded—we do not pose or report hypothesis tests for them); †††, ††, and † indicate significance at the 1-percent, 5-percent, and 10-percent levels, respectively. Scale-use correction through the comprehensive MLE estimator with and without adjustment for top- and bottom-coding.

Table F.4. Estimates of $Var(w_{is})$ with and without Top- and Bottom-Coding Adjustment

SWB	Without adjustment	With adjustment
Life Satisfaction	274.0 (10.9)	300.1 (12.7)
Happiness	264.1 (9.4)	284.2 (10.9)
Worthwhileness	220.8 (9.7)	242.2 (11.1)
No Anxiety	369.1 (12.4)	408.0 (14.4)

Notes: Sample includes 3,358 Baseline respondents who passed QC. Standard errors in parentheses. Scale-use correction through the comprehensive MLE estimator with and without adjustment for top- and bottom-coding.

F.4. Two-Wave CQ-Only MLE

To gauge the persistence of scale-use parameters (and response errors), we extend the CQ-only MLE in Section F.1 above to accommodate CQ responses from two waves of surveys. We start by introducing time (t) to equation (4):

$$r_{ic,t} = \alpha_{i,t} + (1 - \beta_{i,t})\gamma + \beta_{i,t}(w_{c,t} + \epsilon_{ic,t}) + \eta_{ic,t},$$

and extend equations (5) and (6) to

$$\boldsymbol{\alpha}_i \equiv \begin{bmatrix}\alpha_{i,1}\\ \alpha_{i,2}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_\alpha^2 & f_{PV(\alpha)}\sigma_\alpha^2\\ f_{PV(\alpha)}\sigma_\alpha^2 & \sigma_\alpha^2\end{bmatrix}\right)$$

$$\boldsymbol{\beta}_i \equiv \begin{bmatrix}\beta_{i,1}\\ \beta_{i,2}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}\sigma_\beta^2 & f_{PV(\beta)}\sigma_\beta^2\\ f_{PV(\beta)}\sigma_\beta^2 & \sigma_\beta^2\end{bmatrix}\right)$$

$$\begin{bmatrix}\epsilon_{ic,1}\\ \epsilon_{ic,2}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_\epsilon^2 & f_{PV(\epsilon)}\sigma_\epsilon^2\\ f_{PV(\epsilon)}\sigma_\epsilon^2 & \sigma_\epsilon^2\end{bmatrix}\right)$$

$$\begin{bmatrix}\eta_{ic,1}\\ \eta_{ic,2}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_{\eta i}^2 & f_{PV(\eta)}\sigma_{\eta i}^2\\ f_{PV(\eta)}\sigma_{\eta i}^2 & \sigma_{\eta i}^2\end{bmatrix}\right),$$

where $f_{PV(\alpha)}$, $f_{PV(\beta)}$, $f_{PV(\epsilon)}$, and $f_{PV(\eta)}$ denote the fractions of the variances of the shifter, stretcher, and two response errors that are persistent across waves, respectively. We allow w_c to vary across waves so that the normalizations of $E(\alpha_{i,t}) = 0$ and $E(\beta_{i,t}) = 1$ hold in both waves. We do not allow γ to vary across waves so that the gross shifter's variance and its correlation with $\beta_{i,t}$ do not vary across time, just as the stretcher's and error terms' variances and their correlations do not vary across time.[11]

These imply

$$\begin{aligned}\begin{bmatrix}r_{ic,1}\\ r_{ic,2}\end{bmatrix} &\sim \mathcal{N}\left(\begin{bmatrix}\alpha_{i,1} + (1-\beta_{i,1})\gamma + \beta_{i,1}w_{c,1}\\ \alpha_{i,2} + (1-\beta_{i,2})\gamma + \beta_{i,2}w_{c,2}\end{bmatrix}, \begin{bmatrix}\beta_{i,1}^2\sigma_\epsilon^2 + \sigma_{\eta i}^2 & \beta_{i,1}\beta_{i,2}f_{PV(\epsilon)}\sigma_\epsilon^2 + f_{PV(\eta)}\sigma_{\eta i}^2\\ \beta_{i,1}\beta_{i,2}f_{PV(\epsilon)}\sigma_\epsilon^2 + f_{PV(\eta)}\sigma_{\eta i}^2 & \beta_{i,2}^2\sigma_\epsilon^2 + \sigma_{\eta i}^2\end{bmatrix}\right)\\ &\equiv \mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Omega}_c).\end{aligned}$$

The sub-likelihood function for respondent i's responses to CQ c across both waves ($\boldsymbol{r}_{ic} \equiv [r_{ic,1}, r_{ic,2}]'$) is

$$\begin{aligned}&l_{ic}\left(\boldsymbol{r}_{ic}|\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \gamma, \boldsymbol{w}_c, \sigma_\epsilon, f_{PV(\epsilon)}, \sigma_{\eta_i}, f_{PV(\eta)}\right)\\ =\ &|\boldsymbol{\Omega}_c|^{-\frac{1}{2}}\phi\left(\boldsymbol{\Omega}_c{}^{-\frac{1}{2}}(\boldsymbol{r}_{ic} - \boldsymbol{\mu}_c)\right),\end{aligned}$$

where $\boldsymbol{w}_c \equiv [w_{c,1}, w_{c,2}]'$.

[11] In a variant of this MLE that allows γ to vary between the two waves, the cross-wave difference of γ is estimated to be small, 3.6 (SE = 1.0), relative to γ's point estimate of about 60.

Assuming independence across respondents and CQs, the log-likelihood function for CQ ratings from all respondents ($\boldsymbol{r}_{\cdot\mathcal{C}}$) is thus

$$\ln \mathcal{L}\left(\boldsymbol{r}_{\cdot\mathcal{C}} | \sigma_\alpha, f_{PV(\alpha)}, \sigma_\beta, f_{PV(\beta)}, \gamma, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, f_{PV(\epsilon)}, \mu_{\ln\sigma_\eta}, \sigma_{\ln\sigma_\eta}, f_{PV(\eta)}\right)$$
$$= \sum_{i=1}^{I} \ln \int \cdots \int \prod_{c=1}^{C} l_{ic}\left(\boldsymbol{r}_{ic} | \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \gamma, \boldsymbol{w}_c, \sigma_\epsilon, f_{PV(\epsilon)}, \sigma_{\eta_i}, f_{PV(\eta)}\right) f(\boldsymbol{\alpha}_i) f(\boldsymbol{\beta}_i) f\left(\sigma_{\eta_i}\right) d\boldsymbol{\alpha}_i d\boldsymbol{\beta}_i d\sigma_{\eta_i}.$$

Utilizing hierarchical modeling to make the estimation feasible leads to the log-likelihood function

$$\ln\mathcal{L}\left(\boldsymbol{r}_{\cdot\mathcal{C}} | \{\boldsymbol{\alpha}_i\}, \{\boldsymbol{\beta}_i\}, \gamma, \boldsymbol{w}_{\mathcal{C}}, \sigma_\epsilon, f_{PV(\epsilon)}, \sigma_{\eta_i}, f_{PV(\eta)}\right)$$
$$= \sum_{i=1}^{I} \sum_{c=1}^{C} \ln l_{ic}\left(\boldsymbol{r}_{ic} | \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \gamma, \boldsymbol{w}_c, \sigma_\epsilon, f_{PV(\epsilon)}, \sigma_{\eta_i}, f_{PV(\eta)}\right),$$

where $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ will be drawn from their respectively two-wave distributions specified at the start of this subsection, and σ_{η_i} from $\ln \mathcal{N}\left(\mu_{\ln\sigma_\eta}, \sigma^2_{\ln\sigma_\eta}\right)$.

Table F.5 reports estimates from applying this MLE to Baseline and Bottomless responses on the 18 Baseline CQs from 2,472 respondents who provided such responses. The estimates suggest that the stretcher, shifter, and perception error are highly persistent while the trembling-hand error is not persistent.

Table F.5. Estimates From 2-Wave CQ-Only MLE

σ_α	7.48 (0.15)
γ	60.77 (1.02)
σ_β	0.27 (0.01)
σ_ϵ	7.96 (0.28)
$\mu_{\ln\sigma_\eta}$	2.64 (0.01)

$\sigma_{\ln\sigma_\eta}$	0.30 (0.01)
$f_{PV(\alpha)}$	55.41% (2.18%)
$f_{PV(\beta)}$	88.50% (2.67%)
$f_{PV(\epsilon)}$	99.65% (0.16%)
$f_{PV(\eta)}$	3.21% (1.96%)

Notes: The sample is 2,472 respondents who completed the 18 Baseline CQs in both Baseline survey and block 1 of Bottomless survey and passed quality control. Standard errors in parentheses.

Appendix F Reference

Le Cam, Lucien. "On Some Asymptotic Properties of Maximum Likelihood Estimates and Related Bayes' Estimates." *University of California Publications in Statistics* 1 (1953): 277-330.

G. Simulations

This appendix details simulations we ran for investigating the finite-sample performance, robustness, and convergence speed of our scale-use-adjustment estimators.

In our baseline simulations, we assume that our model of the translation function in equation (4) is correct, with parameter values similar to the estimates in Table 4,[12] and we examine a variety of data-generating processes for common-scale SWB (see Section G.1). As expected, the estimators are biased when their assumptions are not met; for example, the MOM and comprehensive MLE estimators for $Cov(x_i, w_{is})$ are biased when the co-skewness between x_i, w_{is} and β_i is non-zero. In all the simulations we examine in which the respective assumptions of the MOM and comprehensive MLE estimators are satisfied, we cannot statistically detect bias. We cannot detect bias for the semi-parametric estimators across any of our baseline simulations.

[12] The simulations were done before the actual data underlying results in the main paper were finalized, leading to small differences between Table 4 and the "Truth" column of Table G.3. A comparison of the two shows that these differences are primarily evident in parameters associated with response errors. The more important parameter values pertaining to the shifter, stretcher, and center are virtually the same.

To examine robustness to misspecification, we conduct three sets of simulations: (i) the data-generating process has a quadratic rather than a linear translation function, with the variance of the quadratic coefficient across individuals equal to the value we estimate from our data (see Web Appendix B); (ii) the error terms in the data-generating process are *t*-distributed (with 5 degrees of freedom, so that they have fatter tails) rather than normally distributed, with unchanged variance; and (iii) common-scale SWB w_{is} in the data-generating process is quadratic in the stretcher β_i (in contrast, we estimate our semi-parametric estimators with $K = 1$, which effectively assumes that the conditional first-moments of w_{is} are linear in β_i). Naturally, in all cases, our misspecified models generate biased parameter estimates. However, in all cases, the semi-parametric estimator generates estimates that are reasonably close to the truth in all scenarios, albeit with standard errors that tend to be somewhat larger than in the baseline simulations.

Our baseline and robustness simulations above use 18 CQs and 3000 respondents to roughly match our Baseline data. To study the convergence speed of our estimators, we also run simulations with $C \in \{2, 3, 9, 27, 81\}$ and $I \in \{100, 1000, 10000\}$. Under our data-generating process, we find that 2 CQs appear to be sufficient for the semi-parametric estimators of $E(w_{is})$ and $Cov(w_{is}, x_i)$ to converge in I at square-root speed. For the comprehensive MLE estimator for $Var(w_{is})$, at least 3 CQs seem necessary to approximate square-root convergence rate in I.

G.1. Data-Generating Processes

We simulate individuals' responses to 12 SWB questions, indexed by s, according to equation (4) and the following process of w_{is}:

$$w_{is} = a_s + b_s x_i + c_s \beta_i + d_s x_i \beta_i + e_s\left(v_{1,s} f_{1,i} + 0.32 v_{2,s} f_{2,i} + \cdots + 0.32 v_{12,s} f_{12,i}\right),$$

where the coefficients a_s, b_s, c_s, d_s and e_s vary across s. We employ a factor structure (the term in parentheses) to model cross-s correlations between the w_{is}'s. The factors are independent of both x_i and β_i, and (for no particular reason) we choose the same number of factors as the number of SWB questions (see below for details about how the $f_{k,i}$'s and $v_{k,i}$'s are determined). To examine our methods' effectiveness in mitigating biases arising from scale-use heterogeneity, we choose the values of the parameters to create simple variations that sometimes induce biases for a naïve estimator that ignores scale-use heterogeneity. Table G.1 details the coefficients' values for each set of SWB-question responses we simulate, as well as what these coefficient values imply for the signs of the covariance between x_i and w_{is}, the covariance between β_i and w_{is}, and the co-skewness between β_i, w_{is}, and x_i. These quantities correspond to key bias terms when scale-use heterogeneity is ignored (see Section V).

Table G.1. a_s, b_s, c_s, d_s and e_s in Simulations

s	a_s	b_s	c_s	d_s	e_s	$Cov(x_i, w_{is})$	$Cov(\beta_i, w_{is})$	$E[(\beta_i - E(\beta_i))(w_{is} - E(w_{is}))(x_i - E(x_i))]$
1	50	0	0	0	$10\sqrt{12}$	0	0	0
2	50	0	0	0	$20\sqrt{12}$	0	0	0
3	50	20	0	0	$10\sqrt{12}$	+	0	0
4	50	10	0	0	$10\sqrt{12}$	+	0	0
5	50	0	10	0	$10\sqrt{12}$	0	+	0
6	50	0	20	0	$10\sqrt{12}$	0	+	0
7	50	20	10	0	$10\sqrt{12}$	+	+	0
8	50	0	0	15	$10\sqrt{12}$	+	+	+
9	50	0	0	30	$10\sqrt{12}$	+	+	+
10	50	20	0	15	$10\sqrt{12}$	+	+	+
11	50	0	10	15	$10\sqrt{12}$	+	+	+
12	50	20	10	15	$10\sqrt{12}$	+	+	+

Distributional assumptions about the shifter, stretcher, and response errors are the same as in equations (5) and (6). The distributional assumptions for x_i and $f_{k,i}$'s are the following:

- x_i is distributed as Bernoulli with $Prob(x_i = 0) = Prob(x_i = 1) = 0.5$.
- $f_{k,i}$'s are mutually independent factors distributed as standard normal, which are also independent from both β_i and x_i.
- $\mathbf{v}_k \equiv [v_{k,1}, v_{k,2}, \cdots, v_{k,12}]'$ is the vector of quasi-loadings for the k^{th} factor with $\|\mathbf{v}_k\| = 1$, and $\langle \mathbf{v}_k, \mathbf{v}_{k'} \rangle = 0$ for $k \neq k'$.
- The matrix

$$[\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{12}]$$

is the orthonormalized version of the following 12×12 dimensional matrix,

$$\begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

through the modified Gram-Schmidt process (Trefethen and Bau, 1997).

G.1.1. Data-Generating Processes When $\boldsymbol{w_{is}}$ Is Quadratic in $\boldsymbol{\beta_i}$

We simulate individuals' responses to 16 SWB questions according to the translation function in equation (4) and the following process of w_{is}:

$$w_{is} = a_s + b_s x_i + c_s \beta_i + d_s x_i \beta_i + g_s \beta_i^2 + e_s \left(v_{1,s} f_{1,i} + 0.32 v_{2,s} f_{2,i} + \cdots + 0.32 v_{16,s} f_{16,i} \right),$$

where the coefficients a_s, b_s, c_s, d_s, e_s, and g_s vary across s. The variations in a_s, b_s, c_s, d_s, and e_s are minimal because the above set of simulations already allows us to study the effects of their variations. Here we focus on allowing g_s to vary for each possible combination of non-zero a_s, b_s, c_s, d_s, and e_s. This allows us to see how the estimators that assume that the conditional first moments of w_{is} are linear in β_i would perform when w_{is} is actually quadratic in β_i. Table G.2 details the values of the parameters. The distributional assumptions for the random variables are the same as above.

Table G.2. $\boldsymbol{a_s}$, $\boldsymbol{b_s}$, $\boldsymbol{c_s}$, $\boldsymbol{d_s}$, $\boldsymbol{g_s}$ and $\boldsymbol{e_s}$ in Simulations

s	a_s	b_s	c_s	d_s	g_s	e_s	$Cov(x_i, w_{is})$	$Cov(w_{is}, \beta_i)$	$E[(\beta_i - E(\beta_i))(w_{is} - E(w_{is}))(x_i - E(x_i))]$
1	50	0	0	0	1	40	0	+	0
2	50	0	0	0	5	40	0	+	0
3	50	10	0	0	1	40	+	+	0
4	50	10	0	0	5	40	+	+	0
5	50	0	10	0	1	40	0	+	0
6	50	0	10	0	5	40	0	+	0
7	50	10	10	0	1	40	+	+	0
8	50	10	10	0	5	40	+	+	0
9	50	0	0	15	1	40	+	+	+
10	50	0	0	15	5	40	+	+	+
11	50	10	0	15	1	40	+	+	+
12	50	10	0	15	5	40	+	+	+

s	a_s	b_s	c_s	d_s	g_s	e_s	$Cov(x_i, w_{is})$	$Cov(w_{is}, \beta_i)$	$E[(\beta_i - E(\beta_i))(w_{is} - E(w_{is}))(x_i - E(x_i))]$
13	50	0	10	15	1	40	+	+	+
14	50	0	10	15	5	40	+	+	+
15	50	10	10	15	1	40	+	+	+
16	50	10	10	15	5	40	+	+	+

G.2. Simulation Results

This subsection reports estimates of the hyperparameters and three SWB moments of interest (to conserve space, we omit tables for $Cov(w_{is}, w_{is'})$) from a variety of simulations. Section G.2.1 reports the baseline simulation with 18 CQs and 3000 respondents. Section G.2.2 reports three sets of robustness simulations, corresponding to data generating processes with a quadratic-translation function, Student-*t* errors, and w_{is} quadratic in β_i. Section G.2.3 reports 15 sets of simulations with $C \in \{2, 3, 9, 27, 81\}$ and $I \in \{100, 1000, 10000\}$.

G.2.1. Baseline Simulation

Our baseline simulation suggests that the six hyperparameters can be estimated very accurately with 3000 respondents and 18 CQs (Table G.3).

Beginning with our estimators for $E(w_{is})$, Table G.4 shows that the naïve estimator ignoring scale use and our MOM estimator are biased when $Cov(\beta_i, w_{is})$ is not zero ($s = 5,6, \ldots, 12$), while both our semi-parametric estimator and comprehensive MLE estimator are unbiased.

Turning to our estimators for $Cov(\beta_i, w_{is})$, Table G.5 shows that when the co-skewness between w_{is}, x_i, and β_i is non-zero ($s = 8,9, \ldots, 12$), the naïve estimator, our MOM estimator and our comprehensive MLE estimator are biased. Only the semi-parametric estimator is able to remove the co-skewness bias.

Finally, we examine our estimators for $Var(w_{is})$. Table G.6 shows that the naïve estimator is heavily biased upward. Estimates from our MOM estimator are closer to the truth because the MOM estimator can remove the biases due to response errors and partially the bias due to scale use. In comparison, our semi-parametric estimator performs better when w_{is} and β_i are not independent, and the advantage is most apparent when the co-skewness between w_{is}, β_i, and x_i is non-zero ($s = 8,9, \ldots, 12$). The simulation's data-generating process does not meet the assumptions of our comprehensive MLE estimator, particularly regarding the normality of w_{is}.

Consequently, its performance is suboptimal, as anticipated. Also as anticipated, when the comprehensive MLE estimator's assumptions are satisfied, it outperforms the other methods, being both unbiased and exhibiting the smallest standard errors, as shown in Table G.7.

Table G.3. Truth and Estimates of Hyperparameters in Baseline Simulation

Parameter	Truth	Estimate
σ_α	7.81	7.85 (0.150)
γ	59.9	59.9 (0.832)
σ_β	0.3	0.300 (0.00722)
σ_ϵ	6.5	6.58 (0.436)
$\mu_{\ln\sigma_\eta}$	2.73	2.73 (0.0152)
$\sigma_{\ln\sigma_\eta}$	0.283	0.283 (0.00778)

Note: Standard errors in parentheses.

Table G.4. Truth and Estimates of $E(w_{is})$ in Baseline Simulation

s	Truth	Ignore scale use/MOM	Semi-parametric	Comprehensive MLE
1	50	50.03 (0.49)	49.98 (0.54)	49.99 (0.49)
2	50	50.06 (0.62)	50.02 (0.66)	50.04 (0.60)
3	60	59.99 (0.52)	60.01 (0.57)	59.98 (0.52)
4	55	54.95 (0.42)	55.00 (0.47)	55.00 (0.44)
5	60	60.91 (0.51)	60.11 (0.48)	60.03 (0.46)
6	70	71.79 (0.49)	70.17 (0.55)	70.09 (0.50)
7	70	70.85 (0.42)	70.09 (0.52)	70.03 (0.44)
8	57.5	58.16 (0.40)	57.56 (0.44)	57.53 (0.41)
9	65	66.27 (0.53)	65.04 (0.55)	64.95 (0.52)
10	67.5	68.12 (0.53)	67.54 (0.56)	67.51 (0.54)
11	67.5	69.03 (0.48)	67.62 (0.49)	67.54 (0.47)
12	77.5	79.04 (0.59)	77.70 (0.60)	77.60 (0.60)

Note: Standard errors in parentheses.

Table G.5. Truth and Estimates of $Cov(x_i, w_{is})$ in Baseline Simulation

s	Truth	Ignore scale use	MOM	Semi-parametric	Comprehensive MLE
1	0	0.02 (0.21)	0.01 (0.21)	0.02 (0.24)	-0.03 (0.20)

s	Truth	Ignore scale use	MOM	Semi-parametric	Comprehensive MLE
2	0	0.03 (0.32)	0.02 (0.32)	0.03 (0.33)	-0.01 (0.31)
3	5	5.01 (0.22)	5.00 (0.22)	5.03 (0.25)	4.98 (0.21)
4	2.5	2.53 (0.23)	2.51 (0.22)	2.52 (0.23)	2.48 (0.21)
5	0	0.06 (0.23)	0.04 (0.23)	0.03 (0.24)	-0.02 (0.22)
6	0	0.04 (0.23)	0.03 (0.22)	0.03 (0.25)	-0.06 (0.21)
7	5	5.00 (0.20)	4.99 (0.20)	5.02 (0.22)	4.95 (0.20)
8	3.75	4.11 (0.23)	4.09 (0.23)	3.80 (0.23)	4.03 (0.23)
9	7.5	8.21 (0.26)	8.20 (0.25)	7.60 (0.29)	8.12 (0.26)
10	8.75	9.13 (0.24)	9.11 (0.23)	8.83 (0.24)	9.06 (0.22)
11	3.75	4.08 (0.25)	4.07 (0.25)	3.74 (0.27)	3.99 (0.23)
12	8.75	9.08 (0.30)	9.07 (0.29)	8.78 (0.31)	9.00 (0.28)

Note: Standard errors in parentheses.

Table G.6. Truth and Estimates of $Var(w_{is})$ in Baseline Simulation

s	Truth	Ignore scale use	MOM	Semi-parametric	Comprehensive MLE
1	212.64	624.13 (19.38)	212.43 (14.92)	220.35 (22.92)	254.45 (14.88)
2	850.56	1323.81 (40.48)	854.43 (38.18)	872.57 (54.02)	896.83 (35.55)
3	312.64	723.32 (20.19)	311.96 (18.11)	328.66 (33.06)	253.59 (13.88)
4	237.64	646.46 (17.59)	239.57 (14.93)	250.88 (25.47)	255.46 (14.70)
5	221.64	628.11 (19.96)	224.08 (16.66)	228.67 (32.69)	264.33 (14.41)
6	248.64	703.34 (21.56)	250.96 (18.45)	268.03 (51.02)	286.47 (16.35)
7	321.64	763.95 (21.63)	322.97 (21.08)	337.38 (55.09)	261.67 (16.22)
8	279.02	709.72 (20.50)	302.18 (18.43)	294.04 (33.52)	268.71 (16.57)
9	478.14	1011.15 (29.76)	560.00 (27.97)	507.75 (66.99)	297.17 (18.49)
10	529.02	1027.85 (28.98)	574.17 (27.55)	557.95 (73.94)	265.07 (16.76)
11	301.52	767.87 (23.39)	325.26 (20.43)	322.95 (56.36)	286.47 (15.72)
12	551.52	1116.30 (36.30)	595.78 (32.54)	584.83 (95.14)	284.27 (16.35)

Note: Standard errors in parentheses.

Table G.7. Truth and Estimates of $Var(w_{is})$ in Simulation Under Normal w_{is}

s	Truth	Ignore scale use	MOM	Semi-parametric	Comprehensive MLE
1	273.81	672.09 (19.36)	275.99 (15.89)	282.96 (35.20)	268.76 (13.14)
2	263.96	666.45 (20.18)	262.59 (18.39)	269.54 (30.74)	256.85 (14.49)
3	222.93	606.94 (17.83)	222.25 (17.01)	231.55 (34.41)	218.64 (12.51)

s	Truth	Ignore scale use	MOM	Semi-parametric	Comprehensive MLE
4	370.46	801.53 (25.02)	370.75 (20.47)	377.43 (36.03)	364.50 (18.97)
5	160.61	536.48 (13.75)	159.75 (13.38)	166.06 (30.57)	159.86 (9.70)
6	121.62	500.27 (12.85)	120.65 (13.14)	127.44 (31.06)	121.57 (8.37)
7	144.17	518.09 (16.74)	142.10 (15.49)	151.62 (33.00)	142.44 (10.26)
8	106.52	485.79 (13.87)	105.85 (12.23)	114.08 (27.38)	104.84 (9.48)
9	239.49	633.36 (16.76)	238.68 (16.61)	245.73 (35.81)	234.74 (12.26)
10	253.74	643.37 (20.96)	252.37 (18.16)	261.14 (32.46)	246.36 (14.49)
11	226.82	630.80 (19.74)	226.96 (16.40)	234.08 (33.97)	222.08 (12.91)
12	250.06	643.59 (16.34)	251.08 (15.48)	265.26 (34.72)	244.30 (14.08)

Notes: w_{is}'s distribution is joint normal with moments calibrated to the estimated moments for the first 12 aspects of Table A.1. Standard errors in parentheses.

G.2.2. Robustness Simulations

We ran three sets of simulations to investigate the robustness of our estimators to misspecifications.

In the first set of simulations, we specify a quadratic translation function for the data-generating process that is similar to the function we estimate in Appendix B:

$$r_{iq} - \gamma = \alpha_i + \beta_i(w_{iq} - \gamma) + \delta_i(w_{iq} - \gamma)^2 + \beta_i \epsilon_{iq} + \eta_{iq},$$

where the distribution of δ_i is calibrated based on our data (see Web Appendix B.2). Our estimators assume the linear translation function of equation (4). Results in Tables G.8-11 suggest that, not surprisingly, the estimates tend to be biased, but the magnitudes of the biases are mild overall.

The second set of robustness simulations we ran changes the response errors' distributions from normal to Student's *t* (with the response-error variances held constant) in the data-generating process, while our estimators still assume normality for the response errors. Tables G.8-11 show that the estimates from our estimators barely change.

The third set of simulations we conducted specifies the data-generating process's w_{is} as being quadratic in β_i (as detailed in Section G.1.1 above), while our semi-parametric method by default continues approximating the (conditional) first moments of w_{is} as being linear in β_i (i.e., we implement the semi-parametric method setting $K = 1$). Tables G.12-14 show that our semi-parametric estimator performs well in these simulations despite the misspecification.

Table G.8. Truth and Estimates of Hyperparameters in Robustness Simulations

Parameter	Truth	Quadratic translation function	Student-*t* errors	w_{is} quadratic in β_i
σ_α	7.81	7.82 (0.129)	7.83 (0.149)	7.84 (0.139)
γ	59.9	60.9 (0.714)	59.8 (0.767)	59.9 (0.663)
σ_β	0.3	0.326 (0.00760)	0.299 (0.00739)	0.300 (0.00658)
σ_ϵ	6.5	5.96 (0.428)	6.38 (0.430)	6.62 (0.431)
$\mu_{\ln \sigma_\eta}$	2.73	2.75 (0.0131)	2.70 (0.0170)	2.73 (0.0155)
$\sigma_{\ln \sigma_\eta}$	0.283	0.274 (0.00785)	0.339 (0.0110)	0.284 (0.00874)

Note: Standard errors in parentheses.

Table G.9. Truth and Estimates of $E(w_{is})$ in Robustness Simulations

		Quadratic translation function in DGP		Student-*t* errors in DGP	
s	Truth	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE
1	50	50.01 (0.44)	50.06 (0.43)	50.02 (0.48)	50.00 (0.44)
2	50	50.56 (0.77)	50.50 (0.70)	49.98 (0.60)	49.97 (0.60)
3	60	60.14 (0.60)	60.02 (0.58)	60.05 (0.49)	60.05 (0.44)
4	55	55.07 (0.54)	55.05 (0.52)	54.99 (0.46)	55.00 (0.45)
5	60	60.19 (0.55)	60.01 (0.53)	60.10 (0.49)	60.08 (0.46)
6	70	70.63 (0.55)	70.13 (0.51)	70.26 (0.52)	70.13 (0.47)
7	70	70.67 (0.58)	70.17 (0.54)	70.15 (0.61)	70.10 (0.58)
8	57.5	57.74 (0.49)	57.55 (0.44)	57.52 (0.57)	57.51 (0.54)
9	65	65.71 (0.71)	65.06 (0.67)	65.15 (0.62)	65.05 (0.58)
10	67.5	68.32 (0.63)	67.74 (0.58)	67.66 (0.57)	67.63 (0.53)
11	67.5	68.19 (0.67)	67.62 (0.59)	67.62 (0.52)	67.53 (0.48)
12	77.5	78.82 (0.72)	77.73 (0.67)	77.66 (0.61)	77.56 (0.55)

Note: Standard errors in parentheses.

Table G.10. Truth and Estimates of $Cov(x_i, w_{is})$ in Robustness Simulations

		Quadratic translation function in DGP			Student-*t* errors in DGP		
s	Truth	MOM	Semi-parametric	Comprehensive MLE	MOM	Semi-parametric	Comprehensive MLE
1	0	0.03 (0.22)	0.05 (0.26)	-0.01 (0.21)	-0.03 (0.25)	-0.04 (0.29)	-0.06 (0.24)
2	0	0.00 (0.36)	0.02 (0.39)	-0.05 (0.35)	-0.02 (0.31)	-0.02 (0.34)	-0.06 (0.30)
3	5	4.99 (0.25)	5.10 (0.26)	4.94 (0.24)	5.00 (0.21)	5.04 (0.23)	4.97 (0.21)
4	2.5	2.50 (0.22)	2.54 (0.25)	2.45 (0.20)	2.52 (0.22)	2.52 (0.23)	2.47 (0.20)
5	0	-0.02 (0.23)	-0.01 (0.25)	-0.09 (0.22)	0.02 (0.23)	0.02 (0.26)	-0.03 (0.23)
6	0	0.03 (0.24)	0.05 (0.28)	-0.08 (0.21)	-0.02 (0.25)	-0.03 (0.27)	-0.10 (0.23)
7	5	5.02 (0.25)	5.22 (0.27)	4.90 (0.22)	4.99 (0.24)	5.03 (0.28)	4.92 (0.22)
8	3.75	4.08 (0.21)	3.88 (0.26)	4.01 (0.21)	4.07 (0.21)	3.75 (0.26)	4.00 (0.20)
9	7.5	8.21 (0.24)	7.89 (0.26)	8.08 (0.23)	8.15 (0.24)	7.54 (0.27)	8.03 (0.23)
10	8.75	9.07 (0.24)	9.07 (0.25)	8.95 (0.22)	9.10 (0.23)	8.82 (0.26)	9.03 (0.22)
11	3.75	4.10 (0.24)	3.95 (0.26)	3.97 (0.23)	4.09 (0.24)	3.78 (0.28)	3.97 (0.22)
12	8.75	9.14 (0.25)	9.34 (0.27)	8.91 (0.25)	9.09 (0.28)	8.82 (0.32)	8.99 (0.28)

Note: Standard errors in parentheses.

Table G.11. Truth and Estimates of $Var(w_{is})$ in Robustness Simulations

		Quadratic translation function in DGP		Student-*t* errors in DGP	
s	Truth	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE
1	212.64	220.00 (31.46)	209.29 (15.49)	218.38 (29.59)	261.84 (16.94)
2	850.56	992.54 (75.87)	905.08 (36.17)	872.09 (50.57)	904.07 (30.04)
3	312.64	350.97 (36.28)	206.42 (14.00)	323.55 (43.13)	261.20 (17.66)
4	237.64	251.39 (30.76)	205.06 (12.93)	249.76 (40.18)	262.21 (16.03)
5	221.64	247.81 (35.84)	212.71 (13.74)	227.61 (40.35)	265.71 (14.71)
6	248.64	307.94 (45.11)	242.75 (15.81)	259.06 (50.40)	296.25 (18.22)
7	321.64	411.74 (47.98)	219.25 (14.66)	340.02 (48.22)	267.18 (16.11)
8	279.02	306.35 (36.92)	217.24 (16.07)	289.07 (40.17)	272.83 (14.33)
9	478.14	565.98 (61.46)	257.79 (15.21)	496.09 (65.19)	305.51 (16.71)
10	529.02	657.28 (62.73)	234.62 (16.80)	545.44 (54.69)	272.86 (16.52)
11	301.52	352.67 (57.62)	234.33 (17.80)	322.51 (52.48)	291.42 (16.82)
12	551.52	810.59 (86.68)	290.41 (19.55)	575.17 (81.85)	293.00 (20.23)

Note: Standard errors in parentheses.

Table G.12. Truth and Estimates of $E(w_{is})$ in Robustness Simulation with w_{is} Quadratic in β_i in DGP

s	Truth	Ignore scale use/MOM	Semi-parametric	Comprehensive MLE
1	51.09	51.34 (0.46)	51.12 (0.51)	51.15 (0.45)
2	55.45	56.37 (0.53)	55.34 (0.57)	55.41 (0.54)
3	56.09	56.30 (0.50)	56.09 (0.55)	56.13 (0.51)
4	60.45	61.32 (0.44)	60.34 (0.51)	60.39 (0.46)
5	61.09	62.15 (0.51)	61.12 (0.58)	61.14 (0.53)
6	65.45	67.26 (0.47)	65.44 (0.49)	65.46 (0.48)
7	66.09	67.13 (0.46)	66.11 (0.52)	66.12 (0.46)
8	70.45	72.29 (0.46)	70.46 (0.51)	70.49 (0.47)
9	58.59	59.44 (0.49)	58.63 (0.56)	58.60 (0.52)
10	62.95	64.49 (0.55)	62.86 (0.59)	62.90 (0.56)
11	63.59	64.43 (0.57)	63.66 (0.62)	63.61 (0.59)
12	67.95	69.47 (0.56)	67.88 (0.61)	67.88 (0.54)
13	68.59	70.37 (0.61)	68.71 (0.65)	68.69 (0.61)
14	72.95	75.40 (0.54)	72.95 (0.62)	72.92 (0.56)
15	73.59	75.35 (0.58)	73.72 (0.60)	73.70 (0.55)
16	77.95	80.38 (0.58)	78.00 (0.59)	77.99 (0.57)

Note: Standard errors in parentheses.

Table G.13. Truth and Estimates of $Cov(x_i, w_{is})$ in Robustness Simulation with w_{is} Quadratic in β_i in DGP

s	Truth	Ignore scale use	MOM	Semi-parametric	Comprehensive MLE
1	0	0.04 (0.26)	0.03 (0.24)	0.03 (0.26)	-0.03 (0.24)
2	0	0.03 (0.23)	0.02 (0.22)	0.02 (0.23)	-0.05 (0.22)
3	2.5	2.53 (0.24)	2.52 (0.22)	2.53 (0.26)	2.46 (0.22)
4	2.5	2.53 (0.27)	2.52 (0.25)	2.53 (0.27)	2.46 (0.26)
5	0	0.00 (0.25)	-0.01 (0.25)	0.00 (0.25)	-0.08 (0.23)
6	0	0.05 (0.23)	0.04 (0.22)	0.05 (0.26)	-0.05 (0.22)
7	2.5	2.50 (0.21)	2.48 (0.19)	2.50 (0.21)	2.41 (0.20)
8	2.5	2.50 (0.20)	2.49 (0.21)	2.52 (0.23)	2.39 (0.20)
9	3.75	4.15 (0.23)	4.14 (0.22)	3.83 (0.26)	4.05 (0.22)
10	3.75	4.11 (0.23)	4.10 (0.23)	3.81 (0.25)	4.00 (0.23)

s	Truth	Ignore scale use	MOM	Semi-parametric	Comprehensive MLE
11	6.25	6.62 (0.26)	6.60 (0.26)	6.32 (0.30)	6.52 (0.26)
12	6.25	6.56 (0.28)	6.55 (0.27)	6.26 (0.29)	6.45 (0.27)
13	3.75	4.11 (0.23)	4.10 (0.22)	3.81 (0.24)	3.99 (0.21)
14	3.75	4.09 (0.26)	4.07 (0.25)	3.80 (0.30)	3.93 (0.25)
15	6.25	6.59 (0.28)	6.58 (0.27)	6.29 (0.30)	6.48 (0.27)
16	6.25	6.57 (0.25)	6.55 (0.24)	6.29 (0.28)	6.43 (0.23)

Note: Standard errors in parentheses.

Table G.14. Truth and Estimates of $Var(w_{is})$ in Robustness Simulation with w_{is} Quadratic in β_i in DGP

s	Truth	Ignore scale use	MOM	Semi-parametric	Comprehensive MLE
1	253.98	664.61 (19.36)	252.91 (16.47)	259.13 (26.07)	246.74 (15.33)
2	263	667.62 (20.42)	267.46 (18.48)	269.65 (30.87)	259.48 (16.47)
3	278.98	689.32 (18.28)	278.95 (14.06)	285.09 (25.29)	247.55 (13.46)
4	288	702.63 (21.90)	290.90 (19.73)	297.80 (35.03)	257.04 (16.88)
5	266.58	677.72 (22.59)	266.37 (19.84)	270.43 (33.20)	258.12 (17.06)
6	290	732.64 (21.66)	297.75 (19.38)	305.87 (37.24)	283.79 (17.25)
7	291.58	720.17 (18.03)	292.99 (16.51)	303.42 (37.44)	259.83 (13.77)
8	315.01	785.92 (23.01)	323.70 (18.29)	333.90 (46.14)	283.80 (15.57)
9	323.05	755.55 (21.22)	343.21 (18.79)	330.31 (33.86)	260.01 (15.04)
10	342.88	798.38 (21.61)	370.99 (18.24)	362.03 (43.18)	281.90 (14.84)
11	423.05	889.27 (27.06)	455.58 (25.77)	429.37 (38.38)	261.93 (16.55)
12	442.88	940.17 (33.32)	479.97 (26.02)	447.74 (55.64)	283.20 (15.89)
13	349.15	828.70 (24.75)	372.64 (22.73)	360.50 (57.03)	285.46 (16.53)
14	383.38	922.70 (28.03)	418.27 (22.43)	410.04 (68.50)	322.56 (16.34)
15	449.15	980.61 (29.91)	485.59 (26.82)	457.38 (65.68)	287.01 (17.70)
16	483.38	1077.68 (31.71)	525.10 (24.74)	503.90 (82.94)	318.27 (16.81)

Note: Standard errors in parentheses.

G.2.3. Simulations Under Various C and I

To study the convergence speed of our estimators, we ran 15 simulations for all combinations of $C \in \{2, 3, 9, 27, 81\}$ and $I \in \{100, 1000, 10000\}$. Theoretically, we expect square-root convergence in I as long as C is large enough, but holding I fixed at any finite value, we expect a positive standard error even as C approaches infinity (because sampling variance remains positive for any finite I). Over the finite range of C we examine, this will create the appearance of slower-than-square-root convergence in C. Our simulations below serve to confirm and quantify these theoretical expectations.

Starting with our estimators of the hyperparameters, Table G.15 shows that the convergence speeds in I of σ_β and σ_ϵ are more sensitive to C than the other hyperparameters are. The convergence of σ_β and σ_ϵ reaches square-root speed in I when $C \geq 9$ and can be slower with smaller C. The average convergence speed in C across all the hyperparameters is slower than square root even at $I = 10{,}000$.

Turning to our estimators of $E(w_{is})$ and $Cov(x_i, w_{is})$, Tables G.16 and G.17 show that, with 2 CQs, we cannot reject the hypothesis of square-root convergence in I of our semi-parametric and comprehensive MLE estimators. The MOM estimator for $Cov(x_i, w_{is})$ also reaches square-root average convergence rate in I at 2 CQs. The convergence rates in C are very slow across all the three levels of I we investigate.

For estimating $Var(w_{is})$, our comprehensive MLE converges at the square-root rate in I when $C \geq 3$, while our semi-parametric estimator converges in I at slower than square-root speed before around 9 CQs (see Table G.18). Both estimators appear to converge in C at slower than square-root speed.

These analyses suggest that the optimal number of CQs varies with the moment of w_{is} of interest. If the interest is in estimating $E(w_{is})$ and $Cov(x_i, w_{is})$, at least 2 CQs are recommended for the convergence in I to reach square-root speed. For estimating $Var(w_{is})$ using our comprehensive MLE estimator, at least 3 CQs are recommended. The tradeoff between additional respondents (for which covariates would need to be collected) and additional CQs depends on the relative costs and the levels of C and I. However, if CQs are an additional module on an existing survey, after meeting the minimum required number of CQs, we recommend prioritizing getting that minimum number of respondents on the full sample.

Table G.15. Truth and Estimates of Hyperparameters in Simulations with Various I and C

Parameter	Truth	$I = 100$	$I = 1{,}000$	$I = 10{,}000$
$C = 2$				
σ_α	7.8	6.1 (3.6)	7.6 (1.2)	7.6 (0.6)
γ	59.9	-9.0 (37.4)	45.4 (28.7)	61.0 (3.9)

Parameter	Truth	$I = 100$	$I = 1{,}000$	$I = 10{,}000$
σ_β	0.3	0.1 (0.1)	0.1 (0.1)	0.3 (0.0)
σ_ϵ	6.5	17.4 (1.8)	12.9 (4.7)	6.9 (1.6)
$\mu_{\ln\sigma_\eta}$	2.7	0.8 (0.7)	2.6 (0.4)	2.7 (0.0)
$\sigma_{\ln\sigma_\eta}$	0.3	0.5 (0.2)	0.3 (0.1)	0.3 (0.0)
$C = 3$				
σ_α	7.8	7.2 (2.9)	7.5 (0.7)	7.7 (0.2)
γ	59.9	0.8 (40.2)	62.8 (21.3)	60.4 (2.3)
σ_β	0.3	0.1 (0.1)	0.2 (0.1)	0.3 (0.0)
σ_ϵ	6.5	16.9 (2.0)	9.8 (4.8)	6.8 (1.2)
$\mu_{\ln\sigma_\eta}$	2.7	1.2 (1.0)	2.7 (0.2)	2.7 (0.0)
$\sigma_{\ln\sigma_\eta}$	0.3	0.4 (0.2)	0.3 (0.1)	0.3 (0.0)
$C = 9$				
σ_α	7.8	8.4 (1.7)	7.9 (0.3)	7.8 (0.1)
γ	59.9	41.4 (26.6)	59.9 (2.2)	59.9 (0.7)
σ_β	0.3	0.2 (0.1)	0.3 (0.0)	0.3 (0.0)
σ_ϵ	6.5	13.2 (5.0)	6.6 (1.4)	6.5 (0.5)
$\mu_{\ln\sigma_\eta}$	2.7	2.2 (0.9)	2.7 (0.1)	2.7 (0.0)
$\sigma_{\ln\sigma_\eta}$	0.3	0.4 (0.2)	0.3 (0.0)	0.3 (0.0)
$C = 27$				
σ_α	7.8	8.6 (0.7)	7.9 (0.2)	7.8 (0.1)
γ	59.9	58.5 (3.1)	59.8 (1.1)	60.0 (0.4)
σ_β	0.3	0.3 (0.0)	0.3 (0.0)	0.3 (0.0)
σ_ϵ	6.5	7.1 (2.0)	6.6 (0.6)	6.5 (0.2)
$\mu_{\ln\sigma_\eta}$	2.7	2.7 (0.1)	2.7 (0.0)	2.7 (0.0)
$\sigma_{\ln\sigma_\eta}$	0.3	0.3 (0.0)	0.3 (0.0)	0.3 (0.0)
$C = 81$				
σ_α	7.8	8.6 (0.6)	7.9 (0.2)	7.8 (0.1)
γ	59.9	58.7 (3.1)	59.8 (1.0)	59.9 (0.4)
σ_β	0.3	0.3 (0.0)	0.3 (0.0)	0.3 (0.0)

Parameter	Truth	$I = 100$	$I = 1{,}000$	$I = 10{,}000$
σ_ϵ	6.5	7.0 (1.5)	6.6 (0.4)	6.5 (0.1)
$\mu_{\ln\sigma_\eta}$	2.7	2.7 (0.1)	2.7 (0.0)	2.7 (0.0)
$\sigma_{\ln\sigma_\eta}$	0.3	0.3 (0.0)	0.3 (0.0)	0.3 (0.0)

Note: Standard errors in parentheses.

Table G.16. Truth and Estimates of $E(w_{is})$ in Simulations with Various I and C

		$I = 100$		$I = 1{,}000$		$I = 10{,}000$	
s	Truth	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE
$C = 2$							
1	50.0	50.1 (2.7)	50.2 (2.3)	50.2 (1.0)	50.2 (0.8)	49.8 (0.3)	50.0 (0.3)
2	50.0	49.8 (4.2)	49.5 (3.7)	50.2 (1.3)	50.0 (1.0)	49.9 (0.4)	50.1 (0.4)
3	60.0	60.6 (2.9)	60.7 (2.7)	59.8 (1.1)	60.0 (0.9)	60.1 (0.3)	60.0 (0.3)
4	55.0	54.9 (2.8)	55.0 (2.6)	55.1 (0.9)	55.1 (0.8)	55.0 (0.3)	55.0 (0.3)
5	60.0	60.3 (2.6)	60.7 (2.5)	60.3 (1.0)	60.7 (0.9)	60.5 (0.3)	60.0 (0.3)
6	70.0	70.5 (3.4)	71.4 (2.8)	70.2 (1.0)	71.3 (1.1)	71.2 (0.6)	70.0 (0.4)
7	70.0	70.6 (3.1)	71.4 (3.0)	69.9 (1.1)	70.7 (1.1)	70.8 (0.5)	70.0 (0.4)
8	57.5	57.8 (3.0)	58.1 (2.7)	57.6 (1.0)	58.1 (0.9)	57.8 (0.3)	57.5 (0.4)
9	65.0	65.5 (3.8)	66.3 (2.8)	65.2 (1.2)	65.9 (1.2)	65.9 (0.5)	65.0 (0.3)
10	67.5	67.4 (3.7)	68.0 (2.9)	67.3 (1.2)	68.0 (1.1)	68.1 (0.5)	67.4 (0.4)
11	67.5	68.3 (3.4)	68.8 (3.0)	67.8 (1.1)	68.7 (1.0)	68.5 (0.5)	67.5 (0.3)
12	77.5	78.4 (4.0)	79.3 (3.7)	77.4 (1.5)	78.6 (1.5)	78.8 (0.7)	77.4 (0.4)
$C = 3$							
1	50.0	50.5 (2.5)	50.3 (2.2)	50.1 (1.0)	50.1 (0.8)	49.9 (0.3)	50.0 (0.3)
2	50.0	50.5 (4.0)	50.3 (3.9)	50.3 (1.3)	50.2 (1.1)	49.9 (0.5)	50.1 (0.4)
3	60.0	60.3 (2.8)	60.2 (2.6)	60.0 (1.0)	60.0 (0.8)	60.0 (0.3)	60.0 (0.3)
4	55.0	55.5 (3.1)	55.4 (2.5)	55.1 (0.9)	55.1 (0.7)	55.0 (0.3)	55.0 (0.3)
5	60.0	60.1 (3.2)	60.6 (2.7)	60.2 (1.0)	60.4 (0.8)	60.4 (0.3)	60.0 (0.3)
6	70.0	70.7 (3.1)	71.5 (2.7)	70.5 (1.3)	71.0 (1.1)	70.8 (0.4)	70.0 (0.3)
7	70.0	69.7 (3.4)	70.3 (3.2)	70.1 (1.4)	70.5 (1.2)	70.6 (0.4)	70.0 (0.3)
8	57.5	57.7 (2.9)	57.9 (2.5)	57.8 (1.1)	57.9 (0.9)	57.7 (0.3)	57.5 (0.3)
9	65.0	65.7 (3.7)	66.3 (3.1)	65.4 (1.3)	65.6 (1.2)	65.6 (0.4)	64.9 (0.3)
10	67.5	67.6 (3.3)	68.1 (3.0)	67.6 (1.3)	67.9 (1.0)	68.0 (0.4)	67.5 (0.3)
11	67.5	68.1 (3.3)	68.7 (2.9)	68.0 (1.3)	68.4 (1.1)	68.2 (0.4)	67.4 (0.3)
12	77.5	77.4 (3.8)	78.3 (3.5)	77.6 (1.6)	78.3 (1.3)	78.4 (0.5)	77.4 (0.4)
$C = 9$							
1	50.0	50.2 (2.9)	49.8 (2.7)	49.9 (0.8)	50.0 (0.8)	50.0 (0.3)	50.0 (0.3)

		$I = 100$		$I = 1{,}000$		$I = 10{,}000$	
s	Truth	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE
2	50.0	50.6 (3.9)	50.0 (3.7)	50.1 (1.4)	50.1 (1.2)	49.9 (0.4)	50.0 (0.4)
3	60.0	61.0 (3.2)	60.7 (2.7)	60.1 (0.9)	60.1 (0.8)	60.0 (0.3)	60.0 (0.3)
4	55.0	55.5 (2.6)	55.2 (2.3)	55.0 (0.9)	55.1 (0.8)	55.0 (0.3)	55.1 (0.3)
5	60.0	60.3 (2.8)	60.8 (2.4)	60.4 (0.9)	60.3 (0.8)	60.1 (0.3)	60.0 (0.3)
6	70.0	70.4 (2.9)	72.1 (2.6)	70.3 (1.0)	70.2 (0.9)	70.4 (0.3)	70.1 (0.3)
7	70.0	70.3 (2.9)	71.4 (2.5)	70.2 (1.0)	70.1 (0.9)	70.3 (0.3)	70.1 (0.3)
8	57.5	57.7 (3.0)	58.1 (2.7)	57.6 (1.1)	57.5 (1.0)	57.6 (0.3)	57.5 (0.3)
9	65.0	66.0 (3.4)	66.7 (3.3)	65.3 (1.2)	65.1 (1.1)	65.3 (0.4)	65.0 (0.3)
10	67.5	68.0 (3.3)	68.3 (2.9)	67.7 (1.2)	67.6 (1.1)	67.7 (0.4)	67.5 (0.4)
11	67.5	68.2 (3.1)	69.6 (3.0)	67.7 (0.9)	67.6 (0.8)	67.9 (0.3)	67.6 (0.3)
12	77.5	78.0 (3.6)	79.5 (3.1)	78.1 (1.2)	77.9 (1.1)	77.9 (0.4)	77.6 (0.4)
$C = 27$							
1	50.0	49.7 (2.7)	49.6 (2.6)	49.9 (0.8)	49.9 (0.8)	50.0 (0.3)	49.9 (0.3)
2	50.0	49.7 (3.6)	49.8 (3.3)	50.2 (1.1)	50.2 (1.1)	49.9 (0.4)	49.9 (0.4)
3	60.0	59.7 (2.9)	59.9 (2.8)	60.0 (0.9)	60.1 (0.8)	60.0 (0.3)	60.0 (0.3)
4	55.0	54.9 (2.6)	54.8 (2.5)	55.0 (0.9)	55.0 (0.9)	55.0 (0.3)	55.0 (0.3)
5	60.0	59.8 (2.8)	60.3 (2.4)	60.0 (0.8)	60.0 (0.8)	60.0 (0.3)	60.0 (0.3)
6	70.0	70.4 (2.8)	71.1 (2.7)	70.1 (0.9)	70.1 (0.9)	70.1 (0.3)	70.0 (0.3)
7	70.0	69.4 (2.9)	70.1 (2.5)	70.2 (0.8)	70.2 (0.8)	70.0 (0.3)	70.0 (0.3)
8	57.5	57.4 (2.8)	57.7 (2.7)	57.5 (0.9)	57.6 (0.8)	57.5 (0.3)	57.5 (0.3)
9	65.0	64.6 (3.2)	65.3 (3.0)	65.1 (1.1)	65.2 (1.0)	65.0 (0.3)	65.0 (0.3)
10	67.5	67.0 (2.8)	67.5 (2.7)	67.7 (1.0)	67.7 (1.0)	67.6 (0.3)	67.6 (0.3)
11	67.5	67.5 (3.1)	68.2 (2.7)	67.5 (0.9)	67.6 (0.9)	67.6 (0.3)	67.5 (0.3)
12	77.5	77.2 (3.2)	78.1 (2.9)	77.6 (1.1)	77.7 (1.1)	77.6 (0.4)	77.5 (0.4)
$C = 81$							
1	50.0	50.1 (2.7)	50.1 (2.5)	49.9 (0.8)	49.9 (0.7)	50.1 (0.3)	50.0 (0.3)
2	50.0	50.4 (4.2)	50.3 (4.0)	50.1 (1.2)	50.1 (1.2)	50.0 (0.4)	50.0 (0.4)
3	60.0	59.7 (2.7)	59.9 (2.5)	59.9 (0.9)	60.0 (0.8)	60.0 (0.3)	60.0 (0.3)
4	55.0	55.0 (2.7)	55.1 (2.6)	55.0 (0.8)	55.0 (0.8)	55.1 (0.3)	55.1 (0.3)
5	60.0	60.4 (2.9)	60.5 (2.8)	60.1 (0.8)	60.1 (0.8)	60.1 (0.3)	60.1 (0.3)
6	70.0	70.0 (2.9)	70.6 (2.7)	70.0 (0.8)	70.1 (0.8)	70.0 (0.3)	70.0 (0.3)
7	70.0	69.9 (2.7)	70.4 (2.6)	70.0 (0.9)	70.0 (0.8)	70.0 (0.3)	70.0 (0.3)
8	57.5	57.6 (2.6)	57.9 (2.3)	57.3 (0.9)	57.4 (0.8)	57.5 (0.3)	57.5 (0.3)
9	65.0	65.1 (3.2)	65.5 (3.0)	65.1 (0.9)	65.1 (0.8)	65.0 (0.3)	65.0 (0.3)
10	67.5	67.3 (3.0)	67.7 (3.0)	67.5 (1.0)	67.6 (1.0)	67.5 (0.3)	67.6 (0.3)
11	67.5	67.6 (3.1)	68.2 (3.0)	67.4 (1.0)	67.5 (0.9)	67.5 (0.3)	67.5 (0.3)
12	77.5	77.6 (3.3)	78.4 (3.4)	77.5 (1.1)	77.7 (0.9)	77.6 (0.4)	77.6 (0.3)

		$I = 100$		$I = 1{,}000$		$I = 10{,}000$	
s	Truth	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE

Note: Standard errors in parentheses.

Table G.17. Truth and Estimates of $Cov(x_i, w_{is})$ in Simulations with Various I and C

		$I = 100$			$I = 1{,}000$			$I = 10{,}000$		
s	Truth	MOM	Semi-parametric	Comprehensive MLE	MOM	Semi-parametric	Comprehensive MLE	MOM	Semi-parametric	Comprehensive MLE
$C = 2$										
1	0.0	0.1 (1.4)	0.2 (1.6)	0.1 (1.3)	0.0 (0.4)	0.0 (0.5)	0.0 (0.4)	0.0 (0.1)	0.0 (0.2)	0.0 (0.1)
2	0.0	-0.1 (1.8)	-0.4 (2.1)	-0.1 (1.8)	0.0 (0.6)	0.0 (0.7)	0.1 (0.5)	0.0 (0.2)	0.0 (0.2)	0.0 (0.2)
3	5.0	5.2 (1.4)	5.0 (1.6)	5.1 (1.1)	5.1 (0.4)	4.8 (0.5)	5.1 (0.4)	5.0 (0.1)	5.1 (0.2)	5.0 (0.1)
4	2.5	2.6 (1.4)	2.5 (1.7)	2.6 (1.1)	2.5 (0.4)	2.4 (0.5)	2.5 (0.4)	2.5 (0.1)	2.6 (0.2)	2.5 (0.1)
5	0.0	0.1 (1.4)	0.1 (1.6)	0.1 (1.2)	-0.1 (0.5)	-0.1 (0.5)	0.0 (0.4)	0.0 (0.1)	0.0 (0.2)	0.0 (0.1)
6	0.0	0.1 (1.7)	0.1 (2.0)	0.1 (1.3)	-0.1 (0.5)	-0.1 (0.5)	0.0 (0.4)	-0.1 (0.1)	0.0 (0.2)	0.0 (0.2)
7	5.0	5.0 (1.5)	4.6 (1.8)	4.8 (1.1)	5.0 (0.5)	4.9 (0.5)	5.1 (0.4)	5.0 (0.1)	5.1 (0.2)	5.0 (0.2)
8	3.8	4.0 (1.3)	3.6 (1.6)	3.9 (1.2)	4.1 (0.4)	3.7 (0.5)	4.0 (0.4)	4.0 (0.1)	4.0 (0.2)	4.1 (0.1)
9	7.5	8.4 (1.7)	7.7 (1.9)	8.2 (1.4)	8.2 (0.5)	7.5 (0.6)	8.0 (0.5)	8.0 (0.1)	8.0 (0.3)	8.2 (0.1)
10	8.8	9.2 (1.6)	8.6 (1.8)	9.0 (1.3)	9.0 (0.5)	8.5 (0.5)	9.0 (0.4)	9.0 (0.1)	9.1 (0.3)	9.1 (0.2)
11	3.8	4.1 (1.8)	3.9 (2.0)	4.0 (1.3)	4.1 (0.5)	3.7 (0.5)	4.0 (0.4)	4.0 (0.1)	4.0 (0.2)	4.1 (0.2)
12	8.8	9.1 (1.9)	8.5 (2.0)	8.9 (1.3)	9.1 (0.6)	8.6 (0.6)	9.1 (0.5)	9.0 (0.2)	9.1 (0.3)	9.1 (0.2)
$C = 3$										
1	0.0	0.0 (1.3)	0.0 (1.4)	0.2 (1.2)	0.0 (0.4)	0.0 (0.5)	0.0 (0.4)	0.0 (0.1)	0.0 (0.2)	0.0 (0.1)
2	0.0	-0.1 (2.0)	-0.3 (2.0)	0.0 (1.9)	0.1 (0.6)	0.0 (0.7)	0.1 (0.6)	0.0 (0.2)	0.0 (0.2)	0.0 (0.2)
3	5.0	5.0 (1.2)	4.8 (1.3)	5.1 (1.2)	5.0 (0.4)	5.0 (0.5)	5.0 (0.4)	5.0 (0.1)	5.1 (0.2)	5.0 (0.1)
4	2.5	2.3 (1.4)	2.1 (1.5)	2.4 (1.2)	2.5 (0.4)	2.5 (0.5)	2.5 (0.4)	2.5 (0.1)	2.5 (0.2)	2.5 (0.1)
5	0.0	-0.2 (1.4)	-0.1 (1.6)	0.0 (1.3)	0.0 (0.3)	0.1 (0.4)	0.0 (0.3)	0.0 (0.1)	0.0 (0.2)	0.0 (0.1)
6	0.0	-0.2 (1.5)	-0.1 (1.6)	0.0 (1.3)	0.0 (0.4)	0.1 (0.6)	0.0 (0.4)	0.0 (0.1)	0.0 (0.2)	0.0 (0.1)
7	5.0	4.9 (1.6)	4.8 (1.8)	5.0 (1.3)	5.0 (0.5)	5.1 (0.6)	5.0 (0.4)	5.0 (0.1)	5.1 (0.2)	5.0 (0.1)
8	3.8	3.9 (1.4)	3.6 (1.7)	4.0 (1.3)	4.2 (0.4)	3.9 (0.5)	4.1 (0.4)	4.1 (0.1)	3.9 (0.2)	4.0 (0.1)
9	7.5	7.7 (1.7)	7.0 (1.7)	7.7 (1.4)	8.2 (0.4)	7.6 (0.6)	8.1 (0.4)	8.2 (0.1)	7.9 (0.2)	8.1 (0.1)
10	8.8	8.7 (1.6)	8.2 (1.6)	8.7 (1.4)	9.2 (0.5)	8.9 (0.7)	9.1 (0.4)	9.1 (0.1)	9.1 (0.2)	9.0 (0.1)
11	3.8	3.8 (1.5)	3.6 (1.8)	3.9 (1.3)	4.1 (0.4)	3.9 (0.5)	4.0 (0.4)	4.1 (0.1)	4.0 (0.2)	4.0 (0.1)
12	8.8	8.8 (2.0)	8.4 (1.9)	8.9 (1.6)	9.2 (0.5)	8.8 (0.7)	9.1 (0.5)	9.1 (0.1)	9.0 (0.2)	9.0 (0.1)
$C = 9$										
1	0.0	0.0 (1.2)	0.0 (1.3)	0.1 (1.2)	0.1 (0.4)	0.0 (0.4)	0.0 (0.4)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)
2	0.0	-0.2 (1.8)	-0.1 (1.9)	-0.1 (1.9)	0.0 (0.7)	0.0 (0.8)	0.0 (0.7)	0.0 (0.2)	0.0 (0.2)	0.0 (0.2)

		$I = 100$			$I = 1{,}000$			$I = 10{,}000$		
s	Truth	MOM	Semi-parametric	Comprehensive MLE	MOM	Semi-parametric	Comprehensive MLE	MOM	Semi-parametric	Comprehensive MLE
3	5.0	5.1 (1.4)	4.9 (1.4)	5.1 (1.4)	5.0 (0.4)	5.0 (0.4)	4.9 (0.4)	5.0 (0.1)	5.1 (0.1)	5.0 (0.1)
4	2.5	2.3 (1.3)	2.3 (1.3)	2.4 (1.3)	2.5 (0.4)	2.6 (0.5)	2.5 (0.4)	2.5 (0.1)	2.5 (0.1)	2.5 (0.1)
5	0.0	0.0 (1.2)	-0.1 (1.4)	-0.1 (1.1)	0.0 (0.4)	0.0 (0.4)	0.0 (0.3)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)
6	0.0	0.1 (1.3)	0.1 (1.4)	0.0 (1.3)	0.0 (0.4)	0.0 (0.5)	-0.1 (0.4)	0.0 (0.1)	0.0 (0.2)	0.0 (0.1)
7	5.0	5.1 (1.1)	5.0 (1.3)	5.2 (1.2)	5.0 (0.4)	5.1 (0.4)	5.0 (0.4)	5.0 (0.1)	5.0 (0.2)	4.9 (0.1)
8	3.8	4.1 (1.2)	3.7 (1.3)	4.1 (1.2)	4.1 (0.4)	3.9 (0.4)	4.0 (0.4)	4.1 (0.1)	3.8 (0.1)	4.0 (0.1)
9	7.5	7.9 (1.3)	7.1 (1.4)	7.9 (1.2)	8.2 (0.4)	7.6 (0.5)	8.1 (0.4)	8.2 (0.1)	7.7 (0.1)	8.1 (0.1)
10	8.8	9.2 (1.3)	8.6 (1.6)	9.3 (1.3)	9.1 (0.4)	8.8 (0.5)	9.0 (0.4)	9.1 (0.1)	8.9 (0.2)	9.0 (0.1)
11	3.8	4.2 (1.3)	3.8 (1.6)	4.1 (1.3)	4.1 (0.5)	3.8 (0.5)	4.0 (0.4)	4.1 (0.1)	3.9 (0.1)	4.0 (0.1)
12	8.8	9.1 (1.4)	8.6 (1.7)	9.1 (1.4)	9.1 (0.5)	8.9 (0.5)	9.1 (0.4)	9.1 (0.1)	8.9 (0.2)	9.0 (0.1)
$C = 27$										
1	0.0	-0.1 (1.2)	-0.1 (1.4)	-0.1 (1.2)	0.0 (0.4)	0.0 (0.4)	-0.1 (0.4)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)
2	0.0	-0.2 (1.9)	-0.2 (1.9)	-0.3 (1.9)	-0.1 (0.6)	-0.1 (0.6)	-0.2 (0.6)	0.0 (0.2)	0.0 (0.2)	-0.1 (0.2)
3	5.0	5.0 (1.2)	5.0 (1.3)	5.1 (1.2)	5.0 (0.4)	5.0 (0.4)	4.9 (0.3)	5.0 (0.1)	5.0 (0.1)	4.9 (0.1)
4	2.5	2.8 (1.1)	2.7 (1.2)	2.8 (1.2)	2.4 (0.4)	2.4 (0.4)	2.4 (0.4)	2.5 (0.1)	2.5 (0.1)	2.4 (0.1)
5	0.0	0.0 (1.1)	-0.1 (1.1)	-0.1 (1.1)	0.0 (0.4)	0.0 (0.4)	-0.1 (0.4)	0.0 (0.1)	0.0 (0.1)	-0.1 (0.1)
6	0.0	0.0 (1.2)	0.0 (1.2)	0.0 (1.2)	-0.1 (0.4)	-0.1 (0.4)	-0.1 (0.4)	0.0 (0.1)	0.0 (0.1)	-0.1 (0.1)
7	5.0	4.9 (1.2)	4.8 (1.4)	4.9 (1.2)	5.0 (0.4)	5.0 (0.5)	5.0 (0.4)	5.0 (0.1)	5.0 (0.1)	4.9 (0.1)
8	3.8	4.3 (1.3)	3.9 (1.4)	4.3 (1.3)	4.0 (0.4)	3.7 (0.4)	4.0 (0.4)	4.1 (0.1)	3.8 (0.1)	4.0 (0.1)
9	7.5	8.3 (1.5)	7.3 (1.5)	8.3 (1.4)	8.1 (0.4)	7.5 (0.4)	8.1 (0.4)	8.2 (0.1)	7.5 (0.1)	8.1 (0.1)
10	8.8	9.3 (1.2)	8.8 (1.4)	9.3 (1.2)	9.1 (0.4)	8.8 (0.5)	9.1 (0.4)	9.1 (0.1)	8.8 (0.1)	9.0 (0.1)
11	3.8	4.3 (1.2)	3.8 (1.4)	4.3 (1.2)	4.0 (0.4)	3.7 (0.4)	4.0 (0.4)	4.1 (0.1)	3.8 (0.1)	4.0 (0.1)
12	8.8	9.2 (1.4)	8.8 (1.7)	9.3 (1.4)	9.1 (0.4)	8.8 (0.4)	9.1 (0.4)	9.1 (0.1)	8.8 (0.1)	9.0 (0.1)
$C = 81$										
1	0.0	-0.2 (1.1)	-0.1 (1.2)	-0.2 (1.2)	0.0 (0.4)	0.0 (0.4)	-0.1 (0.4)	0.0 (0.1)	0.0 (0.1)	0.0 (0.1)
2	0.0	-0.1 (1.5)	0.0 (1.6)	-0.1 (1.5)	0.0 (0.6)	0.0 (0.6)	0.0 (0.6)	0.0 (0.2)	0.0 (0.2)	-0.1 (0.2)
3	5.0	5.0 (1.3)	5.0 (1.4)	5.0 (1.2)	4.9 (0.4)	5.0 (0.4)	4.9 (0.4)	5.0 (0.1)	5.0 (0.1)	4.9 (0.1)
4	2.5	2.5 (1.1)	2.5 (1.2)	2.5 (1.1)	2.5 (0.4)	2.5 (0.4)	2.5 (0.3)	2.5 (0.1)	2.5 (0.1)	2.5 (0.1)
5	0.0	-0.1 (1.2)	-0.1 (1.3)	-0.2 (1.2)	-0.1 (0.4)	-0.1 (0.4)	-0.2 (0.4)	0.0 (0.1)	0.0 (0.1)	-0.1 (0.1)
6	0.0	-0.2 (1.2)	-0.2 (1.3)	-0.3 (1.2)	0.0 (0.4)	0.0 (0.5)	-0.1 (0.4)	0.0 (0.1)	0.0 (0.1)	-0.1 (0.1)
7	5.0	5.0 (1.2)	4.9 (1.4)	5.0 (1.1)	5.0 (0.4)	5.0 (0.4)	5.0 (0.4)	5.0 (0.1)	5.0 (0.1)	4.9 (0.1)
8	3.8	4.0 (1.3)	3.6 (1.3)	4.0 (1.3)	4.1 (0.4)	3.8 (0.4)	4.1 (0.4)	4.1 (0.1)	3.8 (0.1)	4.0 (0.1)
9	7.5	7.9 (1.3)	7.2 (1.5)	7.9 (1.3)	8.1 (0.4)	7.4 (0.4)	8.0 (0.4)	8.2 (0.1)	7.5 (0.1)	8.1 (0.1)
10	8.8	9.1 (1.3)	8.6 (1.4)	9.1 (1.3)	9.1 (0.4)	8.8 (0.5)	9.0 (0.4)	9.1 (0.1)	8.8 (0.2)	9.0 (0.1)
11	3.8	4.1 (1.2)	3.7 (1.4)	4.0 (1.2)	4.1 (0.4)	3.8 (0.4)	4.0 (0.3)	4.1 (0.1)	3.8 (0.1)	4.0 (0.1)
12	8.8	9.0 (1.2)	8.5 (1.4)	9.0 (1.2)	9.1 (0.4)	8.7 (0.5)	9.0 (0.4)	9.1 (0.1)	8.8 (0.1)	9.0 (0.1)

s	Truth	$I = 100$ MOM	$I = 100$ Semi-parametric	$I = 100$ Comprehensive MLE	$I = 1{,}000$ MOM	$I = 1{,}000$ Semi-parametric	$I = 1{,}000$ Comprehensive MLE	$I = 10{,}000$ MOM	$I = 10{,}000$ Semi-parametric	$I = 10{,}000$ Comprehensive MLE

Note: Standard errors in parentheses.

Table G.18. Truth and Estimates of $Var(w_{is})$ in Simulations with Various I and C

s	Truth	$I = 100$		$I = 1{,}000$		$I = 10{,}000$	
		Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE
$C = 2$							
1	213	151 (242)	161 (114)	5 (166)	87 (59)	276 (57)	201 (25)
2	851	733 (388)	762 (185)	613 (201)	668 (119)	1004 (125)	847 (33)
3	313	236 (233)	135 (94)	104 (179)	83 (56)	391 (66)	203 (25)
4	238	143 (259)	142 (99)	24 (161)	82 (58)	298 (60)	203 (26)
5	222	156 (258)	146 (104)	12 (185)	87 (64)	287 (59)	212 (25)
6	249	202 (282)	185 (107)	35 (189)	136 (69)	358 (76)	237 (23)
7	322	237 (327)	152 (98)	120 (194)	100 (60)	433 (79)	210 (25)
8	279	204 (263)	153 (102)	69 (173)	94 (64)	354 (62)	216 (25)
9	478	459 (388)	193 (115)	287 (227)	148 (75)	639 (98)	259 (25)
10	529	474 (369)	171 (114)	320 (216)	120 (69)	683 (95)	225 (25)
11	302	231 (321)	177 (117)	85 (193)	129 (67)	417 (78)	235 (24)
12	552	509 (446)	240 (121)	332 (241)	162 (78)	757 (123)	242 (24)
$C = 3$							
1	213	80 (237)	122 (99)	128 (184)	127 (56)	288 (44)	205 (12)
2	851	708 (374)	696 (208)	754 (286)	745 (99)	1009 (96)	847 (22)
3	313	216 (226)	112 (98)	228 (200)	126 (58)	397 (53)	206 (11)
4	238	104 (255)	127 (98)	143 (184)	120 (57)	314 (47)	207 (11)
5	222	112 (218)	129 (111)	136 (208)	131 (55)	297 (47)	214 (13)
6	249	171 (306)	174 (127)	167 (244)	176 (61)	353 (58)	237 (11)
7	322	227 (337)	134 (102)	244 (228)	141 (57)	423 (64)	210 (12)
8	279	175 (257)	128 (97)	205 (188)	140 (59)	361 (51)	219 (11)
9	478	353 (405)	189 (125)	403 (287)	190 (65)	613 (74)	257 (13)
10	529	374 (365)	154 (98)	465 (285)	153 (59)	671 (86)	224 (13)
11	302	199 (309)	160 (120)	224 (240)	175 (62)	409 (67)	238 (11)
12	552	476 (502)	198 (110)	491 (327)	197 (66)	731 (110)	237 (14)
$C = 9$							
1	213	125 (219)	129 (107)	230 (67)	196 (24)	240 (20)	212 (8)
2	851	636 (426)	755 (196)	903 (136)	838 (64)	914 (42)	849 (18)

		$I = 100$		$I = 1{,}000$		$I = 10{,}000$	
s	Truth	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE	Semi-parametric	Comprehensive MLE
3	313	206 (229)	112 (96)	338 (79)	195 (27)	342 (23)	211 (9)
4	238	133 (250)	115 (108)	262 (65)	196 (28)	265 (17)	212 (9)
5	222	95 (240)	118 (90)	244 (58)	199 (26)	244 (18)	220 (8)
6	249	111 (297)	157 (100)	291 (85)	220 (27)	279 (24)	245 (8)
7	322	201 (260)	128 (93)	348 (86)	199 (30)	361 (28)	218 (8)
8	279	182 (303)	126 (96)	308 (75)	207 (29)	309 (21)	225 (9)
9	478	320 (321)	171 (100)	510 (117)	241 (33)	532 (37)	261 (9)
10	529	425 (343)	146 (104)	574 (104)	206 (28)	578 (32)	228 (9)
11	302	198 (311)	154 (105)	326 (98)	223 (29)	337 (28)	243 (10)
12	552	402 (368)	181 (106)	596 (145)	221 (26)	617 (49)	245 (9)
$C = 27$							
1	213	205 (116)	178 (74)	213 (39)	205 (19)	217 (14)	213 (7)
2	851	816 (246)	792 (173)	847 (85)	833 (58)	863 (27)	851 (17)
3	313	310 (147)	187 (71)	308 (45)	200 (23)	315 (17)	212 (8)
4	238	241 (130)	177 (70)	239 (39)	201 (22)	240 (14)	212 (9)
5	222	233 (150)	191 (71)	217 (46)	209 (23)	227 (14)	223 (8)
6	249	261 (193)	211 (70)	255 (61)	235 (24)	251 (24)	248 (8)
7	322	340 (209)	195 (74)	325 (71)	208 (27)	326 (23)	219 (9)
8	279	258 (158)	187 (76)	286 (54)	211 (27)	284 (19)	222 (7)
9	478	518 (321)	235 (83)	473 (101)	241 (24)	484 (31)	254 (9)
10	529	554 (220)	192 (74)	533 (84)	214 (24)	538 (25)	223 (8)
11	302	309 (248)	210 (82)	306 (82)	231 (27)	305 (29)	244 (8)
12	552	563 (312)	200 (70)	571 (118)	237 (27)	561 (40)	243 (9)
$C = 81$							
1	213	210 (120)	184 (62)	210 (38)	206 (22)	213 (12)	207 (34)
2	851	816 (233)	799 (161)	851 (91)	843 (55)	850 (25)	846 (18)
3	313	321 (142)	196 (72)	312 (45)	206 (25)	312 (17)	211 (7)
4	238	228 (123)	193 (64)	237 (38)	203 (22)	238 (13)	211 (7)
5	222	213 (128)	195 (67)	220 (39)	215 (24)	222 (16)	220 (8)
6	249	247 (184)	230 (77)	248 (65)	237 (21)	249 (24)	246 (8)
7	322	353 (216)	191 (72)	320 (68)	212 (26)	325 (24)	220 (8)
8	279	325 (169)	213 (67)	277 (47)	212 (25)	281 (20)	221 (8)
9	478	496 (248)	219 (77)	487 (100)	248 (25)	476 (34)	250 (8)
10	529	540 (211)	190 (72)	521 (81)	210 (20)	528 (29)	221 (8)
11	302	307 (209)	210 (66)	307 (75)	235 (23)	300 (30)	242 (8)
12	552	562 (276)	206 (66)	549 (117)	237 (24)	551 (40)	243 (7)

Note: Standard errors in parentheses.

Appendix G Reference

Trefethen, Lloyd N., and David Bau. "Numerical Linear Algebra." *Society for Industrial and Applied Mathematics*, 1997.

H. Fraction of Scale-Use Heterogeneity Explained by General Scale Use

This appendix details our estimation of the fraction of scale use accounted for by general scale use.

H.1. The Measure and the Estimates

As noted in Section VII.B, we measure the fraction of scale-use variance in CQ dimension *d* explained by general scale use by the R^2 from a regression of the mean MMB calculated from dimension-*d* CQs on the mean MMB calculated from all other CQs, excluding dimension *d*. We measure the overall fraction of scale-use heterogeneity explained by general scale use as the mean of the dimension-specific R^2's.

Formally, for each dimension *d*, we calculate the analytical R^2 for the following regression equation:

$$\hat{r}_{d,i}(h) = \beta_0 + \beta_1 \frac{1}{G-1} \sum_{k=1,k\neq d}^{G} \hat{r}_{k,i}(h) + \varsigma_i,$$

where, as defined in Section V.C, $\hat{r}_{k,i}(h)$ denotes the dimension-k MMB targeting the height h for person i and G denotes the total number of dimensions used. We calculate the analytical R^2 rather than running the empirical analog of this regression because response errors inflate the MMBs' variances. The analytical R^2 corrects for this by subtracting from the raw variances the response-error variances estimated through a dimension-specific version of the CQ-only MLE.

We conduct this analysis using two sets of CQs. The first set uses all 60 available dimensions in the Bottomless survey (see Appendix A.4 for details). The other set uses the subset of 42 (out of the 60) dimensions that correspond to personal and local-public-good SWB dimensions; it excludes visual CQs and CQs for non-local-public-good SWB dimensions (see

Appendix Table A.4 for the correspondence between each dimension and its respective category).

Table H.1 reports our results with the height used for the MMBs varying between the highest and lowest (semi-parametric-method-estimated) means among the four ONS SWBs (worthwhileness with the mean of 69.90, and no anxiety with the mean of 54.04). Across both MMBs, the mean R^2 is more than 55% in the set of CQs with 60 dimensions and more than 75% in the set of CQs with 42 dimensions.

Table H.1. Means of Dimension-Specific R^2's

Dimensions (CQs)	Mean R^2	MMB Mean
42 (316 CQs)	75.70% (1.53%)	54.04
42 (316 CQs)	75.58% (1.53%)	69.90
60 (388 CQs)	60.03% (1.48%)	54.04
60 (388 CQs)	56.32% (1.52%)	69.90

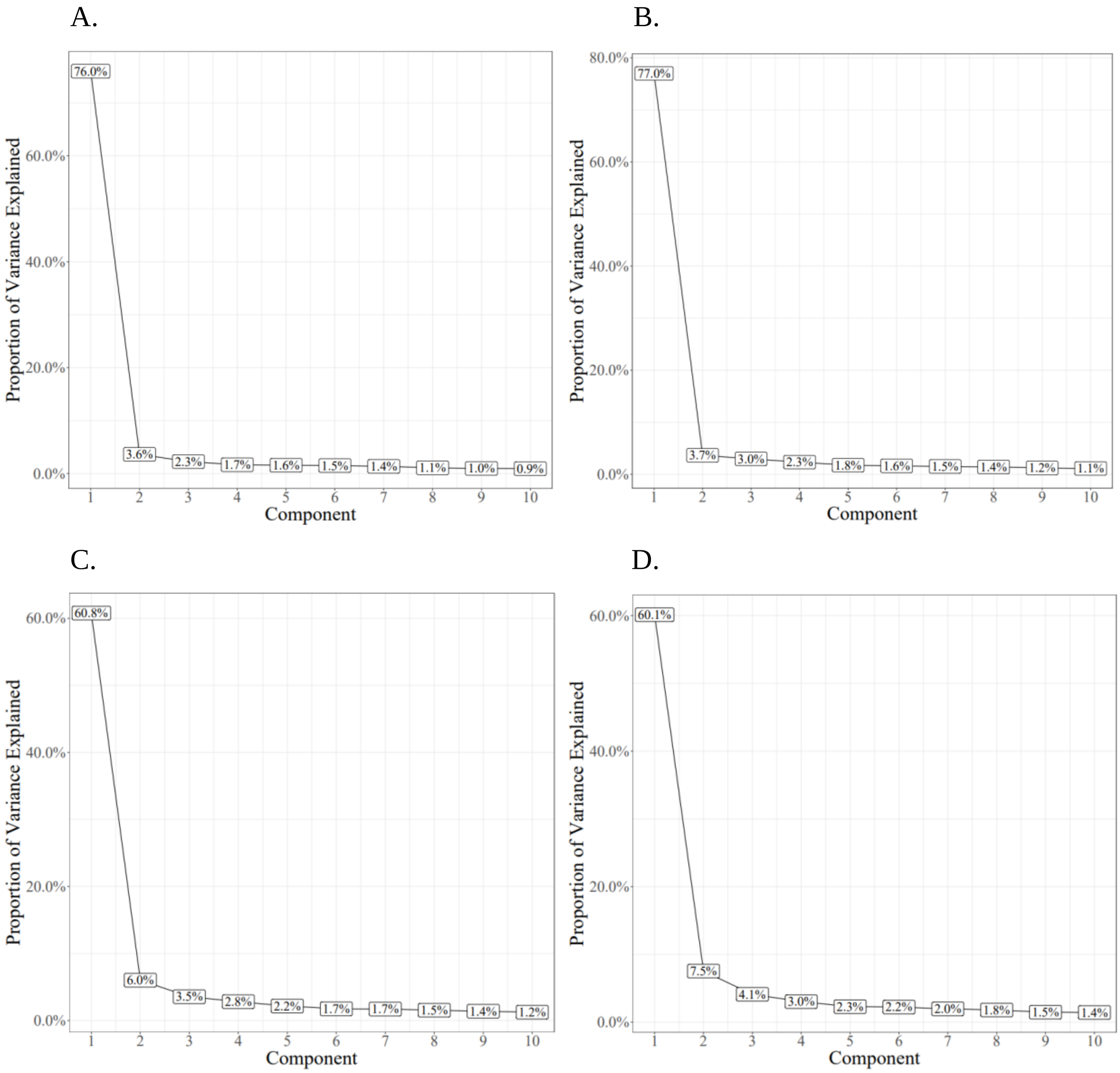
Notes: Sample consists of Bottomless respondents who answered all relevant questions on the survey (Obs. = 701). Includes data on 388 CQs from 60 dimensions, or 316 CQs from the 42 personal and local public good dimensions, collected in Bottomless. The MMBs are matched to the semi-parametric estimate of the mean common-scale SWB of "no anxiety" (54.04), or to the semi-parametric estimate of the mean common-scale SWB of "worthwhileness" (69.90). Standard errors in parentheses.

H.2. An Alternative Measure

As an alternative measure of the fraction of scale use explained by general scale use, we perform principal components analysis (PCA) on the variance-covariance matrix of the dimensional MMBs. After correcting for response errors in the variances, the variance-covariance matrix becomes negative definite. To facilitate the PCA, we substitute the negative eigenvalues with zeros, ensuring positive semidefiniteness of the matrix. To the extent that the majority of the cross-sectional variation in MMBs can be explained by the first principal component, we interpret it as "general" scale use. The relative importance of general scale use can be estimated by the contribution of the first principal component to overall scale-use variation.

Here, we run PCA on the response-error-corrected variance-covariance matrices obtained above from the two sets of CQs. We calculate PCA loadings for 42 or 60 principal components (depending on the number of dimensions we examine) and the eigenvalue for each principal component. In Figure H.1, we plot the proportion of variance explained by the first 10 principal components. As before, we vary the MMB height, using the lowest and highest semi-parametric common-scale SWB means. In all the panels of Figure H.1, the first principal component explains a high proportion (60.1% to 77.0%) of the overall variance in the MMBs. These numbers are very close to what we find using the other measure of the fraction of scale use explained by general scale use in Section H.1.

Figure H.1. Scree Plots for the First 10 Principal Components



Notes: Sample consists of Bottomless respondents who answered all relevant questions on the survey (Obs. = 701).

Panels A and B include data on 316 CQs from 42 personal and local public good dimensions, collected in Bottomless. In Panel A, the MMBs are matched to the semi-parametric estimate of the mean common-scale SWB level of "no anxiety" (54.04), while in Panel B, the MMBs are matched to the semi-parametric estimate of the mean common-scale SWB level of "worthwhileness" (69.90). Panels C and D include data on 388 CQs from 60 dimensions, collected in Bottomless. In Panel C, the MMBs are matched to the semi-parametric estimate of the mean common-scale SWB level of "no anxiety" (54.04), while in Panel B, the MMBs are matched to the semi-parametric estimate of the mean common-scale SWB level of "worthwhileness" (69.90).

I. Validation Analyses

This appendix details validation analyses on our scale-use adjustment methods. The underlying premise of these analyses is that, after adjusting for scale-use heterogeneity, subjective measures should align more closely with their objective counterparts if the scale-use adjustment is performing well.

I.1. Regressing Subjective Measures on Objective Measures

One validation test of our scale-use correction methods examines the relationship between a respondent's "subjective height" (on the 0-100 scale) and "objective height" (in inches, albeit self-reported), as defined in Section VII.A. If scale-use heterogeneity and the response errors we model were the only sources of discrepancy and if there were no measurement error in objective height, then the scale-use-corrected coefficient from a regression of subjective height on objective height, after standardizing both (demeaning and dividing by their standard deviations), would be one. In practice, since both premises are imperfect approximations (for example, subjective height may be judged in comparison to a reference group), we expect the coefficient to be attenuated, but less so after scale-use adjustment than before.

When we regress standardized, unadjusted subjective height on objective height, the regression coefficient is 0.49 (SE = 0.04). To create a scale-use-adjusted analog of this coefficient, we use our MOM estimator to adjust for scale-use heterogeneity that confounds the regression coefficient, and we use our comprehensive MLE to adjust for scale-use heterogeneity that biases the standard deviation used in the procedure of standardization. That is, we benchmark subjective height, r_{is}, using the MMB matching the mean of the common-scale subjective height, $E(w_{is})$, and divide the benchmarked rating by the estimated standard deviation of the common-scale subjective height. This procedure gives us an estimator of the coefficient from a regression of standardized common-scale subjective height on objective height, x_i:

$$Cov\left(\frac{r_{is} - \hat{r}_i\big(E(w_{is})\big)}{\sqrt{Var(w_{is})}}, x_i\right) = Cov\left(\frac{w_{is}}{\sqrt{Var(w_{is})}}, x_i\right) = Cov\left(\frac{w_{is} - E(w_{is})}{\sqrt{Var(w_{is})}}, x_i\right),$$

where the first equality follows from the MOM estimator. The resulting estimate is much closer to, and statistically indistinguishable from, one: 0.85 (SE = 0.09).

However, our scale-use adjustment method adjusts for response errors in addition to correcting for scale-use heterogeneity, and adjusting for response errors alone would increase the regression coefficient by reducing attenuation bias. To isolate the effect of adjusting for scale use, we compare the regression coefficient after scale-use adjustment to the regression coefficient that is obtained by correcting for measurement error only. To correct for measurement error, we model respondent *i*'s raw rating of SWB question *s*, r_{is}, as the true rating, w_{is}, plus response error, v_{is}, that is independent from the true rating and across time:

$$r_{is} = w_{is} + v_{is}.$$

Under this assumption, we can estimate the response-error-corrected standard deviation, $\sqrt{Var(w_{is})}$, using the square root of the test-retest covariance of the subjective ratings of height, $\sqrt{Cov(r_{is1}, r_{is2})}$, where r_{is1} and r_{is2} denote the first rating and the retest rating, respectively, from two pilot waves of our Baseline surveys (see Appendix A.1 for details). When we apply our regression procedure with standardization using the measurement-error-corrected variance estimate $\sqrt{Cov(r_{is1}, r_{is2})}$ (in place of the naïve variance estimate $\sqrt{Var(r_{is})}$), we find a coefficient of 0.62 (SE = 0.06). This coefficient is statistically smaller than the coefficient when both scale use and response errors are corrected for (difference = 0.24 with SE = 0.03).

Looking at subjective and objective measures of weight, we see similar results: the regression coefficient from the unadjusted regression is 0.40 (SE = 0.02), compared with 0.50 (SE = 0.02) after adjusting for only response error, and 0.91 (SE = 0.04) after scale-use adjustment.

As additional validation analyses, we examine four more pairs of subjective and objective measures, one each concerning air quality and crime rate and two concerning financial support.

Table I.1 summarizes the estimates of the slope coefficients from the subjective-objective regressions under no scale-use correction, under only response-error correction, and under our scale-use correction. For all the additional subjective-objective pairs we investigate, we consistently find that each subjective measure corresponds more closely with its objective counterpart after scale-use correction.

Table I.1 Slope Coefficients from Subjective-Objective Regressions

			Regression Coefficient		
Subjective Measure	Objective Measure	Sample size	No scale-use correction	Response-error correction	Scale-use correction
The air in your area not being polluted	Median Air Quality Index	2,810	-0.13 (0.02)	-0.19 (0.02)	-0.21 (0.03)
Your living environment not being spoiled by crime and violence	Violent crime per 100,000	1,345	-0.09 (0.03)	-0.13 (0.04)	-0.14 (0.05)
Financial support for family	Log(HH income/sqrt(HH size))	3,358	0.34 (0.01)	0.40 (0.02)	0.58 (0.02)
Financial support for family	Difficulty paying bills	3,354	-0.49 (0.01)	-0.57 (0.02)	-0.79 (0.02)
Rating of height	Height	3,358	0.49 (0.04)	0.62 (0.06)	0.85 (0.09)
Rating of weight	Weight	3,358	0.40 (0.02)	0.50 (0.02)	0.91 (0.04)

Notes: Subjective measures are from authors' surveys. Median Air Quality Index is measured by the median Air Quality Index in a Metropolitan Statistical Area (MSA), as defined by the U.S. Census Bureau, and comes from the U.S. Environmental Protection Agency. Crime data is measured as the rate of violent crime (defined as murder, nonnegligent manslaughter, rape, robbery, and aggravated assault) per 100,000 residents in an MSA and comes from the F.B.I.'s "Crime in the United States" report. We mapped ZIP Codes (as reported by respondents) to ZIP Code Tabulation Areas (ZCTAs) using the UDS Mapper tool and mapped ZCTAs to MSAs. Sample sizes are smaller for the crime and air quality regressions because not all respondents provided location data that we were able to map to ZIP codes, and because we were not able to map all ZIP codes to the other data sets. Income, household characteristics, height, and weight are self-reported in the Baseline survey demographics. "Difficulty paying bills" comes from the Baseline survey question, "During the last 12 months, would you say you had difficulties paying the bills at the end of the month?" with these response options: Never, Almost never, Occasionally, Most of the time. For this analysis we coded the responses as 0, 1, 2, 3, respectively. All of the subjective and objective measures are standardized. Response-error correction is done using test-retest covariances. For all except "Rating of height" and "Rating of weight," the "test" and "retest" data for the Response-error correction column come from the Baseline survey and Block 1 of the Bottomless survey. For "Rating of height" and "Rating of weight," these data come from a pilot survey with 298 respondents (see Appendix A.1 for details). Standard errors in parentheses.

I.2. Weaker Correlation Between Objective Measures and CQ Ratings

Under our assumptions, relative to the subjective measures that are paired with objective measures (whose correlations are analyzed above), CQ ratings should be much more weakly correlated with the objective measures.

Table I.2 reports the correlations of all the above subjective and objective measures with the mean of our 18 Baseline CQ ratings. A zero correlation with the mean CQ rating cannot be ruled out for four out of the six objective measures (column 2). The other two objective measures' non-zero correlations with the mean CQ rating could be due to possible relationships between the objective measures and correlates of scale use, such as demographics. Column 3 shows that, as hypothesized, relative to the subjective measures, all the objective measures are significantly less (positively) correlated with the mean CQ rating.

Table I.2 Correlation Between Mean CQ Rating and Subjective and Objective Measures

			Correlation with Mean CQ Rating		
			(1)	(2)	(3)
Subjective Measure	Objective Measure	Sample size	Subjective Measure	Objective Measure	Difference
The air in your area not being polluted	Median Air Quality Index	2,810	0.16 (0.01)	0.01 (0.02)	0.15 (0.03)
Your living environment not being spoiled by crime and violence	Violent crime per 100,000	1,345	0.10 (0.02)	-0.01 (0.03)	0.11 (0.03)
Financial support for family	Log(HH income/sqrt(HH size))	3,358	0.23 (0.02)	0.01 (0.02)	0.21 (0.02)
Financial support for family	Difficulty paying bills	3,354	0.23 (0.02)	-0.05 (0.02)	0.28 (0.03)
Rating of height	Height	3,358	0.23 (0.02)	0.00 (0.02)	0.23 (0.02)
Rating of weight	Weight	3,358	0.20 (0.02)	-0.12 (0.01)	0.32 (0.02)

Notes: Subjective measures are from authors' surveys. See notes to Table I.1 for definitions and sources of the objective measures. The subjective and objective measures are not standardized. Standard errors in parentheses.

J. Additional Tables and Figure

This appendix contains tables and a figure that complement the discussion in the paper.

Figure J.1 shows that our distributional assumptions for the scale-use parameters and the response errors provide a reasonable fit between the real and simulated data.

Figure J.1. Density Plots of $\widehat{\alpha}_{i,OLS}$, $\widehat{\beta}_{i,OLS}$, and Standard Deviations of OLS Residuals

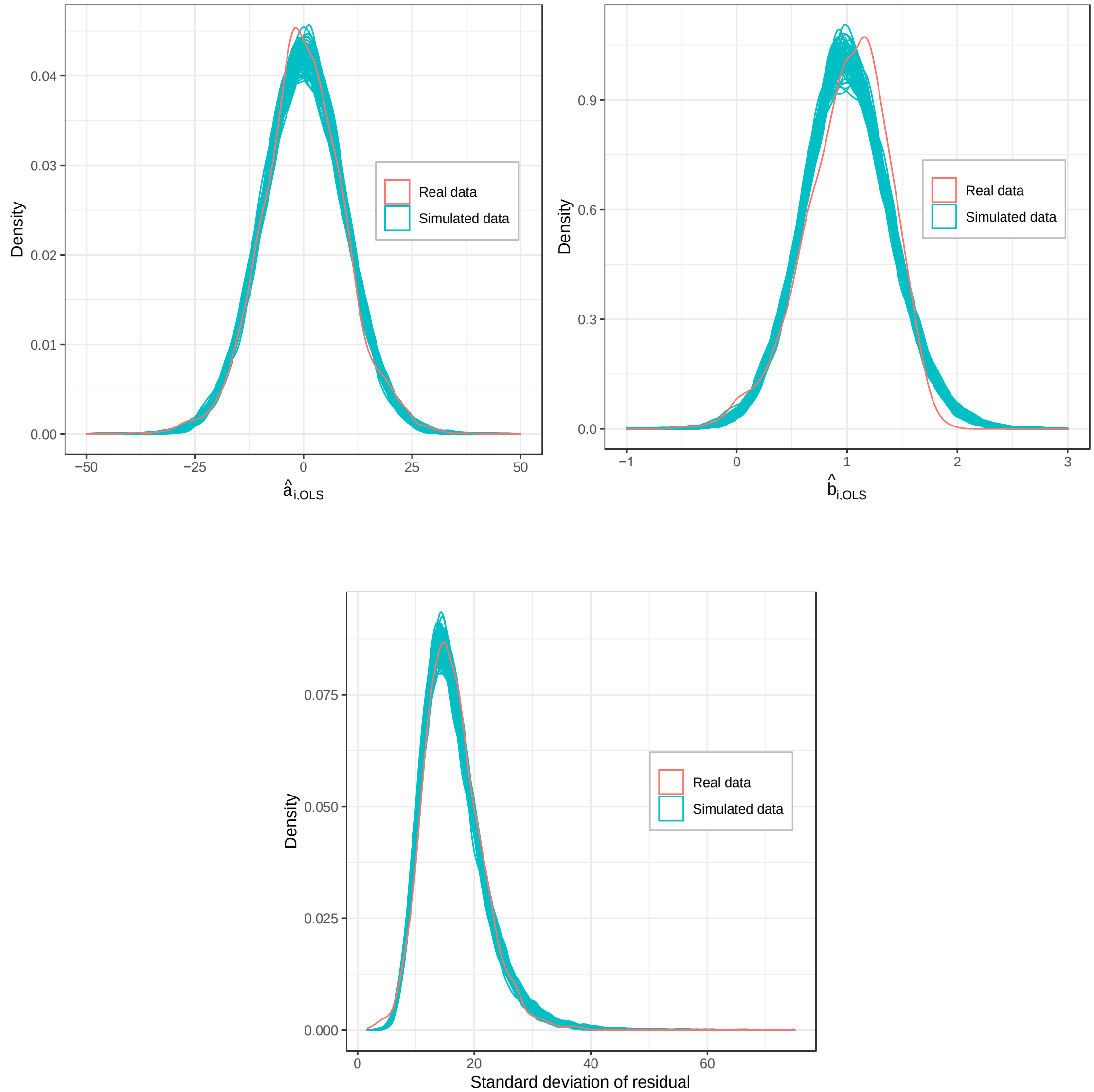


Table J.1 reports the differences in coefficients across our three scale-use adjustment methods and the naïve estimator for demographic regressions, with standard errors calculated from our bootstrap samples. The results are largely similar across the methods.

Table J.1. Cross-Method Differences in Coefficients From Life Satisfaction and "No Anxiety" Regressions

	Life Satisfaction				No Anxiety			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Demographics	No scale-use correction relative to Comprehensive MLE	MOM relative to No scale-use correction	Semi-parametric relative to MOM	Comprehensive MLE relative to Semi-parametric	No scale-use correction relative to Comprehensive MLE	MOM relative to No scale-use correction	Semi-parametric relative to MOM	Comprehensive MLE relative to Semi-parametric
Demeaned age/10	-0.2	0.4	-0.1	-0.1	-0.3	0.6†††	-0.3	0.0
	(0.2)	(0.2)	(0.2)	(0.3)	(0.2)	(0.2)	(0.2)	(0.3)
Demeaned age^2/100	0.2	-0.1	0.1	-0.2	0.2	-0.2†	0.1	-0.1
	(0.1)	(0.1)	(0.1)	(0.2)	(0.1)	(0.1)	(0.1)	(0.2)
Log(HH income)	-0.2	-0.1	0.3	-0.1	-0.7††	0.5†	0.5	-0.2
	(0.3)	(0.3)	(0.4)	(0.4)	(0.3)	(0.2)	(0.4)	(0.5)
Unemployed	-1.3	-1.0	1.8	0.5	-2.2††	0.7	1.0	0.4
	(0.6)	(0.6)	(1.1)	(1.0)	(0.7)	(0.5)	(1.1)	(1.1)
Employed part-time	-0.6	0.1	-0.8	1.3	-1.3†	0.8	-1.1	1.6
	(0.4)	(0.6)	(0.7)	(0.9)	(0.6)	(0.5)	(0.6)	(0.7)
Out of labor force/other	-1.5†	0.3	0.2	0.9	-1.7††	1.6††	0.7	-0.7
	(0.5)	(0.7)	(0.9)	(1.0)	(0.6)	(0.5)	(0.8)	(0.9)
Married, not separated	1.0†	-0.8	-0.4	0.2	1.7†††	-1.6†††	0.0	0.0
	(0.4)	(0.5)	(0.5)	(0.6)	(0.5)	(0.4)	(0.6)	(0.7)
Ever divorced	0.2	-0.1	-0.8	0.7	-0.5	0.2	-1.1	1.4
	(0.6)	(0.6)	(0.7)	(1.0)	(0.6)	(0.5)	(0.8)	(0.9)
Have ≥1 child	0.4	0.4	0.0	-0.8	-0.5	0.5	-0.2	0.3
	(0.4)	(0.6)	(0.6)	(0.7)	(0.5)	(0.5)	(0.5)	(0.7)
Log(HH size)	0.3	-0.5	1.4	-1.1	1.0	-1.2††	0.9	-0.7
	(0.4)	(0.5)	(0.5)	(0.7)	(0.5)	(0.4)	(0.6)	(0.8)
College grad	0.6	-0.8	0.3	-0.1	1.0	-1.1††	0.3	-0.2
	(0.4)	(0.4)	(0.6)	(0.7)	(0.4)	(0.4)	(0.6)	(0.6)
Male	-0.1	-0.7	0.3	0.4	-0.1	-0.4	-0.5	0.9
	(0.3)	(0.4)	(0.5)	(0.5)	(0.4)	(0.3)	(0.5)	(0.5)
Religious attendance (0 to 5, 'Never' to 'More than once a week')	0.5††	-0.4††	0.0	0.0	0.9†††	-0.9†††	0.0	0.0
	(0.1)	(0.1)	(0.2)	(0.2)	(0.1)	(0.1)	(0.2)	(0.2)
Asian	-0.4	1.5	0.2	-1.3	-0.7	1.5†	0.2	-1.0
	(0.7)	(0.8)	(1.0)	(1.1)	(0.8)	(0.7)	(1.0)	(1.1)
N	3,355	3,355	3,355	3,355	3,355	3,355	3,355	3,355

Notes: The differences of the coefficient estimates between the four methods: no scale-use correction, MOM, semi-parametric and comprehensive MLE. Bootstrap standard errors are reported in parentheses. Sample is 3,358 Baseline respondents who passed QC. Daggers signal false-discovery-rate (FDR) significance levels using the Benjamini-Hochberg procedure applied to the 29 *p*-values in each column separately (variables included in FDR correction also include additional race, employment status, region, and day of week indicators; "Other" categories in race and employment status are excluded—we do not pose or report hypothesis tests for them); †††, ††, and † indicate significance at the 1-percent, 5-percent, and 10-percent levels, respectively. Standard errors in parentheses.

Table J.2 reports estimates of the covariances of the 4 U.K. ONS SWB responses before and after general scale-use adjustment by the comprehensive MLE estimator, suggesting that the scale-use adjustment makes a significant difference to the estimates of the covariances.

Table J.2. Estimates of $Cov(w_{is}, w_{is'})$ Before and After Scale-Use Adjustment

Before Scale-Use Adjustment	Life Satisfaction	Happiness	Worthwhileness	No Anxiety
Life Satisfaction	661.8 (14.5)	538.7 (14.4)	453.8 (13.7)	403.5 (13.7)
Happiness	538.7 (14.4)	605.6 (14.0)	427.9 (13.0)	409.1 (12.9)
Worthwhileness	453.8 (13.7)	427.9 (13.0)	554.6 (14.7)	333.8 (13.2)
No Anxiety	403.5 (13.7)	409.1 (12.9)	333.8 (13.2)	830.0 (13.0)
After Scale-Use Adjustment				
Life Satisfaction	273.6 (11.1)	264.6 (9.8)	233.1 (9.4)	212.8 (10.3)
Happiness	264.6 (9.8)	263.6 (9.5)	225.8 (8.7)	224.8 (9.5)
Worthwhileness	233.1 (9.4)	225.8 (8.7)	222.9 (9.9)	173.9 (9.0)
No Anxiety	212.8 (10.3)	224.8 (9.5)	173.9 (9.0)	370.4 (12.4)

Notes: Scale-use adjustment by comprehensive MLE. Standard errors in parentheses.

Table J.3 shows that adjusting for general scale use reduces the estimates of the variances of the 33 SWB questions by around 50%.

Table J.3. SWB Variance Estimates

	(1)	(2)	(3)	(4)
SWB	No scale-use correction	Transitory-error correction	Semi-parametric	Comprehensive MLE
Satisfaction	661.8 (14.5)	512.8 (16.0)	318.8 (26.6)	273.6 (11.1)
Happiness	605.6 (14.0)	456.3 (16.2)	259.9 (25.2)	263.6 (9.5)
Worthwhileness	554.6 (14.7)	401.8 (15.2)	199.8 (27.2)	222.9 (9.9)
No Anxiety	830.0 (13.0)	496.0 (17.4)	446.4 (26.0)	370.4 (12.4)
Ladder	505.6 (11.0)	370.8 (13.0)	123.1 (20.7)	160.5 (8.2)
Well-being of Your Family	363.2 (10.9)	227.0 (11.6)	66.0 (24.1)	121.6 (7.4)
Family Happiness	414.7 (13.5)	270.6 (12.5)	120.0 (24.7)	144.0 (8.8)

SWB	(1) No scale-use correction	(2) Transitory-error correction	(3) Semi-parametric	(4) Comprehensive MLE
Physical Health	435.3 (11.5)	313.5 (12.1)	103.6 (21.5)	106.3 (6.6)
Mental Health	642.0 (12.5)	489.1 (12.6)	320.8 (29.2)	239.5 (8.9)
Sense of Purpose	674.9 (15.1)	504.9 (16.9)	310.9 (25.2)	253.4 (11.1)
Sense of Control	601.0 (13.5)	424.9 (15.8)	229.8 (23.0)	226.8 (10.0)
Having People	628.5 (15.3)	430.8 (15.4)	304.7 (27.0)	249.9 (13.5)
Not Lonely	802.7 (15.4)	501.4 (16.6)	488.5 (30.7)	359.3 (14.0)
No Anger	663.9 (13.4)	350.2 (13.0)	339.7 (23.8)	291.4 (13.3)
No Sadness	755.8 (12.5)	458.7 (17.8)	409.2 (26.4)	328.2 (11.8)
No Stress	763.6 (11.2)	490.2 (12.9)	393.0 (24.3)	330.0 (10.5)
No Worry	772.5 (12.3)	498.4 (14.1)	385.0 (23.4)	338.5 (10.9)
Good Person	266.8 (9.9)	175.9 (6.80)	41.6 (25.4)	72.3 (6.4)
Possibilities	580.5 (13.3)	365.9 (15.9)	223.7 (23.4)	205.9 (9.0)
Time	557.2 (10.8)	332.5 (13.0)	204.4 (25.5)	196.3 (9.4)
Social Status	716.1 (12.7)	535.9 (17.5)	317.7 (22.8)	251.1 (9.7)
Safety	277.5 (9.1)	149.1 (8.17)	-17.3 (22.8)	65.9 (5.6)
Financial Support	807.4 (15.2)	588.1 (22.2)	430.8 (27.9)	271.6 (11.2)
Not Unemployed	951.2 (15.0)	523.5 (20.0)	660.1 (32.4)	476.0 (14.9)
Eat	322.0 (12.3)	168.5 (9.28)	59.3 (23.2)	56.9 (4.5)
Housing Comfort	400.8 (12.7)	260.8 (13.0)	118.0 (26.9)	107.2 (9.2)
Enjoyment	562.9 (13.7)	432.6 (15.0)	236.0 (24.3)	243.3 (9.5)
Knowledge Skills	271.0 (7.5)	177.1 (8.26)	-31.6 (19.8)	66.4 (5.0)
Local Safety	463.4 (14.2)	201.5 (13.2)	204.1 (26.0)	131.1 (14.1)
Local Air	497.6 (12.1)	232.6 (13.5)	218.6 (24.3)	163.7 (11.9)
Citizen Influence	645.2 (10.9)	361.8 (13.3)	236.9 (17.5)	252.6 (10.1)
Citizen Trust	556.1 (9.3)	342.1 (13.2)	183.4 (19.8)	186.9 (8.3)
Culture Being Honored	514.5 (10.6)	250.9 (11.8)	178.9 (23.6)	177.2 (10.2)

Notes: Column (2) is calculated based on our Baseline survey and the repeat SWB questions in Block 1 of our follow-up Bottomless survey. Sample size for Column (2) is N = 2,472. Standard errors in parentheses.

Table J.4 decomposes the raw variances of the four U.K. SWB questions into three components: common-scale SWB (our estimate of $Var(w_{is})$), general scale use (our estimate of $Var\left(\alpha_i + \gamma + \beta_i(w_{is} - \gamma)\right) - Var(w_{is})$), and response errors (our estimate of $Var(r_{is}) - Var\left(\alpha_i + \gamma + \beta_i(w_{is} - \gamma)\right)$). Similar to what is implied by the estimates in Table J.3, the estimates in Table J.4 show that variance in common-scale SWB accounts for around 43% of the variance in reported SWB. Heterogeneity in scale-use accounts for around 12%, and response-error variances account for the remaining about 45%.

Table J.4. Decomposition of Response Variance

SWB	$Var(r_{is})$	Fraction due to common-scale SWB	Fraction due to general scale use	Fraction due to response errors
Life Satisfaction	661.8 (14.5)	0.41 (0.01)	0.11 (0.01)	0.47 (0.01)
Happiness	605.6 (14.0)	0.44 (0.01)	0.12 (0.01)	0.44 (0.01)
Worthwhileness	554.6 (14.7)	0.40 (0.01)	0.13 (0.01)	0.47 (0.01)
No Anxiety	830.0 (13.0)	0.45 (0.01)	0.12 (0.00)	0.44 (0.01)

Notes: Fraction due to common-scale SWB is calculated through $Var(w_{is})/Var(r_{is})$, fraction due to general scale use through $[Var(\alpha_i + \gamma + \beta_i(w_{is} - \gamma)) - Var(w_{is})]/Var(r_{is})$, and fraction due to errors through $[Var(r_{is}) - Var(\alpha_i + \gamma + \beta_i(w_{is} - \gamma))]/Var(r_{is})$. Relevant moments of w_{is} are estimated by the comprehensive MLE method. Standard errors in parentheses.

Tables J.5 and J.6 are the full versions of Tables 3 and 6, respectively, including demographic coefficients that were omitted in the abbreviated versions.

Table J.5. Regression of Mean and Standard Deviation of CQs, and $\hat{\alpha}_i$ and $\hat{\beta}_i$, on Demographics (Full Version)

	(1)	(2)	(3)	(4)
Demographics	Mean of CQs	Std. Dev. of CQs	$\hat{\alpha}_i$	$\hat{\beta}_i$
Demeaned age/10	$-0.7^{\dagger\dagger\dagger}$ (0.2)	$0.2^{\dagger\dagger}$ (0.1)	$-0.51^{\dagger\dagger}$ (0.16)	$0.01^{\dagger}$ (0.01)
(Demeaned age)^2/100	$0.2^{\dagger\dagger}$ (0.1)	-0.1 (0.1)	0.13 (0.08)	$-0.01^{\dagger\dagger}$ (0.004)
Log(HH income)	$-0.7^{\dagger\dagger\dagger}$ (0.2)	$0.4^{\dagger\dagger\dagger}$ (0.1)	-0.23 (0.24)	$0.05^{\dagger\dagger\dagger}$ (0.01)
Unemployed	$-1.3^{\dagger\dagger\dagger}$ (0.5)	$1.0^{\dagger\dagger\dagger}$ (0.3)	0.05 (0.51)	$0.13^{\dagger\dagger\dagger}$ (0.02)
Employed part-time	$-1.1^{\dagger}$ (0.5)	$0.7^{\dagger\dagger}$ (0.3)	-0.50 (0.53)	$0.06^{\dagger\dagger\dagger}$ (0.02)
Out of labor force/other	-2.1 (0.5)	0.7 (0.3)	-1.09 (0.54)	0.12 (0.02)
Married, not separated	$1.9^{\dagger\dagger\dagger}$	$-0.7^{\dagger\dagger\dagger}$	$1.24^{\dagger\dagger}$	$-0.06^{\dagger\dagger\dagger}$

	(0.4)	(0.2)	(0.39)	(0.02)
Ever divorced	-0.4	0.4	-0.07	0.03
	(0.5)	(0.3)	(0.56)	(0.02)
Have ≥1 child	-0.5	0.2	-0.45	0.00
	(0.4)	(0.3)	(0.48)	(0.02)
Log(HH size)	1.4†††	-0.8†††	0.89	-0.05††
	(0.4)	(0.2)	(0.43)	(0.02)
College grad	1.2†††	-0.9†††	0.94†	-0.02
	(0.4)	(0.2)	(0.37)	(0.01)
Male	0.3	-0.1	0.54	0.02
	(0.3)	(0.2)	(0.34)	(0.01)
Religious attendance (0 to 5, 'Never' to 'More than once a week')	1.1†††	-0.4†††	0.66†††	-0.04†††
	(0.1)	(0.1)	(0.10)	(0.004)
Asian	-1.5†	0.2	-1.53	0.00
	(0.7)	(0.4)	(0.73)	(0.02)
Democrat	1.3†††	-0.7†††	0.80	-0.04†††
	(0.3)	(0.2)	(0.36)	(0.01)
Obese	-1.2†††	0.4	-1.16††	0.004
	(0.4)	(0.2)	(0.38)	(0.02)
Black/African American	-0.3	0.8†	-0.23	0.01
	(0.7)	(0.4)	(0.73)	(0.03)
Hispanic/Latino/Spanish	1.0	-0.3	0.66	-0.03
	(0.7)	(0.4)	(0.74)	(0.03)
Other non-white	-0.7	-0.02	-1.19†	-0.05
	(0.6)	(0.5)	(0.72)	(0.03)
Northeast	0.2	0.2	0.11	-0.01
	(0.5)	(0.3)	(0.55)	(0.02)
West	-0.3	0.4	-0.49	-0.01
	(0.4)	(0.3)	(0.48)	(0.02)
South	-0.2	0.6†	0.04	0.02
	(0.4)	(0.3)	(0.39)	(0.02)
High population density	0.5	-0.3	0.20	-0.03
	(0.4)	(0.2)	(0.43)	(0.02)
Sunday	1.4	0.4	1.03	-0.03
	(1.0)	(0.5)	(1.04)	(0.04)
Monday	0.6	-0.2	0.35	-0.03
	(0.5)	(0.3)	(0.54)	(0.02)
Tuesday	0.4	-0.4	0.12	-0.02
	(0.7)	(0.4)	(0.75)	(0.03)
Thursday	0.7	-0.1	0.55	-0.02
	(0.6)	(0.4)	(0.62)	(0.03)
Friday	1.0	-0.6	0.44	-0.05
	(0.5)	(0.3)	(0.57)	(0.02)
Saturday	1.1	-1.0†	0.64	-0.04
	(0.9)	(0.5)	(0.93)	(0.03)

Obs.	3,358	3,358	3,358	3,358

Notes: The dependent variables are constructed from each individual's responses to the 18 Baseline CQs. $\hat{\alpha}_i$ and $\hat{\beta}_i$ are the intercept and slope, respectively, from regression of respondent i's 18 Baseline CQ ratings onto the population means of those 18 ratings. The sample is 3,358 Baseline respondents who passed quality control. Daggers signal false-discovery-rate (FDR) significance levels using the Benjamini-Hochberg procedure applied to the 29 *p*-values in each column separately (variables included in FDR correction also include additional race and employment status indicators, as well as indicators for region, day of week, political party, obesity, and population density; "Other" categories in race and employment status are excluded from FDR correction—we do not pose or report hypothesis tests for them); †††, ††, and † indicate significance at the 1-percent, 5-percent, and 10-percent levels, respectively. Standard errors in parentheses.

Table J.6. Life Satisfaction and "No Anxiety" Regression and General Scale-Use Adjustment (Full Version)

	Life Satisfaction					No Anxiety				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Demographics	No scale-use correction	MOM	Semi-parametric	Compre-hensive MLE	MMB (66.90)	No scale-use correction	MOM	Semi-parametric	Compre-hensive MLE	MMB (54.04)
Demeaned age/10	1.3†††	1.7†††	1.6†††	1.6†††	-0.4	3.6†††	4.2†††	3.9†††	3.9†††	-0.6†††
	(0.4)	(0.4)	(0.5)	(0.4)	(0.2)	(0.5)	(0.4)	(0.5)	(0.4)	(0.2)
Demeaned age^2/100	1.7†††	1.7†††	1.7†††	1.6†††	0.1	1.1†††	0.9†††	1.0†††	0.9†††	0.2†
	(0.2)	(0.2)	(0.3)	(0.2)	(0.1)	(0.3)	(0.3)	(0.3)	(0.3)	(0.1)
Log(HH income)	5.1†††	5.1†††	5.4†††	5.3†††	0.1	2.0††	2.4††	2.9†††	2.7†††	-0.5†
	(0.7)	(0.6)	(0.8)	(0.6)	(0.3)	(0.8)	(0.8)	(0.9)	(0.8)	(0.2)
Unemployed	-8.3†††	-9.2†††	-7.5†††	-7.0†††	1.0	-7.7†††	-7.0†††	-6.0†††	-5.6†††	-0.7
	(1.6)	(1.5)	(1.5)	(1.3)	(0.6)	(2.0)	(1.8)	(2.0)	(1.7)	(0.5)
Employed part-time	-2.9†	-2.8†	-3.6†	-2.3	-0.1	-5.9†††	-5.1†††	-6.2†††	-4.6†††	-0.8
	(1.3)	(1.4)	(1.6)	(1.2)	(0.6)	(1.5)	(1.5)	(1.5)	(1.4)	(0.5)
Out of labor force/other	-6.0†††	-5.6†††	-5.4†††	-4.5†††	-0.3	-6.0†††	-4.4††	-3.7	-4.3††	-1.6
	(1.4)	(1.5)	(1.7)	(1.3)	(0.6)	(1.7)	(1.7)	(1.9)	(1.6)	(0.5)
Married, not separated	9.7†††	8.9†††	8.5†††	8.6†††	0.8	4.7†††	3.1††	3.1†	3.1††	1.6†††
	(1.0)	(1.0)	(1.3)	(1.0)	(0.4)	(1.4)	(1.3)	(1.5)	(1.2)	(0.4)
Ever divorced	2.3	2.2	1.4	2.1	0.1	-0.1	0.2	-0.9	0.5	-0.2
	(1.3)	(1.3)	(1.6)	(1.2)	(0.6)	(1.7)	(1.7)	(1.9)	(1.7)	(0.5)
Have ≥1 child	4.1†††	4.5†††	4.6†††	3.8†††	-0.4	1.1	1.6	1.4	1.7	-0.5
	(1.0)	(1.0)	(1.3)	(0.9)	(0.6)	(1.2)	(1.2)	(1.4)	(1.1)	(0.4)
Log(HH size)	-2.1†	-2.7††	-1.3	-2.4†††	0.5	-1.2	-2.4†	-1.5	-2.2	1.2††
	(1.1)	(1.1)	(1.3)	(1.0)	(0.5)	(1.2)	(1.1)	(1.4)	(1.1)	(0.4)
College grad	2.1†	1.3	1.6	1.5	0.8	3.3†††	2.2†	2.5†	2.3†	1.1††
	(1.0)	(0.9)	(1.2)	(0.9)	(0.4)	(1.0)	(1.0)	(1.2)	(1.0)	(0.4)
Male	0.1	-0.6	-0.3	0.1	0.7	5.7†††	5.2†††	4.8†††	5.7†††	0.4
	(0.9)	(0.9)	(1.1)	(0.9)	(0.4)	(1.0)	(1.0)	(1.2)	(1.0)	(0.3)
Religious attendance (0 to 5, 'Never' to 'More than once a week')	1.9†††	1.6†††	1.5†††	1.5†††	0.4††	1.8†††	0.9†††	1.0††	1.0†††	0.9†††
	(0.2)	(0.2)	(0.3)	(0.2)	(0.1)	(0.3)	(0.3)	(0.4)	(0.3)	(0.1)
Asian	0.0	1.5	1.7	0.4	-1.5	2.5	4.0†	4.2	3.2	-1.5†
	(1.9)	(1.8)	(2.1)	(1.8)	(0.8)	(2.0)	(1.9)	(2.2)	(1.9)	(0.7)

	Life Satisfaction					No Anxiety				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Demographics	No scale-use correction	MOM	Semi-parametric	Compre-hensive MLE	MMB (66.90)	No scale-use correction	MOM	Semi-parametric	Compre-hensive MLE	MMB (54.04)
Democrat	0.2	-0.3	-0.8	-0.1	0.5	-0.1	-1.2	-0.5	-1.1	1.1††
	(0.8)	(0.7)	(0.9)	(0.8)	(0.4)	(1.0)	(0.9)	(1.0)	(0.9)	(0.3)
Obese	-4.0†††	-2.9††	-2.3	-2.9†††	-1.1	-4.6†††	-3.5††	-3.0†	-3.8†††	-1.2†††
	(1.1)	(1.1)	(1.3)	(0.9)	(0.4)	(1.3)	(1.3)	(1.3)	(1.2)	(0.4)
Black/African American	1.6	1.8	1.3	1.6	-0.2	7.3†††	7.6†††	6.8†††	7.2†††	-0.3
	(1.7)	(1.7)	(2.0)	(1.6)	(0.8)	(2.0)	(1.9)	(2.2)	(1.8)	(0.7)
Hispanic/Latino/ Spanish	0.5	0.1	-1.7	-0.5	0.4	2.5	1.6	0.2	1.2	0.9
	(1.6)	(1.6)	(2.1)	(1.6)	(0.8)	(1.8)	(1.8)	(2.2)	(1.9)	(0.7)
Other non-white	-0.3	1.2	2.0	0.5	-1.5†	2.7	3.6	6.2†	3.1	-0.9
	(1.7)	(1.6)	(1.9)	(1.6)	(0.8)	(2.2)	(2.0)	(2.6)	(2.1)	(0.7)
Northeast	1.3	1.3	1.5	0.9	0.1	0.9	0.7	0.8	0.6	0.2
	(1.3)	(1.3)	(1.6)	(1.2)	(0.6)	(1.6)	(1.5)	(1.8)	(1.4)	(0.5)
West	2.9†	3.5††	3.0	3.0††	-0.6	-0.3	0.2	-0.2	0.5	-0.4
	(1.3)	(1.3)	(1.6)	(1.2)	(0.5)	(1.6)	(1.5)	(1.8)	(1.4)	(0.4)
South	1.6	1.4	0.7	1.3	0.2	0.5	0.6	0.5	0.7	-0.1
	(0.9)	(0.9)	(1.2)	(0.9)	(0.4)	(1.2)	(1.2)	(1.4)	(1.1)	(0.4)
High population density	-2.6†	-2.6†	-2.8	-2.6††	-0.03	-0.2	-0.6	-0.7	-1.0	0.4
	(1.1)	(1.1)	(1.3)	(1.0)	(0.5)	(1.2)	(1.3)	(1.5)	(1.2)	(0.4)
Sunday	-2.6	-3.4	-3.3	-4.5	0.8	-4.5	-5.7	-6.0	-5.7	1.2
	(2.9)	(2.9)	(3.4)	(2.8)	(1.2)	(3.7)	(3.6)	(3.9)	(3.6)	(1.0)
Monday	-1.3	-1.4	-0.7	-1.0	0.2	-3.3†	-3.8††	-4.4††	-2.8†	0.5
	(1.6)	(1.4)	(1.8)	(1.4)	(0.6)	(1.4)	(1.4)	(1.7)	(1.4)	(0.5)
Tuesday	1.9	2.0	3.6	1.8	-0.04	0.7	0.5	1.5	0.9	0.3
	(1.7)	(1.7)	(2.0)	(1.6)	(0.8)	(1.9)	(1.9)	(2.2)	(1.7)	(0.7)
Thursday	-1.8	-2.3	-1.6	-2.0	0.4	-2.4	-3.0	-3.6	-2.2	0.7
	(1.8)	(1.7)	(2.1)	(1.6)	(0.7)	(1.8)	(1.8)	(2.0)	(1.7)	(0.6)
Friday	-0.5	-0.6	-1.2	-1.1	0.1	-1.3	-2.0	-3.2	-1.4	0.7
	(1.6)	(1.4)	(1.8)	(1.4)	(0.7)	(1.7)	(1.7)	(2.1)	(1.6)	(0.5)
Saturday	-2.2	-2.5	-2.7	-2.7	0.4	-1.0	-1.9	-3.5	-1.3	0.9
	(2.2)	(2.2)	(2.5)	(2.1)	(1.0)	(2.2)	(2.0)	(2.4)	(2.0)	(0.9)
Obs.	3,358	3,358	3,358	3,358	3,358	3,358	3,358	3,358	3,358	3,358

Notes: The sample is 3,358 Baseline respondents who passed quality control. Dependent variables for columns (5) and (10) are MMBs, matched to the semi-parametric estimates of Life Satisfaction (66.90) and No Anxiety (54.04) means, respectively. Daggers signal false-discovery-rate significance levels using the Benjamini-Hochberg procedure applied to the 29 *p*-values in each column separately. See Table J.5 notes for description of FDR correction procedure and significance levels. Standard errors in parentheses.