

Pre-Analysis Plan

Living Large or Long

Amitabh Chandra
Tomáš Jagelka
Erzo F.P. Luttmer
Joshua Schwartzstein

March 14, 2025

1. Introduction

A variety of resource allocation decisions depend on how much money people are willing to give up to live an extra year of life. We develop a simple approach to estimate this rate of substitution by presenting respondents on the *Prolific* platform with pairs of hypothetical life stories, or “vignettes,” that describe complete life courses, and ask them to choose which life they would prefer for themselves.

Our main measure of people’s valuation of longevity is based on their choices over pairs of vignettes. For comparison purposes, we also ask respondents to make choices over pairs of “tuples.” Tuples are vignettes that are stripped of all context so that the only key information, income and age of death, is shown to the respondent. Two non-key pieces of information are also shown, the name of the protagonist (so that the question sounds less abstract) and the protagonist’s year of birth (so that the respondent knows when the protagonist lived). We treat the valuation of longevity inferred from tuple choices as a comparison metric rather than as our key metric.

2. Data Collection

The study will use *Prolific* respondents who are US residents, have US Nationality, have completed at least 100 surveys, and have an approval rating of 95% or more. We also have *Prolific* select an equal proportion of male and female respondents. The survey consists of two waves. Only respondents who completed the first wave, and were not excluded from the baseline sample (see Section 2.4 for sample exclusion criteria), are invited for the second wave.

2.1 Wave 1

The target sample size for Wave 1 is 3750 completed responses, with the exact number depending on response rates. Respondents who participated in any of our pilots are excluded from the sample frame.

Wave 1 is fielded in three parts. The first part consists of a test run of the survey instrument with 100 respondents. If this test run reveals no implementation issues or mistakes, it becomes part of the main sample. If it does reveal implementation issues or major mistakes, the data from it is discarded and we redo part 1 with 100 different respondents.

The second part consists of 2600 completed responses from respondents who did not participate in any of our pilots or in part 1 of Wave 1. When part 2 of Wave 1 has closed, we launch part 3 of Wave 1, which consists of targeting an additional 1050 respondents who did not participate in any of our pilots, nor in part 1 or part 2 of Wave 1. We have a separate part 2 and part 3 because administrative restrictions related to the source of funding required us to run the survey from two separate *Prolific* accounts. By running surveys from these two accounts sequentially, we can ensure that no respondent takes our survey more than once in a given wave.

2.2 Wave 2

Within one week of the first day that occurs four weeks after Wave 1 part 1 was launched or one week after Wave 1 part 1 was closed (whichever is later), we launch Wave 2 part 1, to which those who completed Wave 1 part 1, and were not excluded from the baseline sample, are invited. We keep Wave 2 part 1 open for four weeks. If Wave 2 part 1 reveals implementation issues or major mistakes, the data from it is discarded, but the data from Wave 1 part 1 will be retained.

Within one week of the first day that occurs four weeks after Wave 1 part 2 was launched or one week after Wave 1 part 2 was closed (whichever is later), we launch Wave 2 part 2, to which those who completed Wave 1 part 2, and were not excluded from the baseline sample, are invited. We keep Wave 2 part 2 open for four weeks.

Within one week of the first day that occurs four weeks after Wave 1 part 3 was launched or one week after Wave 1 part 3 was closed (whichever is later), we launch Wave 2 part 3, to which those who completed Wave 1 part 3, and were not excluded from the baseline sample, are invited. We keep Wave 2 part 3 open for four weeks.

2.3 Completion Sample

We expect about 70% to 80% of respondents to pass our baseline sample inclusion criteria, and we expect the response rate to Wave 2 to be between 80% and 90% of those invited to take Wave 2. Because we invite respondents to Wave 2 only if they passed our baseline sample inclusion criteria, we expect that between 56% and 72% of those invited to Wave 1 will complete Wave 2. Hence, the total number of completed surveys (i.e., respondent-wave observations) should end up somewhere between $1.56 * 3750$ and $1.72 * 3750$. If the number of completed Wave 2 surveys is less than 55% of the number of completed Wave 1 surveys, we will collect a “completion sample” such that we expect to obtain $1.65 * 3750$ completed surveys in total. Specifically, if the realized fraction of completed Wave 2 surveys divided by the number of completed Wave 1 surveys was r and $r < 0.55$, we will field Wave 1 part 4 with a sample size of $(0.65 - r) * 3750 / (1 + r)$.

Within one week of the first day that occurs four weeks after Wave 1 part 4 was launched or one week after Wave 1 part 4 was closed (whichever is later), we launch Wave 2 part 4, to which those who completed Wave 1 part 4, and were not excluded from the baseline sample, are invited. We keep Wave 2 part 4 open for four weeks.

2.4 Respondent Exclusion Criteria

Exclusions occur at the respondent-wave level for data that we collect but, as noted above, we do not invite respondents to Wave 2 if they meet the exclusion criteria for Wave 1. We include only completed surveys in the analyses though respondents are allowed to not respond to certain items. Respondents are excluded from the baseline analyses for a given wave if any of the following conditions hold:

- They do not complete all six vignette pair choices and all six tuple pair choices in a given wave.
- They respond to any of the vignette pair choices in less than 30 seconds in a given wave.
- If we detect, for a given wave, that their *Prolific* ID was used for more than one survey. In that case, we exclude all surveys that used that *Prolific* ID. If we learn of additional ways to detect repeat survey takers within a wave, we will also exclude respondents we flag through such additional detection methods.
- If we detect, for a given wave, that they used AI to answer survey questions. We use three checks to detect AI and exclude a respondent if they fail any of these three checks. First, we have a two-part attention check in which we tell respondents early in the survey what they should answer if they get a particular question. They pass the check if they give the answer they were instructed to give when they receive the particular question later in the survey. Second, there is one question that is invisible on the screen but that still appears in html code, and thus would be answered by bots processing html code. A respondent fails this check if they answer the invisible question. Third, we calculate the within-respondent within-wave standard deviation in the response times to the six vignette pair questions. We consider a standard deviation less than two seconds to be suspicious and an indication of a bot taking the survey. A respondent fails the third check if the standard deviation of their response times to the six vignette pair questions is less than two seconds in a given wave.

The survey also includes in each wave a standard attention check in which we tell respondents to check two particular answer boxes independently of what they truly think. This attention check is not used to screen out participants. Instead, we include it to gauge the response quality of the included sample and of the excluded sample.

3. Empirical Analysis

3.1 Baseline Specification

We have 28 individual vignettes, each with a numerical ID. Without loss of generality, for each pair of vignettes, we refer to the vignette with the smaller ID value as “vignette A” in the pair. The other vignette of the pair, the vignette with the larger ID value, is referred to as “vignette B” in the pair.

The specification for the baseline regression is an OLS regression of the form:

$$ChoseA_{ci} = \beta_1 \Delta \ln(Longevity_{ci}) + \beta_2 \Delta \ln(Income_{ci}) + \beta_3 \Delta YearOfBirth_{ci} + \delta_{v(A,c,i)} - \delta_{v(B,c,i)} + \epsilon_{ci}$$

Where:

- c indexes vignette pairs seen by respondent i , so $c = 1$ for the first choice respondent i makes, $c = 2$ for the second choice respondent i makes, etc. Hence, c runs from 1 to 6 for vignette choices in Wave 1 of the survey and from 7 to 12 for vignette choices in Wave 2 of the survey (given that only those who completed Wave 1 are invited to Wave 2)
- $ChoseA_{ci}$ is a dummy variable that equals 1 if respondent i chose vignette A when they made their choice between the c^{th} pair of vignettes.
- $\Delta \ln(Longevity_{ci}) \equiv \ln(Longevity_{Aci}) - \ln(Longevity_{Bci})$ is the natural logarithm of the age of death of the protagonist in vignette A minus the natural logarithm of the age of death of the protagonist in vignette B for the vignettes shown to respondent i when they made their c^{th} choice over pairs of vignettes.
- $\Delta \ln(Income_{ci}) \equiv \ln(Income_{Aci}) - \ln(Income_{Bci})$ is the natural logarithm of the income of the protagonist in vignette A minus the natural logarithm of the income of the protagonist in vignette B for the vignettes shown to respondent i when they made their c^{th} choice over pairs of vignettes.
- $\Delta YearOfBirth_{ci} \equiv YearOfBirth_{Aci} - YearOfBirth_{Bci}$ is year of birth of the protagonist in vignette A minus the year of birth of the protagonist in vignette B for the vignettes shown to respondent i when they made their c^{th} choice over pairs of vignettes. This control holds the year of birth constant, which ensures that increases in longevity imply later years of death rather than earlier years of birth.
- $\delta_{v(A,c,i)}$ is a fixed effect for vignette v , which was shown as the “A” vignette when respondent i made their c^{th} choice over pairs of vignettes. There are 28 vignettes used in the study, so v runs from 1 to 28.
- $\delta_{v(B,c,i)}$ is defined analogously. It enters with a negative sign because if vignette B is more attractive, the respondent is less likely to choose vignette A.
- ϵ_{ci} is an error term that is clustered by respondent.

We cluster standard errors by respondent. The baseline regression is run using the observations of all respondents that meet the inclusion criteria, pooling respondents across both waves of the survey.

The Life Year Value Elasticity (LYVE) is the percentage change in income that respondents value as much as a 1% increase in longevity, and is formally defined by:

$$LYVE \equiv \frac{\partial ChoseA / \partial \ln Longevity_A}{\partial ChoseA / \partial \ln Income_A},$$

where, in slight abuse of notation, $Longevity_A$ denotes the longevity of the protagonist in vignette A and $Income_A$ denotes the income of the protagonist in vignette A. In the baseline specification, the LYVE is given by β_1 / β_2 . We use the LYVE as our baseline metric for expressing how much respondents value marginal increases in longevity.

3.2 Validation Analyses

3.2.1 Attitudes and Behaviors Predicting the LYVE

We create an index based on five respondent attitudes and five self-reported behaviors that we believe, on a priori grounds, should be predictive of the LYVE. Each of the attitudes and behaviors is first standardized, and reverse-coded where necessary to ensure that higher values correspond to an attitude or behavior that we expect to be associated with a higher LYVE. Next, we take the average over the responses for respondents who answered an index question in both waves. For those respondents who answered an index question in only one wave, we use that answer. Next, we create the index for each respondent by averaging over non-missing index questions for that respondent. Finally, we standardize the index and replace any missing values with zero.

The components of the index are (text in italics indicates actual question wording):

Attitudes (Respondents rate how well the statement describes them on a scale from 0 to 10)

- *Living a long life is very important to me.*
- *I would rather die than live with debilitating pain.* (Reverse coded)
- *Living a life that is financially comfortable is important to me.* (Reverse coded)
- *I'm willing to work extra to increase my pay.* (Reverse coded)
- *I would rather live a longer life with fewer luxuries than a shorter life with greater material comforts.*

Behaviors (Converting categorical answers to a linear numeric scale)

- Health checkup: *How often do you miss your recommended checkup with your primary care physician?* (Reverse coded)
- Healthy eating: *How often do you eat fruits and vegetables as part of your meals?*
- Exercise: *In a typical week, how many days do you exercise?*
- Preferred healthy eating: *How often would you PREFER to eat fruits and vegetables as part of your meals?*
- Preferred exercise: *In a typical week, how many days would you PREFER to exercise?*

We calculate $dLYVE/dIndex$ using the delta method after adding the following three terms to the baseline regression:

- $\beta_4 Index_i * \Delta \ln (Longevity_{ci})$
- $\beta_5 Index_i * \Delta \ln (Income_{ci})$
- $\beta_6 Index_i$

We will also report how the LYVE varies with two subindices and with the ten individual components. The two subindices are formed analogously to the main index, but average only either the five attitudes or the five behaviors. The derivative of the LYVE with respect to the two subindices or the ten components is calculated analogously to $dLYVE/dIndex$ by replacing *Index* by a subindex or a component.

3.2.2 Within-Person Consistency of Choices Over Time and Across Contexts

We use two approaches to evaluate whether respondents have a stable underlying valuation of longevity.

First, we calculate the test-retest correlation for the probability of choosing vignette A for the vignette pair that was randomly selected to be asked in an identical fashion in both waves of the survey for a given respondent. As a comparison, we will also show the test-retest correlation for the tuple that was randomly selected to be asked in an identical fashion in both waves of the survey for a given respondent. As additional comparison, we will show test-retest correlations for the following attitudes and behaviors collected in both waves of the survey (text in italics indicates actual question wording):

- *Living a long life is very important to me.*
- *I would rather die than live with debilitating pain.*
- *Living a life that is financially comfortable is important to me.*
- *I'm willing to work extra to increase my pay.*
- *I would rather live a longer life with fewer luxuries than a shorter life with greater material comforts.*
- Health checkup: *How often do you miss your recommended checkup with your primary care physician?*
- Healthy eating: *How often do you eat fruits and vegetables as part of your meals?*
- Exercise: *In a typical week, how many days do you exercise?*
- Preferred healthy eating: *How often would you PREFER to eat fruits and vegetables as part of your meals?*
- Preferred exercise: *In a typical week, how many days would you PREFER to exercise?*
- Life satisfaction
- Self-reported optimism
- *I have difficulty picturing the details of a situation I have not previously experienced.*
- *The things I imagine are not very detailed.*
- Willingness to give to good causes without expecting anything in return
- Willingness to give up something that is beneficial today in order to benefit more in the future
- Willingness to take risks
- Others' intentions: *I assume that people have only the best intentions.*
- Present bias: *I tend to postpone things even though it would be better to take care of them right away.*

- Locus of Control, which is an index that equals the standardized mean of all seven locus-of-control questions on the survey (outlined in subsection 3.5.2). Each locus-of-control question is asked on a 5-point Likert scale from “Strongly disagree” to “Strongly agree.”
- Life expectancy
- Youngest age that the respondent considers as “having lived a long life.”

Second, we estimate non-linear mixed effects regressions that allow for random coefficients. The models provide estimates of the correlation of the latent random coefficient for choices across different domains. Domains can be defined in different ways, for example, the first domain being vignette choices made in Wave 1 and the second domain being vignette choices made in Wave 2. Alternatively, one domain could be all vignette choices and the second domain could be all tuple choices.

We estimate random coefficient models by adjusting our baseline specification to allow for unobserved heterogeneity in the LYVE ($= \beta_1 / \beta_2$). Because the LYVE is positive for well-defined preferences, we model $\ln(LYVE) \equiv \ln(\phi) = \ln(\beta_1 / \beta_2)$ as having a latent normal distribution for choices in each domain.

Specifically, we impose that unobserved heterogeneity in the LYVE symmetrically affects the coefficients β_1 and β_2 , but in opposite directions:

- $\beta_1 \equiv \gamma_1 (\phi)^{1/2}$
- $\beta_2 \equiv \gamma_1 (\phi)^{-1/2}$
- $\ln(LYVE) = \ln(\phi) \sim N(\mu, \sigma^2)$

Using this random-coefficient specification, we will estimate the correlation between the latent person-specific LYVE across two domains for the following partitions of domains:

- Domain 1: vignette choices in Wave 1. Domain 2: vignette choices in Wave 2 by the same respondent. A high correlation would indicate that the person’s latent LYVE is stable over time.
- Domain 1: vignette choices in both waves. Domain 2: tuple choices in both waves. A high correlation would indicate that the latent LYVE is stable across different elicitation methods.
- Domain 1: vignette choices in both waves for the first, third, and fifth vignette pairs in a given wave. Domain 2: vignette choices in both waves for the second, fourth, and sixth choice pairs in a given wave. A high correlation indicates consistency of respondents’ latent LYVE within waves.
- (Placebo) Domain 1: vignette choices in Wave 1. Domain 2: vignette choices in Wave 2 by a randomly chosen (so generally different) respondent. Because the identity of the respondent in Wave 1 and Wave 2 is different, the correlation in the latent LYVE should be close to zero. We will randomly match Wave 2 respondents to Wave 1 respondents 100 times, and report results from the randomization that yielded a correlation between the two waves that was closest to the mean correlation across those 100 randomizations.

3.2.3 The Role of Context

Our main estimate of the valuation of longevity is based on choices over vignettes because we think vignettes help respondents in thinking about, and thus incorporating, pathways through which income and longevity contribute to lifetime utility. Moreover, choices over vignettes feel less artificial than choices over tuples because respondents in their day-to-day lives encounter stories or descriptions of people’s lives and may therefore be accustomed to evaluating other people’s lives. Hence, based on these a priori reasons, we believe choices over vignettes are more likely to reflect true preferences than choices over tuples.

We expect the LYVE based on tuple choices to be different than the LYVE based on vignette choices, and to be less informative about people’s true preferences. Still, the degree to which the LYVE varies between tuple choices and vignette choices is informative about the extent to which the context affects choices over income and longevity. Hence, as a first way to investigate the role of context, we will report the baseline specification estimated on choices over tuples.

All vignettes include information intended to make respondents think about ways in which income, or the lack of it, can affect utility. In addition, all vignettes include information intended to make respondents think about ways one can derive utility from non-material sources, such as companionship or appreciation of beauty. We then randomize the inclusion of additional sentences highlighting material sources of utility (“things money can buy”) or non-material sources of utility (“things money can’t buy”). Because there is already information about both sources of utility in each of the vignettes, additional information about these sources of utility should not meaningfully affect the estimated LYVE if we are accurately capturing preferences for longevity. In addition, we randomly include filler sentences to test whether the LYVE is sensitive to the length of the vignette. We test the effect of material-utility sentences, non-material-utility sentences, and filler sentences by adding the following nine terms to the baseline regression:

- $\beta_4 \text{MaterialSentences}_i * \Delta \ln(\text{Longevity}_{ci})$ where the variable *MaterialSentences* takes on the values -1, 0, or 1 depending on the number of added sentences that highlight material-utility. A value of 0 corresponds to a vignette with its baseline number of material-utility sentences (two). A value of -1 means the vignette has one less material-utility sentence, and a value of 1 means that the vignette has one more material-utility sentence.
- $\beta_5 \text{MaterialSentences}_i * \Delta \ln(\text{Income}_{ci})$
- $\beta_6 \text{MaterialSentences}_i$
- $\beta_7 \text{NonMaterialSentences}_i * \Delta \ln(\text{Longevity}_{ci})$ where the variable *NonMaterialSentences* takes on the values -1, 0, or 1 depending on the number of added sentences that highlight non-material-utility. A value of 0 corresponds to a vignette with its baseline number of non-material-utility sentences (two). A value of -1 means the vignette has one less non-material-utility sentence, and a value of 1 means that the vignette has one more non-material-utility sentence.
- $\beta_8 \text{NonMaterialSentences}_i * \Delta \ln(\text{Income}_{ci})$
- $\beta_9 \text{NonMaterialSentences}_i$
- $\beta_{10} \text{FillerSentences}_{ci} * \Delta \ln(\text{Longevity}_{ci})$ where the variable *FillerSentences* takes on the values 0 or 1 depending on whether a filler sentence was added to both vignettes in the pair, or neither.

- $\beta_{11} \text{FillerSentences}_{ci} * \Delta \ln(\text{Income}_{ci})$
- $\beta_{12} \text{FillerSentences}_{ci}$

We then will report $dLYVE/dMaterialSentences$, $dLYVE/dNonMaterialSentences$, and $dLYVE/dFillerSentences$, where each is calculated using the delta method.

3.3 Possible Bounds and Corrections to the LYVE

In this section, we consider four potential pathways through which our estimate of the LYVE may systematically deviate from the respondents' true preference for longevity. We propose tests to examine whether these pathways are operative, and propose corrections (which in one case is a bound) for pathways that are operative.

3.3.1 Possible Differential Attention to Longevity and Income

We aim to present vignettes that do not disproportionately draw attention to either income or longevity. We therefore expect that, if remembered perceived longevity or income differences are attenuated by the time the respondent makes a choice between the vignette, this attenuation is similar in proportional terms for longevity and income differences and the resulting LYVE will be unbiased because β_1 (in the numerator) and β_2 in the denominator will be attenuated by the same proportion.

To assess the degree to which respondents noticed, and remembered, income differences and longevity differences between vignette pairs, we ask respondents to rate their perception of the income difference and the longevity difference of the two vignettes in the last pair shown in a given survey wave on a 5-point Likert scale. We ask these ratings on a Likert scale rather than in terms of years (for longevity) or dollars (for income) because we want to measure how respondents perceived longevity or income differences intuitively rather than whether they were able to calculate or recall a specific figure. We ask respondents for these Likert ratings on a screen that does not show the text of the two vignettes. Hence, these answers reflect both the respondent's perception of the income and longevity differences as well as their recall of it just after they made a choice between the two vignettes.

We investigate how ratings of remembered longevity and income differences between two vignettes depend on longevity and income randomizations in the vignettes by running the following two regressions:

$$\text{LonLikertVig}_{ci} = \gamma_1^L \Delta \ln(\text{Longevity}_{ci}) + \gamma_2^L \Delta \ln(\text{Income}_{ci}) + \gamma_3^L \Delta \text{YoB}_{ci} + \delta_{v(A,c,i)}^L - \delta_{v(B,c,i)}^L + \epsilon_{ci}^L$$

$$\text{IncLikertVig}_{ci} = \gamma_1^I \Delta \ln(\text{Longevity}_{ci}) + \gamma_2^I \Delta \ln(\text{Income}_{ci}) + \gamma_3^I \Delta \text{YoB}_{ci} + \delta_{v(A,c,i)}^I - \delta_{v(B,c,i)}^I + \epsilon_{ci}^I$$

Elsewhere in the survey, we present respondents with a pair of tuples that differ in their income and longevity. *On the same screen*, we ask respondents to rate the longevity difference and income difference on the same 5-point Likert scale as used for the last vignette pair. Because tuples do not contain additional context (beyond the protagonists' names and years of birth), we can be confident that the respondents perceived the longevity and income differences equally well. Moreover, because the rating happens on the same screen, recall is not an issue. Hence, the questions on the

tuples serve to measure how respondents translate quantitative longevity and income differences into 5-point Likert scales. We measure how quantitative differences get transformed to differences on a Likert scale, by running the following two regressions:

$$\begin{aligned} LonLikertTup_{ci} &= \mu_1^L \Delta \ln(\widehat{Longevity}_{ci}) + \mu_2^L \Delta \ln(\widehat{Income}_{ci}) + \mu_3^L \Delta YoB_{ci} + \delta_{v(A,c,i)}^L - \delta_{v(B,c,i)}^L + \epsilon_{ci}^L \\ IncLikertTup_{ci} &= \mu_1^I \Delta \ln(\widehat{Longevity}_{ci}) + \mu_2^I \Delta \ln(\widehat{Income}_{ci}) + \mu_3^I \Delta YoB_{ci} + \delta_{v(A,c,i)}^I - \delta_{v(B,c,i)}^I + \epsilon_{ci}^I \end{aligned}$$

The hats above *longevity* and *income* indicate that we assume that respondents perceive longevity and income differences perfectly in the tuples presentation because there is no potentially distracting information and the information is shown on the same screen as the Likert question (hence, there is also no issue with recall).

If there are no cross effects (so $\gamma_2^L = \gamma_1^I = \mu_2^L = \mu_1^I = 0$), then the degree by which the latent longevity perception is attenuated relative to the randomized longevity manipulation in the vignette choice is given by γ_1^L / μ_1^L (one can think of this ratio as the first stage). The coefficient on *longevity* in the baseline regression (β_1) gives the effect of the randomized longevity manipulation on choice (so one can think of β_1 as a reduced form coefficient). Hence, the effect of the latent longevity perception on choice is given by $\beta_1 / (\gamma_1^L / \mu_1^L)$, i.e., it is given by β_1 divided by the degree of attenuation (one can think of this effect as the IV). The effect of interest, the effect of latent perceived longevity on vignette choice is given by $\beta_1 / (\gamma_1^L / \mu_1^L) = \beta_1 * \mu_1^L / \gamma_1^L$. Analogously, the effect of latent perceived income on vignette choice is given by $\beta_2 * \mu_2^I / \gamma_2^I$.

The LYVE, corrected for attenuation of longevity or income perceptions, is given by the ratio of the effects of interest:

$$LYVE_{Perception\ adjusted} = \frac{\beta_1 \mu_1^L / \gamma_1^L}{\beta_2 \mu_2^I / \gamma_2^I}$$

We interpret this corrected LYVE as the value the LYVE from the vignettes would have been if (i) respondents had had the identical perception and recall of income and longevity in the vignettes as they had in the tuples where they saw the income and longevity on the same screen as the subjective rating questions, and (ii) imperfect perception and recall is not endogenous to the respondent's underlying LYVE (e.g., they may not notice or recall income differences as much if they have a lower valuation of income). Hence, this calculation overstates the required true correction to the LYVE in case of endogenous perception and recall. We therefore think of it as a bound on the LYVE.

We will perform a joint test of the cross effects being zero (so $\gamma_2^L = \gamma_1^I = \mu_2^L = \mu_1^I = 0$). If we fail to reject this test at the 5% significance level, we implement the LYVE correction as described by the equation above. Otherwise, we calculate the analogous correction that allows for cross effects, as described below.

We first find the correspondence between latent perceptions and actual manipulations by setting $LonRatingVig_{ci} = LonRatingTup_{ci}$ and $IncRatingVig_{ci} = IncRatingTup_{ci}$, limiting ourselves to the causal effects of the longevity and income manipulations. This yields:

$$\gamma_1^L \Delta \ln(Longevity_{ci}) + \gamma_2^L \Delta \ln(Income_{ci}) = \mu_1^L \Delta \ln(\widehat{Longevity}_{ci}) + \mu_2^L \Delta \ln(\widehat{Income}_{ci})$$

and

$$\gamma_1^I \Delta \ln (\text{Longevity}_{ci}) + \gamma_2^I \Delta \ln (\text{Income}_{ci}) = \mu_1^I \Delta \ln (\widehat{\text{Longevity}}_{ci}) + \mu_2^I \Delta \ln (\widehat{\text{Income}}_{ci})$$

Next, we solve these two equations so that we get expressions of actual longevity manipulations in terms of latent perceptions of these manipulations (the variables with the hat). This yields:

$$\Delta \ln (\text{Longevity}_{ci}) = \frac{\mu_1^I - \mu_1^I * \gamma_2^I / \gamma_2^I}{\gamma_1^I - \gamma_1^I * \gamma_2^I / \gamma_2^I} \Delta \ln (\widehat{\text{Longevity}}_{ci}) + \frac{\mu_2^I - \mu_2^I * \gamma_2^I / \gamma_2^I}{\gamma_1^I - \gamma_1^I * \gamma_2^I / \gamma_2^I} \Delta \ln (\widehat{\text{Income}}_{ci})$$

and

$$\Delta \ln (\text{Income}_{ci}) = \frac{\mu_1^I - \mu_1^I * \gamma_1^I / \gamma_1^I}{\gamma_2^I - \gamma_2^I * \gamma_1^I / \gamma_1^I} \Delta \ln (\widehat{\text{Longevity}}_{ci}) + \frac{\mu_2^I - \mu_2^I * \gamma_1^I / \gamma_1^I}{\gamma_2^I - \gamma_2^I * \gamma_1^I / \gamma_1^I} \Delta \ln (\widehat{\text{Income}}_{ci})$$

Next, we substitute these two expressions into the estimating equation for vignette choice in the baseline regression (suppressing the terms not related to longevity and income for readability):

$$\text{Chose}A_{ci} = \beta_1 \Delta \ln (\text{Longevity}_{ci}) + \beta_2 \Delta \ln (\text{Income}_{ci}) + \text{other terms}$$

This yields:

$$\begin{aligned} \text{Chose}A_{ci} = & \frac{\beta_1(\mu_1^I - \mu_1^I * \gamma_2^I / \gamma_2^I)}{\gamma_1^I - \gamma_1^I * \gamma_2^I / \gamma_2^I} \Delta \ln (\widehat{\text{Longevity}}_{ci}) + \frac{\beta_1(\mu_2^I - \mu_2^I * \gamma_2^I / \gamma_2^I)}{\gamma_1^I - \gamma_1^I * \gamma_2^I / \gamma_2^I} \Delta \ln (\widehat{\text{Income}}_{ci}) + \\ & \frac{\beta_2(\mu_1^I - \mu_1^I * \gamma_1^I / \gamma_1^I)}{\gamma_2^I - \gamma_2^I * \gamma_1^I / \gamma_1^I} \Delta \ln (\widehat{\text{Longevity}}_{ci}) + \frac{\beta_2(\mu_2^I - \mu_2^I * \gamma_1^I / \gamma_1^I)}{\gamma_2^I - \gamma_2^I * \gamma_1^I / \gamma_1^I} \Delta \ln (\widehat{\text{Income}}_{ci}) + \\ & \text{other terms} \end{aligned}$$

Collecting terms yields:

$$\begin{aligned} \text{Chose}A_{ci} = & \left(\frac{\beta_1(\mu_1^I - \mu_1^I * \gamma_2^I / \gamma_2^I)}{\gamma_1^I - \gamma_1^I * \gamma_2^I / \gamma_2^I} + \frac{\beta_2(\mu_1^I - \mu_1^I * \gamma_1^I / \gamma_1^I)}{\gamma_2^I - \gamma_2^I * \gamma_1^I / \gamma_1^I} \right) \Delta \ln (\widehat{\text{Longevity}}_{ci}) + \\ & \left(\frac{\beta_1(\mu_2^I - \mu_2^I * \gamma_2^I / \gamma_2^I)}{\gamma_1^I - \gamma_1^I * \gamma_2^I / \gamma_2^I} + \frac{\beta_2(\mu_2^I - \mu_2^I * \gamma_1^I / \gamma_1^I)}{\gamma_2^I - \gamma_2^I * \gamma_1^I / \gamma_1^I} \right) \Delta \ln (\widehat{\text{Income}}_{ci}) + \text{other terms} \end{aligned}$$

The LYVE, corrected for attenuation of longevity or income perceptions, is given by the ratio of the effects of interest, which in turn are given by the expressions in front of $\Delta \ln (\widehat{\text{Longevity}}_{ci})$ and $\Delta \ln (\widehat{\text{Income}}_{ci})$, respectively:

$$\text{LYVE}_{\text{Perception adjusted}} = \frac{\frac{\beta_1(\mu_1^I - \mu_1^I * \gamma_2^I / \gamma_2^I)}{\gamma_1^I - \gamma_1^I * \gamma_2^I / \gamma_2^I} + \frac{\beta_2(\mu_1^I - \mu_1^I * \gamma_1^I / \gamma_1^I)}{\gamma_2^I - \gamma_2^I * \gamma_1^I / \gamma_1^I}}{\frac{\beta_1(\mu_2^I - \mu_2^I * \gamma_2^I / \gamma_2^I)}{\gamma_1^I - \gamma_1^I * \gamma_2^I / \gamma_2^I} + \frac{\beta_2(\mu_2^I - \mu_2^I * \gamma_1^I / \gamma_1^I)}{\gamma_2^I - \gamma_2^I * \gamma_1^I / \gamma_1^I}}$$

Note that if the cross effects are zero (so $\gamma_2^L = \gamma_1^I = \mu_2^L = \mu_1^I = 0$), this expression reduces to:

$$LYVE_{Perception\ adjusted} = \frac{\beta_1 \mu_1^L / \gamma_1^L}{\beta_2 \mu_2^L / \gamma_2^L}$$

3.3.2 Possible Effect of Income on Perceived Disutility of Labor

Income in a given vignette is randomized while the description of the protagonist's occupation in a given vignette is fixed, thus holding constant any compensating differentials associated with fixed characteristics of the occupation. However, a higher income within a given occupation may, in the perception of respondents, be associated with changes in work amenities of the protagonist, and thereby affect the respondent's perception of the compensating differential for that protagonist's work. The direction of this compensating differential is not clear a priori. For example, higher earnings may signal that the protagonist works for a firm that compensates their employees highly both in wages and non-wage benefits. In that case, the value of the higher earnings exceeds the value of a cash transfer in the same amount. Alternatively, higher earnings may signal that the protagonist experiences onerous or stressful work conditions, in which case the value of the higher earnings is less than the value of a cash transfer of the same amount.

To test whether variation in income for a given vignette affects the respondent's perceived disutility of the work of the vignette protagonist, we include a question about perceived work disamenities after some of the vignette pairs. Specifically, we ask respondents to compare the work disamenities (which we phrase as the job involving "stress, worries, irregular hours, or other hardships") of the two jobs in the vignette pair. We also randomized whether a sentence describing work disamenities was added to one of the vignettes in a vignette pair. Thus, any given vignette could be shown with or without this sentence, which was tailored to each vignette (rather than identical for all vignettes). This sentence serves two purposes. First, it allows us to test whether respondents perceive work disamenities. Second, as explained below, we need it if we want to correct for any effect of income on perceived work disamenities.

Next, we run the following two regressions in a SUR framework. First, we re-run the baseline regression, augmented with $\Delta WorkDisamenity_{ci}$, but instead have as the dependent variable the respondent's subjective rating (on a 3-point scale) of whether the job in vignette A had more work disamenities, the jobs had about equal work disamenities, or the job in vignette B had more work disamenities. The variable $\Delta WorkDisamenity_{ci}$ is the difference between the work disamenity indicator variables between vignette A and vignette B that correspond to the extra work disamenity sentence.

$$DisamenityRating_{ci} = \gamma_1 \Delta \ln (Longevity_{ci}) + \gamma_2 \Delta \ln (Income_{ci}) + \gamma_3 \Delta YearOfBirth_{ci} + \gamma_4 \Delta WorkDisamenity_{ci} + \delta_{v(A,c,i)} - \delta_{v(B,c,i)} + \epsilon_{ci}$$

In this regression, γ_2 measures the impact of income, conditional on occupation, on perceived work disamenities. The LYVE should express the value of longevity relative to a change in income from a cash transfer rather than from a change of income that could potentially have associated

amenities or disamenities. If γ_2 is then statistically significant at the 5% level (so apparently income is perceived to affect work amenities or disamenities), we calculate the following adjustment to the LYVE for the effect of income on perceptions of work amenities or disamenities.

Second, we rerun our baseline regression but include among the right-hand side variables an indicator variable for whether a work disamenity sentence was added. Hence, we add the following term to the baseline regression: $\beta_4 \Delta \text{WorkDisamenity}_{ci}$.

Finally, we calculate the LYVE adjusted for any perceived changes in work disamenities as:

$$LYVE_{Disamenity\ Adjusted} = \frac{\beta_1}{\beta_2 - \beta_4 * \gamma_2 / \gamma_4}$$

This adjustment is based on the following reasoning. The effect of income on the perception of work disamenities is γ_2 and the effect of the disamenity sentence is γ_4 . Hence, the effect of income on perceived work disamenities is as large as γ_2 / γ_4 times the effect of the disamenity sentence. The effect of the disamenity sentence on the likelihood of choosing the vignette is given by β_4 . Hence, the effect of income, that runs via disamenity perceptions, on choosing the vignette is $\beta_4 * \gamma_2 / \gamma_4$. This effect is subtracted out from the total effect of income on choosing the vignette (β_2) to obtain the “pure” effect of income, i.e., net of any effects running through the perceived disamenity channel.

3.3.3 Possible Effect of Life Satisfaction of the Protagonist in the Vignettes

We use a broad set of vignettes with the aim of describing a range of completed life stories that broadly represent experienced lives in the US. However, if these stories represent lives that lead to especially high life satisfaction, respondents may value additional longevity more than for a typical life in the US. To address this possible concern, we ask respondents to rate one randomly selected vignette that they have not seen in that wave as part of the vignette pair choices. We refer to this vignette as the “rating vignette.” We ask respondents to rate how they perceived the life satisfaction of the protagonist in the vignette when the protagonist had a specified age. The wording of the life satisfaction question, as well as the response scale, corresponds to the wording of the life satisfaction question in the Behavioral Risk Factor Surveillance System (BRFSS). This allows us to compare the perceived life satisfaction of the protagonists in the vignettes to life satisfaction in the US population.

Specifically, we calculate mean life satisfaction in the BRFSS for a sample that is matched in terms of age to the distribution of vignette protagonists’ ages at which the respondents rated the protagonists’ life satisfaction. We calculate this age-adjusted mean life satisfaction using the untransformed 4-point rating scale for life satisfaction as defined by the BRFSS. We define Δu as the difference between the age-adjusted mean life satisfaction in the BRFSS and the perceived mean life satisfaction of the protagonists in the vignettes.

If there is a difference, significant at the 5% level, between the mean age-adjusted life-satisfaction in the BRFSS and the mean life satisfaction that respondents perceived the vignette protagonists to have, we adjust the LYVE as follows for this life satisfaction difference.

First, we run the baseline regression on the rating vignette using the perceived life satisfaction

of the protagonist as the outcome variable, and adding dummies for the decade of age of the protagonist at which their life satisfaction was evaluated. We take the vignette fixed effects of this regression (including a value of zero for the omitted fixed effect) and demean the fixed effects. This gives us each vignette's deviation from mean life satisfaction. We denote the demeaned fixed effect for vignette v by $\bar{u}_v$.

Next, we investigate the effect of a vignette's deviation from mean life satisfaction on the LYVE. We do so by running a variant of the baseline regression in which we interact both longevity and income with the vignette's deviation from mean life satisfaction as measured by the demeaned vignette dummies from the perceived life satisfaction regression. Specifically, we run:

$$\begin{aligned} ChoseA_{ci} = & \beta_1 \Delta \ln(Longevity_{ci}) + \beta_2 \Delta \ln(Income_{ci}) + \beta_3 \Delta YearOfBirth_{ci} + \\ & \beta_4 (\bar{u}_{v(A,c,i)} * \ln(Longevity_{Aci}) - \bar{u}_{v(B,c,i)} * \ln(Longevity_{Bci})) + \\ & \beta_5 (\bar{u}_{v(A,c,i)} * \ln(Income_{Aci}) - \bar{u}_{v(B,c,i)} * \ln(Income_{Bci})) + \\ & \delta_{v(A,c,i)} - \delta_{v(B,c,i)} + \epsilon_{ci} \end{aligned}$$

Note that the direct effects of $\bar{u}_{v(A,c,i)}$ and $\bar{u}_{v(B,c,i)}$ are absorbed by the fixed effects δ . Next, we calculate $dLYVE/d\bar{u}$ using the delta method.

We calculate the LYVE adjusted to the mean life satisfaction in the US as follows:

$$LYVE_{LSAdjusted} = \frac{\beta_1}{\beta_2} + \Delta u * \frac{dLYVE}{d\bar{u}}$$

3.3.4 Possible Perception of Ill Health Prior To Final Years Before Death

We want the LYVE to reflect the value respondents place on additional years in good health. The experiment aims to achieve this by explicitly stating any illness, and its duration, that ultimately causes death. The duration of this illness is fixed, and hence its starting age moves up in lock-step with longevity, which is randomized. As a result, the additional years of life that are added when longevity is randomized to be longer are additional years prior to this final illness, which presumably are perceived to be in good health.

To test the assumption that respondents perceive no ill health prior to the onset of the final disease, we randomize whether we insert a sentence explicitly stating that health was good until the onset of the final disease. If respondents already assumed health was good, this sentence should have no effect on vignette choice. We test this prediction by running the baseline regression to which we add an indicator variable for the good-health sentence. If this indicator is not statistically significant at the 5% level, there is no evidence of perceptions of ill health prior to the onset of the final disease affecting vignette choice.

If the indicator is significant at the 5% level, we calculate an adjustment as follows. We run a variant of the baseline regression in which we interact both longevity and income with an indicator of whether the vignette has the good-health sentence. Specifically, we run:

$$\begin{aligned}
ChoseA_{ci} = & \beta_1 \Delta \ln(Longevity_{ci}) + \beta_2 \Delta \ln(Income_{ci}) + \beta_3 \Delta YearOfBirth_{ci} + \\
& \beta_4 (Health_{Aci} * \ln(Longevity_{Aci}) - Health_{Bci} * \ln(Longevity_{Bci})) + \\
& \beta_5 (Health_{Aci} * \ln(Income_{Aci}) - Health_{Bci} * \ln(Income_{Bci})) + \\
& \beta_6 Health_{v(A,c,i)} - \beta_7 Health_{v(B,c,i)} + \delta_{v(A,c,i)} - \delta_{v(B,c,i)} + \epsilon_{ci}
\end{aligned}$$

where $Health_{Aci}$ is an indicator variable for the good-health sentence being included for vignette A of the c^{th} vignette pair seen by respondent i and $Health_{Bci}$ is an indicator variable for the good-health sentence being included for vignette B of the c^{th} vignette pair seen by respondent i .

Next, we calculate $dLYVE/dHealth$ using the delta method.

We calculate the LYVE adjusted so that health prior to the onset of the final disease is defined as good:

$$LYVE_{Health\ Adjusted} = \frac{\beta_1}{\beta_2} + \frac{dLYVE}{dHealth}$$

3.4 More Flexible Functional Form

The baseline analysis uses an OLS specification on a Cobb-Douglas functional form (in log form) because (i) we observed that this fits the data reasonably well in the pilots, (ii) it is simple and transparent, and (iii) it allows preferences over lifetime income and longevity to be described by a single parameter, the LYVE, which does not vary by income or longevity.

However, we investigate the extent to which the baseline functional form fits the data by estimating the following more flexible functional form, using non-linear least squares. We will use the Schwartz criterion to evaluate goodness of fit in order to take into account varying numbers of parameters of the considered utility functions. This specification is identical to the baseline specification, except for two changes.

First, we replace $\ln(Longevity_{ci})$ by $(Longevity_{ci}^{1-\gamma_1} - 1)/(1 - \gamma_1)$. Second, we replace $\ln(Income_{ci})$ by $(Income_{ci}^{1-\gamma_2} - 1)/(1 - \gamma_2)$.

This more flexible functional form nests the following commonly used functional forms:

- Cobb-Douglas in log form, for $\gamma_1 = \gamma_2 = 1$
- Linear, for $\gamma_1 = \gamma_2 = 0$
- Quasilinear, with a constant valuation of each additional year of life, for $\gamma_1 = 0$
- Quasilogarithmic, with a valuation of an additional year of life that is proportional to the inverse of longevity, for $\gamma_1 = 1$

We compare the model fit of the more flexible functional form and the four more restricted functional forms that it nests (as listed above) using the Schwarz Information Criterion.

For this more flexible functional form, we report two values of the LYVE. First, we report the LYVE evaluated at the median income and the median longevity of the vignettes shown to

respondents. Second, we calculate the LYVE at all income and longevity combinations that were shown in vignettes to the respondents, and then report the mean of these LYVEs.

We assess the fit of our baseline functional form by reporting the LYVE, both for our baseline specification and for the flexible functional form, at five combinations of longevity and income:

- 25th percentile of longevity and 25th percentile of income
- 25th percentile of longevity and 75th percentile of income
- 50th percentile of longevity and 50th percentile of income
- 75th percentile of longevity and 25th percentile of income
- 75th percentile of longevity and 75th percentile of income

In addition, we will plot the indifference curve defined by the baseline specification and by the more flexible functional form in each of the four quadrants separated by 50th percentile of longevity and 50th percentile of income. In each of these four quadrants, we plot the two indifference curves going through the relevant combination of 25th or 75th percentile of income and longevity for that quadrant.

3.5 Heterogeneous Effects

We use four approaches to analyze heterogeneity in the valuation of longevity.

The first approach examines heterogeneity in the LYVE separately for each single dimension of interest. The dimensions can be categorical respondent characteristics, univariate respondent characteristics, or vignette characteristics (other than income, longevity, or year of birth) possibly interacted with respondent characteristics. This approach is followed in subsections 3.5.1, 3.5.2, and 3.5.3. The second approach examines unobserved heterogeneity in the LYVE by estimating a random coefficients model. This approach is followed in subsection 3.5.4. The third approach examines heterogeneity in the dollar value of people's last life year. This approach is followed in subsection 3.5.5. The fourth approach estimates a random utility model that jointly examines multiple measured dimensions of heterogeneity and combines that with unobserved heterogeneity. We refer to this approach as the "structural analysis" and describe it in subsection 3.9.

3.5.1 Heterogeneity by Categorical Respondent Characteristics

For categorical characteristics, we define a set of categorical variables for each category of respondents within a given dimension. We estimate heterogeneity for a single dimension at a time by adding to the baseline regression the following three terms for all categorical variables of that dimension, other than for the omitted category:

- $\beta_{1dk} \text{Category}_{dki} * \Delta \ln(\text{Longevity}_{ci})$
- $\beta_{2dk} \text{Category}_{dki} * \Delta \ln(\text{Income}_{ci})$
- $\beta_{3dk} \text{Category}_{dki}$

The variable $Category_{dki}$ is the categorical variable that indicates whether respondent i is part of category k in dimension d . For example, if the dimension is “age”, we have five categorical variables for the five age categories defined below. Hence, we add a total of 12 terms to the baseline regression, three variables each for the four non-omitted age categories.

For each dimension, we report the p-value on the test that the LYVE is equal for all categories of respondents in that dimension.

We examine heterogeneity with respect to the following dimensions:

- Age (Under 40; 40-49; 50-59; 60-69; 70 and over)
- Gender (Male; Female)
- Income (Under \$40K; \$40K-\$75K; \$75K-\$100K; \$100K-\$150K; over \$150K)
- Race/Ethnicity (Only White; White and other race; Black; Hispanic; Other)
- Education (HS or less; Some college; Bachelor’s and up)
- Employment status (Employed; Unemployed; Other)
- Political Affiliation (Democrat; Republican; Other)
- Urbanicity (Urban; Rural)
- Marital status (Married; Unmarried)
- Having had children (No children; Any children)
- Having had grandchildren (No grandchildren; Any grandchildren)
- Desire to have any children (Ideal number of children > 0 ; Ideal number of children = 0)
- Religion (Protestant and Mormon; Catholic and Orthodox; Atheist, Agnostic, and no religion; All else)
- Belief in afterlife (Yes; No)

3.5.2 Heterogeneity by Univariate Respondent Characteristics

For univariate characteristics, we define a single variable for each dimension. We estimate heterogeneity for a single dimension at a time by adding to the baseline regression the following three terms:

- $\beta_{1d} Characteristic_{di} * \Delta \ln (Longevity_{ci})$
- $\beta_{2d} Characteristic_{di} * \Delta \ln (Income_{ci})$
- $\beta_{3d} Characteristic_{di}$

where $Characteristic_{di}$ is the univariate characteristic associated with dimension d for respondent i .

We examine heterogeneity with respect to the following dimensions:

- Age (demeaned)
- Life expectancy (demeaned)
- Youngest age that the respondent considers as “having lived a long life” (demeaned)
- Income (log, demeaned)

The income buckets are “less than \$9,999”, “\$10,000-\$19,999”, “\$20,000-\$29,999”, “\$30,000-\$39,999”, “\$40,000-\$49,999”, “\$50,000-\$59,999”, “\$60,000-\$74,999”, “\$75,000-\$99,999”, “\$100,000-\$149,999”, “\$150,000-\$199,999”, “\$200,000 or more”, and “prefer not to say.” The continuous income measure is the midpoint of the range of income buckets. For example, if the respondent chose the “\$50,000-\$59,999” option, we would assign a numerical value of \$55,000 to the respondent. The “\$200,000 or more” category is assigned a numerical value of \$400,000. We exclude participants that choose the “prefer not to say” category from this analysis.

- Years of education (demeaned)

Years of education correspond to education level categories where “less than high school” and “GED” correspond to 10 years. “High school diploma” = 12 years. “Some college but no degree” and “Associate’s degree from a 2-year college” = 14 years. “Bachelor’s degree from a 4-year college” = 16 years. “Master’s degree, Professional degree (MD, JD, etc.)” = 18 years. “PhD” = 20 years.

- Self-reported grades in last year of high school (standardized)

The seven categories are “All As”, “Mostly As with a few Bs”, “A mix of As and Bs”, “Mostly Bs”, “A mix of Bs and Cs”, “Mostly Cs”, “Mostly Cs and Ds”, and “prefer not to say.” Numerical values increase as grades improve where “All As” has a value of 7 and “Mostly Cs and Ds” has a value of 1.

- Self-reported optimism (1 to 5 scale from “Not at all optimistic” to “Very optimistic”, standardized)
- Imaginativeness

The Imaginativeness Index is comprised of the following two questions. Each question is asked on a 6-point Likert scale from “Very inaccurate” to “Very accurate”. The index is the standardized mean of the two questions.

- *I have difficulty picturing the details of a situation I have not previously experienced.* (Reverse coded)
- *The things I imagine are not very detailed.* (Reverse coded)

- Life satisfaction (0 to 10 scale from “Not satisfied at all” to “Extremely satisfied”, standardized)
- Self-reported health (0 to 10 scale from “Very poor health” to “Very good health”, standardized)
- Locus of Control index

The Locus of Control index is comprised of the following seven questions. Each question is asked on a 5-point Likert scale from “Strongly disagree” to “Strongly agree”. The index is the standardized mean of all seven questions.

- *I have little control over the things that happen to me.* (Reverse coded)
- *There is really no way I can solve some of the problems I have.* (Reverse coded)
- *There is little I can do to change many of the important things in my life.* (Reverse coded)
- *I often feel helpless in dealing with the problems of life.* (Reverse coded)
- *Sometimes I feel that I am being pushed around in life.* (Reverse coded)
- *What happens to me in the future mostly depends on me.*
- *I can do just about anything I really set my mind to do.*
- Willingness to give to good causes without expecting anything in return (altruism) (0 to 10 scale from “Completely unwilling to do so” to “Very willing to do so”, standardized)
- Willingness to give up something that is beneficial today in order to benefit more in the future (patience) (0 to 10 scale from “Completely unwilling to do so” to “Very willing to do so”, standardized)
- Willingness to take risks (0 to 10 scale from “Completely unwilling to do so” to “Very willing to do so”, standardized, reverse coded)
- Others’ intentions (trust): *I assume that people have only the best intentions.* (0 to 10 scale from “Does not describe me at all” to “Describes me perfectly”, standardized)
- Present bias: *I tend to postpone things even though it would be better to take care of them right away.* (0 to 10 scale from “Does not describe me at all” to “Describes me perfectly”, standardized)
- The percent of comprehension checks answered correctly by the respondent

3.5.3 Heterogeneity Based on Vignette Characteristics (possibly interacted with respondent characteristics)

First, we examine how the LYVE varies by characteristics of the c^{th} vignette *pair* seen by respondent i .

We use the specification for categorical variables defined in subsection 3.5.1. We examine heterogeneity by the following categorical dimensions of vignette *pairs*:

- Marital status of the vignette protagonists (Both married; Both unmarried; One married and one unmarried)
- Whether the vignette protagonists had children (Both had children; Neither had children; One had children and the other did not)
- Gender of vignette protagonists (Both male; Both female)

Second, we examine how the LYVE varies by univariate characteristics of a *single* vignette. Let the characteristic of dimension d of vignette A of the c^{th} pair seen by respondent i be denoted by $Char_{Adci}$, and $Char_{Bdci}$ is defined analogously. We then estimate:

$$\begin{aligned}
ChoseA_{ci} = & \beta_1 \Delta \ln(Longevity_{ci}) + \beta_2 \Delta \ln(Income_{ci}) + \beta_3 \Delta YearOfBirth_{ci} + \\
& \beta_4 (Char_{Adci} * \ln(Longevity_{Aci}) - Char_{Bdci} * \ln(Longevity_{Bci})) + \\
& \beta_5 (Char_{Adci} * \ln(Income_{Aci}) - Char_{Bdci} * \ln(Income_{Bci})) + \\
& \beta_6 Char_{Adci} - \beta_7 Char_{Bdci} + \delta_{v(A,c,i)} - \delta_{v(B,c,i)} + \epsilon_{ci}
\end{aligned}$$

We examine heterogeneity by the following univariate dimensions of a *single* vignette:

- Marital status of vignette protagonist (1=Married; 0=Unmarried)
- Gender of the vignette protagonist (1=Female; 0=Male)
- Whether the vignette protagonists had children (1=Had children; 0=Did not have children)
- The desirability of the vignette as measured by the vignette's fixed effect (standardized)
 - We use the estimated fixed effects from the baseline regression, and first demean these fixed effects separately for the fixed effects of vignettes with female protagonists and the fixed effects of vignettes with male protagonists. Next, we standardize the combined set of demeaned fixed effects. We need to do the demeaning separately by vignette protagonist gender because the fixed effects for female protagonists are not directly comparable to those for male protagonists since respondents never make a choice between a vignette with a female and a male protagonist. Hence, the average values of fixed effects for vignettes for male protagonists can have an arbitrary level difference with the average values of the fixed effects for the vignettes with female protagonists because the choice of omitted fixed effect is arbitrary for both types of vignettes. (The regression will omit exactly one fixed effect for a vignette with a female protagonist and exactly one fixed effect for a vignette with a male protagonist). The demeaning removes this arbitrariness and sets the mean values of the fixed effects equal to zero for the two sets of fixed effects. The subsequent standardization is only for interpretability purposes.

Finally, we examine whether the LYVE varies by measures of similarity between the respondent and the characteristics of a single vignette. These similarity measures are univariate and vignette-specific. We use the same specification as mentioned above, replacing $Char_{Adci}$ by $Similarity_{Adci}$, and similarly for vignette B.

- The respondent and the vignette protagonist share the same gender (1=Yes; 0=No)
- The respondent and the vignette protagonist both are married or both are unmarried (1=Yes; 0=No)
- The respondent and the vignette protagonist both had kids or both had no kids (1=Yes; 0=No)
- The respondent and the vignette protagonist both had kids or both had no kids (1=Yes; 0=No), where respondents who answer that they wish to have kids are counted as having had kids (some respondents may not yet have reached the stage in life when they'll have kids).

- The respondent and the vignette protagonist both have at least a college degree or neither has at least a college degree (1=Yes; 0=No)
- A predicted relatability index (standardized), which will be constructed as follows:
 - We run an OLS regression of each respondent’s relatability to the rating vignette on the set of explanatory variables selected by LASSO out of the set of potential explanatory variables listed below. Specifically, we measure relatability to the rating vignette as the respondent’s answer to the question *How similar are your experiences to those of [protagonist’s name]?*, where answer choices are measured on a 5-point Likert scale.
 - We use LASSO using cross-validation with 10 folds to select the penalty parameter (λ).
 - Using the estimated OLS model, as defined above, we predict the relatability of each respondent for each vignette they see. We then standardize this predicted relatability measure.
 - The LASSO procedure will select explanatory variables from the following set of variables:
 - * A constant term
 - * The respondent’s Imaginativeness Index (standardized index defined above based on two underlying questions)
 - * Dummy variables for whether the respondent and the vignette protagonist share the same: sex, marital status (married / not married), having kids (yes/no), predicted to have kids (yes/no)(here respondents who answer that they wish to have kids are counted as predicted to have kids because some respondents may not yet have reached the stage in life when they’ll have kids), having grandchildren (yes/no), educational status (at least college degree/not), and occupation category (same first digit of the International Standard Classification of Occupations (ISCO) occupational categories/not).
 - * A two-piece linear function, with a slope break and a discontinuous jump at zero, of the difference between an underlying continuous variable for the respondent and the corresponding underlying continuous variable for the protagonist. A special case of this spline is the absolute difference between the two underlying continuous variables (if the slopes are equal but of opposite sign and the discontinuous jump is zero). Specifically, the two-piece linear function is implemented by including three regression variables, defined below, for each underlying continuous variable. Let Z be the difference between an underlying continuous variable for the respondent and the corresponding underlying continuous variable for the protagonist, and let D be an indicator that is 1 if and only if $Z > 0$, and is 0 otherwise. Then the three regression variables are defined as: $X_1 = Z * D$, $X_2 = Z * (1 - D)$, and $X_3 = D$. These three regression variables are created for each of the following underlying continuous variables: year of birth, self-reported life expectancy (vs. protagonist age of death), youngest age that the respondent considers as “having lived a long life” (vs. protagonist age of death), respondent age (vs. protagonist age of death), $\ln(\text{income})$ (as specified above based on the categorical income variable collected from respondents), and life satisfaction (standardized). For the protagonist’s life satisfaction, we use the average life satisfaction Likert-scale rating from all respondents who rated the life satisfaction of that protagonist, and then standardize it across the 28 vignettes.

- * We also include interactions of all of the variables specified in the four bullet points above with the following variables describing the respondent: sex, age, marital status (married / not married), having kids (yes/no), wanting to have kids (yes/no), $\ln(\text{income})$, and Imaginativeness Index.
- * We standardize all variables listed above (other than the constant term) because LASSO is not scale-independent.
- * We set any missing variables equal to zero and include an indicator variable for each variable that indicates whether a missing value was replaced by a zero.

3.5.4 Unobserved Heterogeneity in Preferences

If all respondents saw the exact same vignette pairs, misspecification of the functional form would not translate into unobserved heterogeneity in preferences. However, given the randomization of vignette pairs to respondents, and randomizations of income and longevity within vignettes, there is variation across respondents in the location in income-longevity space of the vignettes they consider. Hence, misspecification of the functional form in income-longevity space could show up as unobserved heterogeneity in respondent preferences.

To limit the role of functional form misspecification on unobserved heterogeneity estimates, we use the functional form that is preferred by the Schwarz Information Criterion in subsection 3.4, treating the functional form parameters of that specification as fixed (rather than as to be re-estimated). Using this functional form, we estimate a non-linear mixed effects model in which there is one dimension of unobserved heterogeneity in the preference parameters. As in subsection 3.2.2, we impose that unobserved preference heterogeneity symmetrically affects the coefficients β_1 and β_2 , but in opposite directions:

- $\beta_1 \equiv \gamma_1 (\phi)^{1/2}$
- $\beta_2 \equiv \gamma_1 (\phi)^{-1/2}$
- $\ln(\phi) \sim N(\mu, \sigma^2)$

Because the LYVE varies by the income and longevity of the vignette if the functional form is not Cobb-Douglas, we need to hold income and longevity at which the LYVE is evaluated constant to show heterogeneity that is solely due to heterogeneity in preferences. As in subsection 3.2.2, we use two ways of evaluating the LYVE at a constant (set of) values of income and longevity. First, we report the LYVE evaluated at the median income and the median longevity of the vignettes shown to respondents. Second, we calculate the LYVE at all income and longevity combinations that were shown in vignettes to the respondents, and then report the mean of these LYVEs (so this is a single value for any given preference). Next, we show the heterogeneity of these two LYVE values due to heterogeneity in preferences, as estimated by the distribution of the ϕ parameter.

3.5.5 Heterogeneity in the Dollar Value of an Additional Life Year

Many policy decisions are based on the dollar value per life year. In our setting, we estimate the dollar value of the marginal life year by the Life Year Value (LYV). The LYV is the change in income

that respondents value as much as a one-year increase in longevity, and is formally defined by:

$$LYV \equiv \frac{\partial ChoseA / \partial Longevity_A}{\partial ChoseA / \partial Income_A},$$

where, in slight abuse of notation, $Longevity_A$ denotes the longevity of the protagonist in vignette A and $Income_A$ denotes the income of the protagonist in vignette A.

There are two sources of heterogeneity in the LYV in a population. First, other than for linear functional forms (which we predict fit poorly), the LYV varies by the income level and longevity at which it is evaluated. Thus, heterogeneity in income and longevity in the population leads to heterogeneity in LYVs. Second, there is heterogeneity in the population in their preferences, even if evaluated at the same income and longevity combination. We show the contribution of both sources to the distribution of LYVs as follows.

We first isolate the contribution of heterogeneity in income and longevity in the population. To do so, we take the population distribution of lifetime incomes from [Guvenen et al. \(2022\)](#). We take the distribution of longevity from the [U.S Centers for Disease Control and Prevention \(2020\)](#). We assume (as an approximation) that these distributions are independent and evaluate the LYV at the resulting joint distribution of life-time incomes and longevity using the fixed parameter estimates from the functional form that is preferred by the Schwarz Information Criterion in subsection 3.4. We then plot the resulting distribution of LYVs.

Second, we use the same joint distribution of lifetime incomes and longevity as above but, in addition, allow for heterogeneity in preferences as estimated in the previous subsection. Assuming that preference heterogeneity is independent of lifetime income and longevity, we take independent draws from the distribution of ϕ and from the joint distribution of lifetime incomes and longevity, and calculate the LYV for each pair of draws. We then plot the resulting distribution of LYVs.

3.6 Specification Checks

We plan to study the robustness of our analysis along several dimensions, including econometric specifications.

3.6.1 Testing Sensitivity to Linear Probability Models

The baseline specification is a linear probability model. We use a linear probability model for transparency and ease of estimation. To check that this specification does not meaningfully affect the estimate of the LYVE, we also estimate the baseline model for the following alternative regression specifications:

- OLS with vignette-pair fixed effects, rather than vignette fixed effects. We do so as a comparison for the conditional logit below because conditional logit can only use a single set of fixed effects (which is the case for vignette-pair fixed effects, but not for vignette fixed effects, which are two-way fixed effects).
- Conditional logit with vignette-pair fixed effects (i.e., the group or conditioning variable is the set of vignette-pair fixed effects).

- Regular logit with vignette-pair fixed effects included as right-hand side variables. Differences in this estimate and the estimate from the conditional logit would be indicative of an incidental parameter problem.
- Regular logit with vignette fixed effects included as right-hand side variables. This specification would only merit consideration if there is no incidental parameter problem.
- Probit with vignette fixed effects included as right-hand side variables. This specification would only merit consideration if there is no incidental parameter problem.

3.6.2 Testing Possible Effects of Vignette Pair Imbalance

We investigate whether the effect of longevity and the effect of income on vignette choice is differentially affected if one vignette is more attractive based on the context of the story than the other vignette. Such differential effects of imbalance between vignette attractiveness would affect the estimate of the LYVE. We measure the degree to which one vignette is more attractive than the other based on the context of the vignette's story by a variable called $Imbalance_{ci}$, which is defined as the absolute value of the difference between the value of the vignette fixed effect for vignette A and the value of the vignette fixed effect for vignette B, as estimated in the baseline regression. We standardize $Imbalance_{ci}$ for ease of interpretation. We test the effect of the vignette imbalance on the LYVE by including the following interaction terms to the baseline regression:

- $\beta_4 Imbalance_{ci} * \Delta \ln (Longevity_{ci})$
- $\beta_5 Imbalance_{ci} * \Delta \ln (Income_{ci})$
- $\beta_6 Imbalance_{ci}$

We calculate the derivative of the LYVE with respect to the imbalance variable using the delta method. We also estimate the LYVE in separate regressions, splitting the sample by quartiles of $Imbalance_{ci}$, and test whether we can reject that the LYVE is the same in those four subsamples.

3.7 Survey Response Patterns

We conduct a number of analyses to better understand how survey respondents approached the survey. The key question surrounding this analysis is the following: To what degree do respondents' choices reflect their underlying preferences as opposed to test-taking strategies to economize on time or effort to complete the survey?

If the latter is important (i.e, using heuristics to finish the survey), we might expect order effects in the estimated LYVE, with later vignettes answered differently than earlier ones. For example, respondents might develop rules of thumb for approaching the survey or exert less effort on later questions.

3.7.1 Testing for Possible Heuristic Survey-Taking Strategies

Our LYVE estimates are meaningful reflections of respondents' preferences only if respondents follow survey-taking strategies that express those preferences. We measure whether respondents

follow certain simple rules of thumb and heuristics that are unlikely to reflect their underlying preferences. We also use estimates from the baseline model of vignette choice to provide a base rate for how choice *appears* to follow a heuristic purely by chance (i.e., even if in truth no heuristic was used). We compare the measured incidence of heuristics to this model-based predicted base rate to assess the degree to which respondents actually use heuristics.

We examine the following six heuristics:

- Always choosing the left vignette
- Always choosing the right vignette
- Alternating between the left and the right vignette, starting with the left vignette
- Alternating between the left and the right vignette, starting with the right vignette
- Always choosing a vignette with a weakly higher level of longevity
- Always choosing a vignette with a weakly higher level of income

For each of these six heuristics, we calculate the fraction of respondents whose choices are consistent with that heuristic in a given wave. So, the six items above are calculated at the respondent-wave level (over six vignette pair choices), and the fractions are averages over respondent-wave-level observations on whether choices were consistent with a particular heuristic.

For each respondent, we define six indicator variables for whether their choices were consistent with a particular heuristic. For example, the “always chose left vignette” indicator will equal 1 if the respondent chose the left vignette for all six vignette pairs in a given wave, and 0 otherwise. The “weakly higher income” indicator equals 1 if the respondent chose the vignette with the higher income for all six vignette pairs in a given wave.

For each wave and heuristic, we then report the share of respondents whose choices were consistent with a particular heuristic. This table shows the rates for six heuristics (averaged over respondent-wave observations).

We then compare these shares to what we’d expect under a null of truly heuristic-free responses, i.e., how often choices are consistent with a heuristic by chance when no heuristic is used. To construct this base rate, we assume that a vignette is chosen if the baseline model predicts that the probability of it being chosen is above 0.50. This is done separately by wave and by respondent for the six vignette pairs that a respondent saw in a given wave. We then calculate the number of times the six heuristics (listed above) are present in the predictions from the baseline model, and also report those base rates in the table above. The difference between the fraction of respondents whose choices are consistent with a heuristic and the base rate for that heuristic forms our estimate of the fraction of respondents that actually use that heuristic.

3.7.2 Testing for Possible Discontinuities in Longevity and Income Differences on Choice

We test whether the probability of choosing a vignette is discontinuous at the point where the two vignettes have equal longevity or have equal incomes. If so, this would indicate that respondents’ choices are inconsistent with an underlying model in which the utility of each vignette solely depends on its own characteristics (including longevity and income). To test for such discontinuities, we add the following two terms to the baseline regression:

- $\beta_4 \text{HigherLongevity}_{A_{ci}}$
- $\beta_5 \text{HigherIncome}_{A_{ci}}$

The variable $\text{HigherLongevity}_{A_{ci}}$ equals 1 if vignette A has a higher longevity than vignette B, equals 0 if both vignettes have the same longevity, and equals -1 otherwise. Similarly, the variable $\text{HigherIncome}_{A_{ci}}$ equals 1 if vignette A has a higher income than vignette B, equals 0 if both vignettes have the same income, and equals -1 otherwise. We perform a joint test of the coefficients on these two variables to determine the presence of discontinuities in choice.

3.7.3 Testing for Possible Effects of Comprehension Checks on Subsequent Choices

After each vignette pair, respondents are given a comprehension check that requires them to compare an aspect between the two vignettes they have just seen. The aspects are randomized between protagonist income, longevity, marital status, whether they had children, and the perceived work disamenities of the protagonists' jobs.

We are interested to know whether respondents' choices become more responsive to longevity difference in the current vignette pair if we have just asked them a comprehension question comparing the longevity of the previous vignette pair. Similarly, we examine whether respondents become more sensitive to the income difference of the current vignette pair if they received a comprehension check about the income of the previous vignette pair. Any such dependence indicates that the implied LYVE is sensitive to manipulations that draw attention to longevity or income, implying that it is critical that respondents face choice environments that do not systematically draw attention to either longevity or income.

We test the effect of comprehension checks focused on income and longevity on subsequent vignette choices by including the following terms to our baseline regression:

- $\beta_4 \text{LongevityCheck}_{c-1,i} * \Delta \ln(\text{Longevity}_{ci})$
- $\beta_5 \text{LongevityCheck}_{c-1,i} * \Delta \ln(\text{Income}_{ci})$
- $\beta_6 \text{IncomeCheck}_{c-1,i} * \Delta \ln(\text{Longevity}_{ci})$
- $\beta_7 \text{IncomeCheck}_{c-1,i} * \Delta \ln(\text{Income}_{ci})$
- $\beta_8 \text{LongevityCheck}_{c-1,i}$
- $\beta_9 \text{IncomeCheck}_{c-1,i}$

In this bulleted list, $\text{LongevityCheck}_{c-1,i}$ is an indicator for whether the comprehension check immediately before the current vignette pair asks about longevity differences of the previous vignette pair. For example, if the respondent is on the fifth vignette pair, $\text{LongevityCheck}_{4i}$ equals 1 if, after the fourth vignette pair, the comprehension check asked about the protagonists' longevity. $\text{IncomeCheck}_{c-1,i}$ is an indicator for the comprehension check immediately before the current vignette pair asking about income differences. This analysis excludes the first vignette choice in each wave since there are no comprehension checks before the first vignette pair in a given wave. We calculate the effects on the LYVE of the longevity and income checks separately using the delta method.

3.7.4 Testing for Possible Effects of Comprehension Checks on Perceived Longevity and Income Differences

Similarly, we examine whether subjective perceptions of longevity or income differences (on a Likert scale) can be manipulated by asking a comprehension check about longevity or income immediately prior to the current vignette pair. If so, this would imply that special care has to be taken not to draw disproportionate attention to either longevity or income prior to asking about subjective perceptions of longevity or income differences.

We test this by running the following regression, abbreviating *Longevity* to *Lon* and *Income* to *Inc*:

$$\begin{aligned} LonLikertVig_{ci} = & \beta_1 \Delta \ln(Lon_{ci}) + \beta_2 \Delta \ln(Inc_{ci}) + \beta_3 \Delta YearOfBirth_{ci} + \\ & \beta_4 LonCheck_{c-1,i} * \Delta \ln(Lon_{ci}) + \beta_5 LonCheck_{c-1,i} * \Delta \ln(Inc_{ci}) + \\ & \beta_6 IncCheck_{c-1,i} * \Delta \ln(Lon_{ci}) + \beta_7 IncCheck_{c-1,i} * \Delta \ln(Inc_{ci}) + \\ & \beta_8 LonCheck_{c-1,i} + \beta_9 IncCheck_{c-1,i} + \delta_{v(A,c,i)} - \delta_{v(B,c,i)} + \epsilon_{ci}. \end{aligned}$$

where *LonLikertVig_{ci}* is the subjective perception of the longevity difference between vignette A and vignette B on a 5-point Likert scale. We also run this regression with *IncLikertVig_{ci}* as the dependent variable, i.e., with the subjective perception of the income difference between vignette A and vignette B on a 5-point Likert scale as the dependent variable. The vignette income and longevity perception questions occur after the last vignette (vignette six in Wave 1 or vignette 12 in Wave 2), so this regression is run only for $c = 6$ or $c = 12$.

3.7.5 Testing for Possible Order Effects of Survey Blocks

The survey contains three main survey blocks: vignettes, tuples, and index questions. The vignette block is randomized to have an equal chance of being either the first or the second main block a respondent sees. If the vignette block is randomized to be second, it has an equal chance of being preceded by the index block or the tuples block.

We want to test whether our baseline estimate of the LYVE (which is based on vignette questions) is sensitive to it being preceded by either the tuples block or the block with the index questions. If so, we put more faith in the estimate of the LYVE derived from the vignette questions that were asked in the first block.

We first test for order effects by estimating the LYVE for each of the following four subsamples (defined at the respondent-wave level) of our baseline sample:

- Respondents who saw vignettes in the first main block in a given wave
- Respondents who saw vignettes in the second main block in a given wave
- Respondents who saw vignettes in the second main block and saw tuples in the first main block in a given wave
- Respondents who saw vignettes in the second main block and saw index questions in the first main block in a given wave

To test for such order effects in a regression framework, we add the following interaction terms to our baseline regression:

- $\beta_4 TupsBeforeVigs_{ci} * \Delta \ln (Longevity_{ci})$
- $\beta_5 TupsBeforeVigs_{ci} * \Delta \ln (Income_{ci})$
- $\beta_6 IdxBeforeVigs_{ci} * \Delta \ln (Longevity_{ci})$
- $\beta_7 IdxBeforeVigs_{ci} * \Delta \ln (Income_{ci})$
- $\beta_8 TupsBeforeVigs_{ci}$
- $\beta_9 IdxBeforeVigs_{ci}$

In this bulleted list, $TupsBeforeVigs_{ci}$ is an indicator that the tuple block is shown before the vignette block. Similarly, $IdxBeforeVigs_{ci}$ is an indicator that the index block is shown before the vignette block. We calculate the derivative of the LYVE with respect to each of these two indicator variables using the delta method. We then perform a joint test of whether these two derivatives are equal to zero.

3.7.6 Testing for Possible Order Effects by Survey Wave

We want to test whether our baseline estimate of the LYVE is stable across the two waves of the survey. Stability of the estimates of the LYVE would increase our confidence in the LYVE estimate being externally valid for other time periods.

We first test for order effects by estimating the LYVE for each of the following three subsamples (defined at the respondent-wave level) of our baseline sample:

- Observations from Wave 1 only
- Observations from Wave 1 only, but limited to those respondents who are also part of the baseline sample in Wave 2
- Observations from Wave 2 only

The comparison of the latter two bullets isolates the effect of the survey wave without confounding it with changes in the sample composition (due to attrition and possible failure to meet the sample inclusion criteria in the second wave). Recall that observations in Wave 2 are part of the baseline sample in Wave 1 by construction.

To test for wave effects in a regression framework, we add the following interaction terms to our baseline regression, limiting the sample to respondents who were part of the baseline sample both in Wave 1 and in Wave 2:

- $\beta_4 Wave2_{ci} * \Delta \ln (Longevity_{ci})$
- $\beta_5 Wave2_{ci} * \Delta \ln (Income_{ci})$
- $\beta_6 Wave2_{ci}$

In this bulleted list, $Wave2_{ci}$ is an indicator variable for the second wave of the survey. We calculate the derivative of the LYVE with respect to this indicator variable using the delta method.

3.7.7 Testing for Question Order Effects

If respondents' preferences are stable and respondents' effort in answering the questions is constant, the estimate of the LYVE should not vary systematically by the order of the vignette pair.

We first test for order effects by estimating the LYVE in six subsamples of the baseline sample, separating the sample by the vignette pair order number within each wave. This yields six separate estimates, and we test whether we can reject that the six LYVE estimates from these six regressions are equal.

Next, we test the stability of the LYVE by the vignette pair order in a regression framework by adding the following interaction terms to the baseline regression:

- $\beta_4 \text{VignetteOrder}_{ci} * \Delta \ln(\text{Longevity}_{ci})$
- $\beta_5 \text{VignetteOrder}_{ci} * \Delta \ln(\text{Income}_{ci})$
- $\beta_6 \text{VignetteOrder}_{ci}$

In this bulleted list, $\text{VignetteOrder}_{ci}$ is a value from 0 to 1 that encodes the order of the vignette pair shown. The variable is defined as the vignette-pair order within a wave, minus 1, divided by 5, so that the first vignette pair in a wave has a value of 0 and the sixth vignette pair within a wave has a value of 1. We calculate the derivative of the LYVE with respect to this order variable using the delta method.

3.7.8 Testing the Sensitivity of Our Estimates to Sample Exclusions

We examine the LYVE separately for five samples defined by the following measures of sample quality:

- The entire sample (including those who were excluded due to the exclusion criteria)
- The excluded sample
- Respondents in the baseline sample who answered at least 75% of the comprehension checks correctly
- Respondents in the baseline sample who answered all comprehension checks correctly
- Respondents in the baseline sample who took at least 60 seconds on all vignette questions (the baseline criteria is at least 30 seconds)

For each sample, we also estimate the derivative of the LYVE with respect to the Index introduced in subsection 3.2.1, as well as report the percent of comprehension checks answered correctly in that sample.

3.8 Suggestive Evidence on Mechanisms from Ratings

As mentioned in subsection 3.3.3, we show the respondents one "rating vignette" and ask them to rate various aspects of the vignette protagonist's life. Specifically, we ask about the following dimensions of the vignette protagonist:

- Ability to spend money
- Life satisfaction
- Health
- Meaningfulness of life
- Work satisfaction, disregarding income
- Relationship satisfaction
- Quality of life toward the end of their life

We also ask about the respondent's relatability to the vignette protagonist, which was detailed in subsection 3.5.3. All rating vignette questions, except for the "Quality of life toward the end of their life", are asked to be evaluated during the protagonist's last healthy decade. For example, if the protagonist lived to 81 and they had cancer which lasted for two years, their last healthy decade would be their 60s since the onset of the condition that led to their death occurred in their 70s. We conduct two sets of analyses on these ratings.

3.8.1 Effect of Randomizations on Ratings

We analyze the effect of income and longevity randomizations, as well as other vignette-text-specific randomizations on the ratings. Any such effects provide suggestive evidence of possible pathways by which the randomizations affected vignette choice. We run the following regression:

$$Rating_{di} = \beta_1 \Delta \ln (Longevity_i) + \beta_2 \Delta \ln (Income_i) + \beta_3 \Delta YearOfBirth_i + \beta_4 MaterialSentences_i + \beta_5 NonMaterialSentences_i + \beta_6 FillerSentences_i + \beta_7 \Delta WorkDisamenity_i + \beta_8 Health_i$$

where $Rating_{di}$ is respondent i 's rating of the rating vignette protagonist on dimension d (listed in the bulleted list above). This analysis is limited only to the rating vignette, so there will only be one observation per respondent.

3.8.2 Effect of Ratings on Choices

To provide suggestive evidence on life aspects that the vignette fixed effect may be capturing, we analyze the effect of average ratings by vignette on choice by including the following term to our baseline regression, from which we exclude the vignette fixed effects:

- $\beta_4 \Delta Rating_{dci}$

where $\Delta Rating_{dci}$ is the difference between vignette A's average rating on a given dimension d and vignette B's average rating on the same dimension d for the c^{th} vignette pair seen by respondent i . The average rating for a given dimension for a specific vignette is averaged over all respondents

who saw that vignette as their rating vignette. For example, if a respondent is choosing between vignette 17 (vignette A) and vignette 28 (vignette B), we would calculate the average value of vignette 17's rating on a given dimension, say, their "ability to spend money", averaged over all the respondents who received vignette 17 as their rating vignette. We would do the same for vignette 28. We will do this for each vignette number, over each dimension. Then we will run separate regressions where the difference between the average ratings between the two vignettes for a dimension is added as an additional explanatory variable one at a time for each dimension. We also run one regression where every dimension's difference in average ratings between the two vignettes will be added all at once. The rationale for this analysis is to understand how different dimensions of a protagonist's life affect a respondent's choice.

3.9 Structural Analyses

The goal of the structural analysis is to allow for more flexible functional forms than our baseline analysis, to jointly analyze multidimensional observed heterogeneity and unobserved heterogeneity, and to increase statistical precision of the estimates by modeling heterogeneity in the accuracy of respondents' choices.

We provide a general outline of the planned structural analyses, but are unable to pre-specify the full details because (i) we want to condition the variables we include on the incremental predictive power they bring to the model and (ii) we cannot predict the degree to which the data will allow us to estimate many dimensions of unobserved heterogeneity and very flexible functional forms.

We plan to use the following outline for the random-utility model in the structural analysis:

- The utility of choosing vignette v with randomizations seen by respondent i takes the form:

$$U_{vi} = \beta_{1vi} \frac{Longevity_{vi}^{1-\gamma_1} - 1}{1 - \gamma_1} + \beta_{2vi} \frac{Income_{vi}^{1-\gamma_2} - 1}{1 - \gamma_2} + \beta_{3vi} YoB_{vi} + \delta_{vi} + \varepsilon_{vi}$$

where $Longevity_{vi}$ is the value of the longevity that respondent i was randomized to see for the protagonist in vignette v , $Income_{vi}$ is the value of the income that respondent i was randomized to see for the protagonist in vignette v , and YoB_{vi} is the value of the year of birth that respondent i was randomized to see for the protagonist in vignette v .

- The additive error is distributed normally: $\varepsilon_{vi} \sim N(0, \sigma_{vi}^2)$.
- Because utility is identified up to a constant and up to a scale parameter, we make two normalizations without loss of generality. First, we set the level of utility by normalizing one of the vignette fixed effects to 0. Second, we set the scale of utility by normalizing the variance of ε_{vi} to 1. In models with heteroskedasticity, we normalize the geometric mean of the variance to 1.
- We model observed, and possibly unobserved, heterogeneity in β_{1vi} and β_{2vi} .
- We model observed heterogeneity in β_{3vi} , δ_{vi} , and σ_{vi}^2 .
- We model any heterogeneity in β_{1vi} as:

$$\beta_{1vi} = \exp(\theta_1 X_{1vi} + \zeta_{1vi})$$

with ξ_{1vi} reflecting unobserved heterogeneity. The parameter vector θ_1 may contain: characteristics of vignette v or of the other vignette shown when respondent i considered vignette v , characteristics of respondent i , responses or response times of respondent i when they considered vignette v , characteristics of the other vignette shown when respondent i considered vignette v , or interactions of the preceding terms. The vector X_{1vi} also contains a constant term.

- We model β_{2vi} symmetrically to β_{1vi} with the analogous explanatory variables.
- We may restrict the error term in the equation for β_{1vi} to be equal to minus the error term in the equation for β_{2vi} so that there is only one dimension of unobserved heterogeneity that governs both β_{1vi} and β_{2vi} .
- We may restrict all the elements of θ_2 other than the constant term to be equal to minus the corresponding element in θ_1 . This restriction would make elements of X_{1vi} have symmetric effects, though in opposite directions, on β_{1vi} and β_{2vi} .
- We model any observed heterogeneity in β_{3vi} as:

$$\beta_{3vi} = \theta_3 X_{3vi}$$

where X_{3vi} is a vector that may contain: characteristics of vignette v or of the other vignette shown when respondent i considered vignette v , characteristics of respondent i , responses or response times of respondent i when they considered vignette v , characteristics of the other vignette shown when respondent i considered vignette v , or interactions of the preceding terms. The vector X_{3vi} also contains a constant term.

- We model any observed heterogeneity in δ_{vi} as:

$$\delta_{vi} = \theta_4 X_{4vi}$$

where X_{4vi} is a vector that may contain: characteristics of vignette v or of the other vignette shown when respondent i considered vignette v , characteristics of respondent i , responses or response times of respondent i when they considered vignette v , characteristics of the other vignette shown when respondent i considered vignette v , or interactions of the preceding terms. The vector X_{4vi} also contains a constant term.

- We model observed heterogeneity in σ_{vi}^2 as:

$$\sigma_{vi}^2 = \exp(\theta_5(X_{5vi} - \overline{X_{5vi}}))$$

where X_{5vi} is a vector that may contain: characteristics of vignette v or of the other vignette shown when respondent i considered vignette v , characteristics of respondent i , responses or response times of respondent i when they considered vignette v , characteristics of the other vignette shown when respondent i considered vignette v , or interactions of the preceding terms. The vector X_{5vi} does not contain a constant term. The sample mean of X_{5vi} is denoted by $\overline{X_{5vi}}$. Subtracting out the sample mean ensures that the geometric mean of σ_{vi}^2 is equal to 1.

- The vectors X will generally have different explanatory variables because we will tailor the explanatory variables to those that we consider a priori plausible for the parameter in which we are modeling heterogeneity. Specifically, vectors X_{1vi} and X_{2vi} contain predictors of the

respondent's valuation of longevity and income. The vector X_{3vi} contains predictors of respondent i 's valuation of the time period in which the protagonist in the vignette lived. The vector X_{4vi} includes predictors that we expect to affect respondent i 's valuation of the contextual information in vignette v . Finally, in X_{5vi} , we include predictors that we expect to affect the accuracy of a respondent's choice in the vignette pair that includes vignette v .

- Because heterogeneity in *all* preference parameters (i.e, all β s and δ s) cannot be separately identified from heterogeneity in σ_{vi}^2 , the variables in X_{5vi} do not overlap with the variables in X_{1vi} , X_{2vi} , X_{3vi} or X_{4vi} .
- We may choose values for some of the parameters specified above such that specification becomes more restrictive. For example, the functional form could become more restrictive if we restrict $\gamma_2 = 1$ so that longevity takes a log form. For example, unobserved heterogeneity in one of the parameters would disappear if we set the variance of the error term of that parameter to zero, and we may not have unobserved heterogeneity in any of the parameters. For example, observed heterogeneity would disappear if we restrict an X-vector to contain only a constant term, and we may not have observed heterogeneity in any of the parameters.

3.10 Value of a Statistical Life

As a comparison for our LYVE estimate, we estimate respondents' Value of a Statistical Life (VSL) in Wave 2 of the survey with questions based on [Hakes and Viscusi \(2007\)](#) and [Viscusi \(2008\)](#). Respondents are given one of two risk scenarios and asked how much they would be willing to pay to reduce their risk of dying. The two risk scenarios are randomized equally between a "stadium" setting and a "car crash" setting. The reduction in the probability of dying in the scenario, as well as how it is presented to respondents, is randomized across respondents with equal probability between 1/20,000, 5/100,000, and 1/100,000.

The two possible scenarios and three risk reduction options create six unique VSL questions, where only one is shown to the respondent. The six possible VSL questions are shown below with the portions relating to the risk reduction option shown in bold (in this document, but not for the respondents):

- Stadium – 1/20,000: *How much would you agree to pay to eliminate a one-time only risk of death of **1/20,000**? To conceptualize this risk, it is helpful to imagine that you are attending a sporting event in a crowd of **20,000**, where **one** of you will not survive. Suppose you have to pick from one of the following dollar categories as the most that you would agree to pay to get out of this **1/20,000** one-time only risk of death*
- Stadium – 5/100,000: *How much would you agree to pay to eliminate a one-time only risk of death of **5/100,000**? To conceptualize this risk, it is helpful to imagine that you are attending a sporting event in a crowd of **100,000**, where **five** of you will not survive. Suppose you have to pick from one of the following dollar categories as the most that you would agree to pay to get out of this **5/100,000** one-time only risk of death*
- Stadium – 1/100,000: *How much would you agree to pay to eliminate a one-time only risk of death of **1/100,000**? To conceptualize this risk, it is helpful to imagine that you are attending a sporting event in a crowd of **100,000**, where **one** of you will not survive. Suppose you have to pick from one of the following dollar categories as the most that you would agree to pay to get out of this **1/100,000** one-time only risk of death*

- Car crash – 1/20,000: Your annual risk of death in a car crash is roughly **2/20,000**. To illustrate this risk, the following picture depicts **two** black dots against a background of **20,000** dots. Suppose you could reduce your annual risk of death in a car crash in **half**, that is lower the risk from **2/20,000** to **1/20,000**. Thus, if there were **20,000** people just like you, there would be **one** fewer expected death per year in your group. How much would you agree to pay each year either for a safer car or for improved highway safety measures that would reduce your annual risk of death in a car crash by **half**?
- Car crash – 5/100,000: Your annual risk of death in a car crash is roughly **10/100,000**. To illustrate this risk, the following picture depicts **ten** black dots against a background of **100,000** dots. Suppose you could reduce your annual risk of death in a car crash in **half**, that is lower the risk from **10/100,000** to **5/100,000**. Thus, if there were **100,000** people just like you, there would be **five** fewer expected deaths per year in your group. How much would you agree to pay each year either for a safer car or for improved highway safety measures that would reduce your annual risk of death in a car crash by **half**?
- Car crash – 1/100,000: Your annual risk of death in a car crash is roughly **10/100,000**. To illustrate this risk, the following picture depicts **ten** black dots against a background of **100,000** dots. Suppose you could reduce your annual risk of death in a car crash by **a tenth**, that is lower the risk from **10/100,000** to **9/100,000**. Thus, if there were **100,000** people just like you, there would be **one** fewer expected death per year in your group. How much would you agree to pay each year either for a safer car or for improved highway safety measures that would reduce your annual risk of death in a car crash by **a tenth**?

The answer choices for the VSL question are “Under \$50,” “\$50-\$199,” “\$200-\$499,” “\$500-\$1,000,” “Above \$1000,” and “Infinite amount: the total of all your present and future resources.” Answer choices are converted into numeric values by taking the midpoint of the range of the answer choice. For example, the numeric value of the “\$500-\$1,000” choice is \$750. The “Above \$1,000” choice is given a value of \$1,500 following [Hakes and Viscusi \(2007\)](#). The “Infinite amount” choice is not converted into a numeric value and observations making this choice are excluded in the reports of the mean and regression analyses. The answer choice order is randomized across respondents with equal probability so that the “Under \$50” choice is shown either first or last.

On the question page, we show a picture of black dots in an array of lighter dots to visually depict the risk reduction option listed in the respondent’s question. The 1/20,000 picture is shown in Figure 1, and the other two risk reduction options have analogous figures.

Respondent i ’s VSL is calculated as:

$$VSL_i = \frac{ChoiceValue_i}{p_i}$$

where p_i is the reduction in the probability of dying as presented to respondent i . If the respondent chose the “\$500-\$1,000” option for a risk reduction of 1/20,000, then their VSL would be $\$750/(1/20,000) = \$15,000,000$.

To analyze the effects of the risk scenario and the effect of the risk reduction option, we run the following regression:

$$\log(VSL_i) = \beta_1 LowerRisk_i + \beta_2 LargerNumeratorOfRisk_i + \beta_3 CarCrash_i + \epsilon_i$$

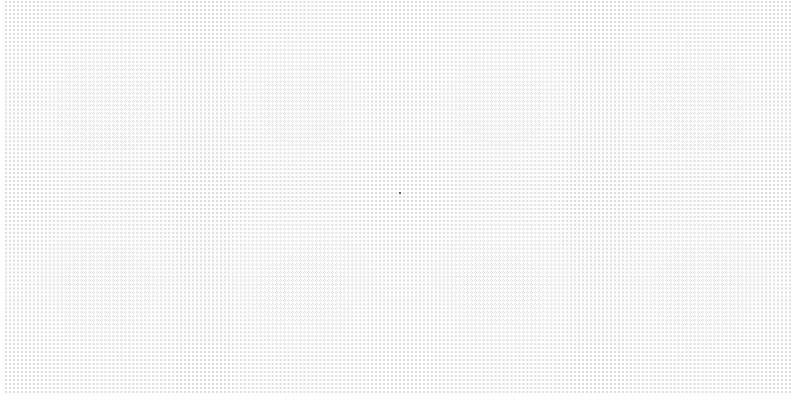


Figure 1

where $LowerRisk_i$ is an indicator variable for the risk reduction option being 1/100,000 (rather than 1/20,000 or 5/100,000), $LargerNumeratorOfRisk_i$ is an indicator for denominator of the risk reduction option being displayed as 100,000 (rather than 20,000), and $CarCrash_i$ is an indicator for the risk scenario being the car crash setting (rather than the stadium setting).

We also test whether the *Index* affects an individual's VSL. Specifically, we run:

$$\log(VSL_i) = \beta_1 LowerRisk_i + \beta_2 LargerNumeratorOfRisk_i + \beta_3 CarCrash_i + \beta_4 Index_i + \epsilon_i$$

where $Index_i$ refers to the respondent's *Index* value detailed in subsection [3.2.1](#).

References

- Guvenen, F., Kaplan, G., Song, J., and Weidner, J. (2022). Lifetime earnings in the united states over six decades. *American Economic Journal: Applied Economics*, 14(4):446–79.
- Hakes, J. K. and Viscusi, W. K. (2007). Automobile seatbelt usage and the value of statistical life. *Southern Economic Journal*, 73(3):659–676.
- U.S Centers for Disease Control and Prevention (2020). Available at https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/NVSR/71-01/Table01.xlsx.
- Viscusi, W. K. (2008). How to value a life. *Journal of Economics and Finance*, 32(4):311–323.