A more refined version of this preprint has been published:
Kratz, F. and J. Brüderl. 2025. Assessing age trajectories (of subjective well-being): clarifying estimands, identification assumptions, and estimation strategies. European Sociological Review; https://doi.org/10.1093/esr/jcaf038
Please refer to the published version, when citing this work.

The Age Trajectory of Happiness:
How Lack of Causal Reasoning has Produced the Myth of a U-Shaped Age-Happiness Trajectory

Fabian Kratz and Josef Brüderl

LMU Munich

March 2021

fabian.kratz@soziologie.uni-muenchen.de

Abstract

A large interdisciplinary literature on the relationship between age and subjective well-being (happiness) has produced very mixed evidence. Virtually every conceivable age-happiness trajectory has been supported by empirical evidence and theoretical arguments. Sceptics may conclude that the social science of happiness can only produce arbitrary results. In this paper we argue that this conclusion is premature. Instead, the methodological toolbox that has been developed by the modern literature on causal inference gives scholars everything they need to arrive at valid conclusions: the causal inference toolbox only must be applied by happiness researchers. We identify four potential sources of bias that may distort the assessment of the age-happiness relationship. By causal reasoning we derive a model specification that avoids these biases. For an empirical illustration, we use the longest running panel study with information on happiness, the German Socio-Economic Panel (1984-2017; N persons=70,922; N person-years =565,703). With these data we demonstrate the relevance of the four biases and how combinations of different biases can reproduce almost any finding from the literature. Most biases tend to produce a spuriously U-shaped age trajectory, the most prominent finding from the literature. In contrast, with our specification we find a (nearly monotonic) declining age-happiness trajectory.

Keywords: Aging, Happiness, Subjective Well-being, Life Satisfaction, Causal Inference

The social and behavioral sciences have been shaken by a replication crisis that generated much doubt concerning the credibility of research results. While most of these discussions refer to experimental studies, meanwhile it has become clear that non-experimental (observational) studies also have a credibility problem (Author citation, 2021). Recent papers identified several reasons for the low credibility of observational studies. Ferguson & Heene (2012: 555) argue that the coexistence of "undead theories" hampers theoretical advancement. What we need are decisive hypothesis tests that eliminate incorrect rival hypotheses. Grosz, Rohrer, & Thoemmes (2020) identify the "taboo against explicit causal inference" as a leading source that hinders progress in non-experimental research. Explicit causal reasoning could help in developing precise research questions and minimizing the mis-specification of estimation models. Precisely defined research questions are essential for decisive hypothesis tests as are correctly specified estimation models (Lundberg et al., forthcoming).

A premier example where such issues have produced a serious credibility crisis is research on the age trajectory of subjective well-being (SWB)[1]. As the ongoing highly polarized debate highlights (see e.g., Galambos, Krahn, Johnson, & Lachman, 2020; Blanchflower & Graham, 2020), there is – despite dozens of studies over the last decades – no consensus on how the age-happiness trajectory evolves. A literature review (see below) discloses a puzzling pattern of mixed empirical evidence. This is especially disturbing since the question on how happiness is affected by ageing is a fundamental question on the "conditio humana". The social and behavioral sciences seem unable to answer such a fundamental question despite much research effort!

In this paper we argue that the main reason for this unsettling state of affairs is a lack of causal reasoning in the research literature on the age-happiness trajectory. Identifying the causal effect of aging on happiness[2] needs a precisely defined research question and model specifications that are well founded in causal inference methodology. We will show how clarifying subtle nuances in research questions (i.e., distinguishing between the question "how does aging affect happiness" and "how happy are individuals at different ages"), and choosing

[1] In this paper we often use "happiness" as the every-day synonym for the scientific construct of subjective well-being (SWB). SWB is commonly seen as comprising three indicators (Jebb et al., 2020): the cognitive assessment of SWB (life satisfaction), positive affect, and negative affect. Our own empirical work below will focus on the single-item measure "life satisfaction".

[2] In a narrow counterfactual sense age cannot have a causal effect, because it is not manipulable. In a wider sense aging affects mechanisms that cause (because they are manipulable) happiness. Grosz et al. (2020) argued convincingly that it is not productive to avoid the "C-word" ("causal"). In our context, many authors state that they investigate only the association between age and happiness. The consequence of this self-restriction is, however, that they can only answer the question on how happiness is distributed by age in a population. However, the much more interesting question on how age (-related mechanisms) affects happiness cannot be answered.

model specifications according to modern causal inference methodology can reconcile much of the divergent findings in the literature.

Admittedly, conceptualizing age as a cause is challenging. Identifying aging effects on happiness involves intriguing argumentations on model specifications: Researchers must find an adequate functional form, they must deal with statistical confounding (covariate selection and unobserved heterogeneity bias), they must tackle natural confounding (controlling for period and cohort when estimating age trajectories), and they are confronted with endogenous selection of the population (i.e., the probability to survive depends on age and happiness). Nevertheless, the methodology is out there. We only have to use it. This is what we want to show in this paper.

We begin with illustrating how studies from demography, economics, gerontology, psychology, and sociology contributed to a puzzling pattern of mixed empirical evidence begged by conflicting theorizing that coexists in fragmented research disciplines. We proceed by drawing on a causal inference framework to outline subtle nuances in research questions that must be distinguished to arrive at valid conclusions. Based on these research questions (and effect definitions) we discuss four stylized statistical pitfalls (i.e., quadratic specification bias, undercontrol bias, overcontrol bias and mortality selection bias) that may bias results from happiness regressions. For our empirical demonstration of the relevance of these four biases we use the longest running panel data set with information on life satisfaction: the German Socio-Economic Panel study. First, we show how the four pitfalls may bias estimated age-happiness trajectories. Further, we demonstrate how different combinations of these biases may have produced the picture of mixed empirical evidence. Finally, we use a specification that avoids the four biases to estimate the age-happiness trajectory with our data.

Overall, this contribution aims to clarify arguments in the age-happiness debate. By this clarification we hope to settle the fierce discussion about the U-curve. In addition, we believe that the general approach of causal reasoning that we use in this paper may also prove to be useful for other research fields in the social and behavioral sciences.

Mixed Evidence and Coexisting “Undead” Theories

Despite the fundamental importance of the age-happiness profile for interdisciplinary research there is a puzzling picture of mixed empirical evidence (see our Literature Review in Supplementary Table 1, for other literature reviews and similar conclusions see Galambos et al., 2020 or Ulloa et al., 2013). Although a U-shaped age-SWB profile appears to be the most prominent finding, almost every conceivable trajectory has been supported by empirical evidence. Given this outcome, a sceptic could conclude that the social science of happiness is unable to provide valid results. A different sceptic might conclude that statistics can prove anything. Most confusing is the fact that most authors provide a (more or less plausible) theory for explaining the pattern found.

We illustrate the state of research by using six stylized age-happiness trajectories (see Figure 1). For convenience we define three life course stages: *adulthood* (18 - 54), the *golden ages* (55 - 64), and *old age* (65+).

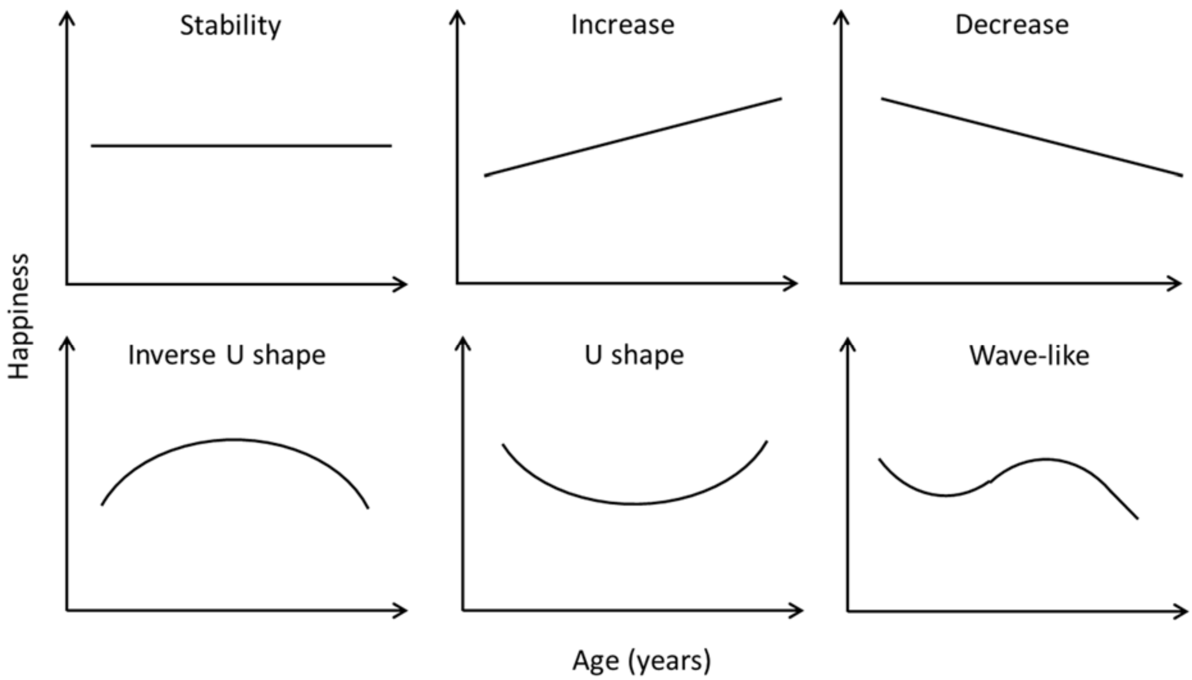


Fig. 1 | Stylized age-SWB trajectories, as reported by previous studies.

Studies reporting <u>SWB stability</u> (Realo & Dobewall, 2011) refer to the psychological *set point theory* and argue that people possess a fixed set point of SWB. Critical life events (e.g., marriage, divorce, widowhood, unemployment) may cause temporary deviations from this set point, but psychological adaptation mechanisms regulate SWB such that the initial set point is reached again after an adaptation process (Lykken & Tellegen, 1996). Gerontological studies additionally refer to the *'stability-despite-loss paradox'* (Kunzmann et al., 2000) to explain why SWB remains stable during the golden ages and through old age. They reason that mental strategies help older people to cope with physical and social losses such that SWB remains stable despite setbacks such as decreasing health or the death of friends or partners.

Studies reporting an increase in SWB (Keyes et al., 2002; Sutin et al., 2013; Yang, 2008) mainly argue that mental strategies help older people to cope better with different situations. According to this *age-as-maturity hypothesis,* psycho-sociological traits such as *self-esteem*, *self-integration*, and *insight* increase with rising age, which in turn increases SWB (Yang, 2008). Studies reporting an increase in SWB during the golden ages also refer to the *anticipation of retirement* or psychological adaptation mechanisms, such as *reduced expectations* and *adaptation to more realistic life goals* (Wunder et al., 2013).

Studies reporting a decrease in SWB (Kassenboehmer & Haisken-DeNew, 2012; Wolbring et al., 2013) predominantly argue that declining physical health causes a decrease in SWB with rising age. Studies aiming to explain this decrease during adulthood additionally mention *reduced biological activity* and *lower episodic memory performance* (Wunder et al., 2013), while others reporting a decline during old age argue that the growing awareness of the finiteness of life decreases SWB (Gerstorf et al., 2010).

Studies reporting an inverse U-shape (Easterlin, 2006; Glenn, 2009) relate a growing SWB during adulthood to a growing satisfaction with the employment and family situation. In turn, they explain the diminishing satisfaction after midlife by decreased satisfaction in these life domains.

Studies reporting a U-shape (Blanchflower, 2020; Blanchflower & Oswald, 2008, 2009; Cheng et al., 2017; Stone et al., 2010; Weiss et al., 2012) deliver neurological, evolutionary, and behavioral explanations. The *neurological explanation* relates the seemingly U-shaped age-SWB profile to age-specific changes in brain structures. The *evolutionary explanation* argues that species display a higher fitness if their genes support the U-shaped age-SWB profile. The *behavioral explanation* is consistent with psychological theories that postulate a midlife crisis. To explain an increase in SWB during old age, Blanchflower and Oswald additionally mention the following *social comparison mechanism*: "I have seen school-friends die and come eventually to value my blessings during my remaining years" (Blanchflower & Oswald, 2008: 1747).

Finally, studies reporting a more wave-like, curvilinear pattern (Baird et al., 2010; Bauer et al., 2017; Frijters & Beatton, 2012; McAdams et al., 2012; Wunder et al., 2013) explain these patterns by combining several of the life course stage-specific mechanisms mentioned above.

The multitude of contradictory findings is unsettling for the social science of happiness. Thus far, there is no definite answer to such a seemingly simple question: How does aging affect happiness? The literature reports widely diverging empirical results substantiated by (more or less) plausible theoretical arguments. These arguments are hardly scrutinized against each other, but rather coexist as "undead theories" in different fragmented research disciplines.

In the following we will argue that the reason for these divergent results lies in the difficulties of drawing causal inferences from observational data. Many potential sources of bias threaten causal inferences from observational data; thus, age-happiness trajectory estimates may be incorrect if sources of bias are not accounted for (Diener et al., 2018). Methodological discussions over the past decades have provided all elements necessary for an "optimal" estimation approach – researchers have only been remiss to include these elements into their work. In the following we try to bring these elements together.

Applying Causal Reasoning: How to Specify Age-Happiness Regressions?

Scholars must avoid four main methodological problems to arrive at valid conclusions regarding the age happiness trajectory: *quadratic-specification bias* (i.e., imposing a quadratic form although the observed pattern is more complex), *under-control bias* (i.e., not controlling for variables that distort the age-SWB trajectory), *over-control bias* (i.e., controlling for variables that induce the age-SWB trajectory), and *mortality selection bias* (i.e., not taking into account that people with lower SWB set points die earlier). We will show in the empirical part of the paper that these biases can have drastic consequences: They may even affect qualitative conclusions and they have most likely produced much of the divergence in research findings.

The quadratic specification bias results if scholars impose an inadequate quadratic functional form. The popularity of the U-shape produced a research praxis in which scholars use a quadratic age specification as "the default", although more complex curvilinear relationships cannot be ruled out on theoretical grounds. With quadratic specifications the age-SWB profile is forced to follow a specific trajectory, ruling out the identification of more complex trajectories. Furthermore, with a quadratic age-SWB specification identifying when SWB levels are lowest generally will fail (Jebb et al., 2018). This is of fundamental importance, as it involves the test of theoretical mechanisms: Is there a quarter-life or a midlife crisis, or a pension hump? Therefore, modern statistical analysis has moved beyond imposing simple functional forms towards semi-parametric or even non-parametric modelling (Ranjbar & Sperlich, 2019). Despite this progress in the flexibility of statistical modelling, our literature review indicates that many studies continue to use a quadratic age specification.

Undercontrol bias results from not controlling for confounding variables, i.e., those that affect both the treatment (age) and the outcome (SWB). In certain disciplines, this is also known as “confounding bias”, “omitted variable bias”, or “unobserved heterogeneity bias”. Defining confounding variables when estimating the age-SWB profile is not straightforward. In a strict sense of causality, no variable can have a causal effect on age, because age cannot be manipulated. Therefore, one could argue that no confounders exist. This, however, is a pitfall; “affect” not only encompasses causal effects, but also implies correlations. Age could correlate with many variables due to age-specific compositional effects. For example, in many Western societies, migrants are younger than native-born residents. If migrants report lower SWB than natives (Safi, 2010), it will erroneously appear as if SWB increases with age. Therefore, all (ascriptive) characteristics that affect SWB and are distributed non-randomly over age constitute potential confounders and should be controlled for when estimating age-SWB profiles. Typically, these include sex, migrant status, social origin, and region of origin.

To avoid *undercontrol bias,* scholars must in addition control for period effects (i.e., societal changes affecting the whole population, e.g., economic situation) and cohort effects (i.e., historical effects that influence generations even when the events are over, e.g., wars). Completely controlling for period and cohort effects when estimating age effects is not possible, as all three are linearly dependent (Baetschmann, 2013; De Ree & Alessie, 2011; Ferrer-i-Carbonell & Frijters, 2004; Glenn, 2009; Wunder et al., 2013; Yang, 2008). As entirely technical solutions will not suffice (Bell & Jones, 2014a, 2014b, 2015; Luo, 2013a, 2013b; Luo & Hodges, 2016; Pelzer et al., 2015), a growing consensus of age-period-cohort scholars advocate for theory-guided restrictions (Bell & Jones, 2015; Chauvel & Schröder, 2015; Fienberg, 2013; Glenn, 2005; Heckman & Robb, 1985). Finding suitable theory-guided restrictions involves two considerations: First, deciding whether period or cohort effects create more systematic biases and, second, deciding whether period or cohort effects can be more easily captured by explicitly measured variables.

In most cases, cohort effects cause more systematic bias, whereas period effects can be captured more easily by macro variables. Period effects are usually erratic (Easterlin & Plagnol, 2008; Schupp et al., 2013), causing SWB to increase in one year by a nationwide event (e.g., winning a football world-championship) and decrease again in another year (e.g., due to an economic crisis). Such erratic variation rarely biases age-SWB profiles in a systematic way. In addition, because there is consistent evidence regarding period effects (Di Tella et al., 2001, 2003; Frijters et al., 2004) (e.g., GDP/wage growth increases SWB; unemployment rate, unemployment benefit cuts, economic crises all reduce SWB) the set of period variables that one should control for is also well determined.

In contrast, cohort effects can more easily cause systematic distortions. Experiences during adolescence might have an effect on the SWB level all over life. Whether experiencing hard times (e.g., war) has positive (via lower aspirations) or negative effects (via psychological scars) on SWB is still up for debate (Sutin et al., 2013; Wunder et al., 2013; Yang, 2008). In the German case the war generation is happier presumably because of lower aspirations (see below). Thus, it might be that a higher level of SWB in old age is not the result of age but still reflects the impact of the war. Therefore, it is important to control for birth year in happiness regressions.

Finally, scholars must also control for method effects. In particular, SWB responses may be biased due to social desirability (Chadi, 2013; Heintzelman et al., 2014). Generally, society expects us to be happy. The tendency to report inflated SWB values is higher if an unknown person (i.e., an interviewer) is asking. We therefore expect over-reporting in all survey modes for which an interviewer is present. If the interviewee self-completes the questionnaire, the pressure to appear happy is lower. Many panel studies begin with an interviewer-guided wave followed by paper-and-pencil waves once respondents are familiar with the survey. Furthermore, over the course of a panel, respondents become familiar with the interviewer, resulting in more truthful, lower SWB reports (Frijters & Beatton, 2012; Kassenboehmer & Haisken-DeNew, 2012; Wunder et al., 2013). Such method effects heavily bias age-SWB profiles at the age of panel entry (usually 18), resulting in an upward "kink". In addition, these method effects could lead to a strong decrease in the age-SWB profile in the first years of the study. Therefore, we suggest controlling for interviewer presence during questionnaire completion and a "years in the panel" variable. However, as this effect rapidly wears off (Supplementary Table 6, model (3)), dummy variables for the first waves are sufficient. If the social desirability reasoning holds, over-reporting should not only occur in the first few waves of a panel, but also after each interviewer change. We therefore suggest also including dummies for the first waves after an interviewer change.

Overcontrol bias results from controlling for mediating variables, i.e., variables that result from the treatment (age) and affect the outcome (SWB). A growing causal inference literature clearly demonstrates that controlling mediating mechanisms distorts total causal effect estimates (Elwert, 2013; VanderWeele, 2015)[3]. The rationale is clear: If such variables are

[3] To be clear: Overcontrol bias is only a threat if one is interested in the "total causal effect". Since our research question is on the total effect of aging on happiness, this threat applies. This argument would not apply if one would be interested in the "direct causal effect" (the effect that remains, after controlling for mediating mechanisms). Below we will comment on this point more extensively.

included, the causal effect will be distorted by explaining away (part) of its mechanisms. Although Glenn (2009) has illustrated how erroneously controlling for family events distorts age-SWB profiles, overcontrol bias is still very common in this field (Supplementary Table 1) and some even argue that critical life events must be controlled for to uncover the causal effect of age on SWB (Blanchflower & Oswald, 2009). To illustrate the erroneousness of this reasoning consider the following examples: If family events are included, important reasons for happiness during adulthood are explained away. If we include employment characteristics and income, favorable economic circumstances that boost happiness (Jebb et al., 2018; Wolbring et al., 2013) are lost, especially when peak household income is reached during midlife. If health is included, biological aging, one key aspect that brings about decreasing health, is lost, biasing SWB estimates especially during golden and oldest age. Thus, some typical control variables (Supplementary Table 1) severely distort the trajectory by selectively controlling various age-specific sources of SWB away.[4]

Mortality selection bias results from not accounting for biases due to the lower longevity of respondents with low SWB. Mortality selection is a special case of what is known as "Berkson's Bias" (Berkson, 1946) in biology, or "endogenous selection bias" in sociology (Elwert & Winship, 2014). These insights have not been taken systematically into account when studying the age-SWB profile. That higher SWB levels throughout life result in longer lifetimes is a well-established finding in happiness research (Diener et al., 2018; Diener & Chan, 2011). Consequently as happier people live longer, respondents with higher SWB set points will be over-represented in the sample at older ages (Frijters & Beatton, 2012; Kassenboehmer & Haisken-DeNew, 2012). Thus, mortality selection will bias the age-happiness trajectory upwards in old ages.

As a result of such composition effects due to mortality selection, scholars must differentiate between two research questions. First, "How happy are people (who are alive) at different ages?". This descriptive research question can be answered using cross-sectional data and averaging SWB levels (after controlling for confounders) at different ages. The second, more causal research question asks: "How does aging affect happiness?". If everyone would die at the same age, this question could be answered analog to the first. As unhappier people die earlier, however, this is not possible. Instead, panel data following the same respondents over

[4] Overcontrol bias (post-treatment conditioning) is very widespread in sociological research more general. Lundberg et al. (forthcoming) argue that 10 out of 18 articles using observational data in the 2018 volume of the American Sociological Review condition on a post-treatment variable though they are interested in a total causal effect.

time must be utilized to assess how SWB develops for each person, and then average the trajectories.

If mortality selection operates along a fixed SWB level (i.e., the set point), implementing such a within approach will remedy the bias. For instance, a regression with individual-specific fixed effects (FE regression) will do the job: Here the age effect is estimated by comparing SWB at different ages for each individual (for details, see Methods section). Contrary to popular belief, random effects approaches (RE model, multilevel model, latent growth curve model, etc.) do not solve the problem, but rather result in a measure of weighted between and within effects that has no substantial meaning (Rohwer, 2016). Because mortality selection removes the less happy from the sample, between-person variation will bias the age-SWB profile upwards at older ages.

Altogether, causal reasoning provides a model specification that avoids all four biases. Quadratic specification bias is avoided by including a dummy for every age – the most flexible age parametrization available. Further, our preferred specification controls for confounding variables to avoid undercontrol bias and does not control for mediating variables to avoid overcontrol bias. To tackle mortality selection bias, it is suggested using a fixed effects regression that relies only on within-person variation.

Data, Measures, and Analytical Approach

Data. Data used are from the German Socio-Economic Panel Study (SOEP V34, 1984-2017) (Socio-Economic Panel, 2019; Wagner et al., 2007). The SOEP is a large-scale representative survey of the adult German population (ages 18 and older) residing in private households. These data are especially appropriate for this paper for various reasons: The data cover a large time span with a maximum of 34 yearly observations (person-years) per respondent, they provide a large number of cases (interviews) per wave (about 30,000 in the most recent waves), and contain a single-item life satisfaction measure in each wave from the beginning.

We restricted our analysis sample to persons between 18 and 90. Due to some data peculiarities, we do not use all available SOEP subsamples. In particular, we exclude the following: SOEP Innovation Samples, “Families in Germany” (different questionnaire), and the migration sample from 2013 (different procedures and questionnaire). Person-years with missing values on any variable included in the respective model are also dropped (listwise deletion on the person-year level). After these restrictions 70,922 respondents provide 565,703 person-year observations for our analyses. Supplementary Table 1 lists summary statistics.

Estimates from models that include mediating variables are based on a reduced sample without the years 1984-1990, and 1993, because the health variables were not collected in these years ("only" 49,610 respondents providing 425,476 person-years).

Measures. Subjective well-being is measured by the following single-item satisfaction question: "How satisfied are you with your life, all things considered?" Respondents answered this life satisfaction question on an 11-point scale ranging from 0 to 10 (0: "completely dissatisfied", 10: "completely satisfied").

Age is a (technical) survey variable measured in years. We include a separate dummy for each age year from 18 to 90 (reference: age 18).

We control for the following time-constant confounders: nationality as recorded in the first interview (German dummy) and sex (female dummy). These variables are automatically controlled for in fixed effects (FE) models. As a time-varying confounder, we include an indicator as to whether a person lives in former East or West Germany (East dummy).

Period: To avoid the age-period-cohort (APC) problem, we do not include (all) survey year dummies. Instead, we control for period effects by including three macro variables: unemployment rate, GDP growth, and wage growth. As the SOEP interviews are carried out early in the year (80% of interviews completed between February and April), these period variables are lagged, thus controlling for the economic situation of the previous year. In addition, we include dummies for two exceptional years: In 2004, Germany recorded an all-time life satisfaction low due to the political debates about the "Agenda 2010" reforms (Schupp et al., 2013). Similarly, in 2009, shortly after the peak of the financial crisis, life satisfaction was exceptionally low.

Method effects: A dummy for interviewer presence controls for the interview mode effect. The social desirability effects due to a new interviewer are captured by dummies for the first three waves of each respondent (note that the age at which respondents enter the SOEP varies widely) as well as the first three years after an interviewer change (these dummies are exclusive).

Mediating mechanisms: Human capital variables are years of education, employment status (full-time (ref.), part-time, marginal (minor) employment, apprenticeship, out of labor force, unemployed, retired), and the deflated net household income (ln). To capture family events, we include dummies for relationship status (single (ref.), married, separated/divorced, widowed) and the log household size. As health-related measures we use a dummy for disability status, the number of nights a respondent spent in a hospital during the last year, and self-rated health.

Analytical Approach. We assume that a linear regression is appropriate, although life satisfaction is measured on an ordinal scale (some authors argue that this assumption is innocuous (Ferrer-i-Carbonell & Frijters, 2004)). We avoid a quadratic specification bias by including age dummies. Further, we control for confounding variables to avoid undercontrol bias and do not control for mediating variables to avoid overcontrol bias. To tackle mortality selection bias, we use a fixed effects regression. The following regression specification results from these considerations:

$$SWB_{it} = \alpha_i + \sum_{n=19}^{90} \theta_n Age_{n,it} + \boldsymbol{\beta}'\boldsymbol{X}_{it} + \gamma C_i + \pi P_t + \boldsymbol{\delta}'\boldsymbol{M_{it}} + \varepsilon_{it}. \qquad (1)$$

SWB_{it} is the SWB of individual i at time t. We model SWB as a function of age (also called a "growth curve"). To ensure the age effect is not biased by an over-restrictive specification, we use a completely flexible approach in which each age is captured by a dummy variable (72 dummies from age 19 to 90; reference: 18). θ_n estimates the difference in satisfaction between age 18 and age n. These are the parameters of interest, given our research question. Their development over age gives us the effect of aging on SWB.

The confounding variables are represented by four terms. $\boldsymbol{X}_{it}$ are (ascriptive) time-constant (e.g., sex and nationality) and time-varying (e.g., region) characteristics. C_i represents birth cohort (dummies for each birth year), whereas P_t represents period. As argued above, due to the APC problem we do not use a linear period term, but proxy variables. Finally, $\boldsymbol{M_{it}}$ represents variables that capture method effects.

Equation (1) includes an idiosyncratic error term ε_{it} and an individual-specific constant α_i (unobserved SWB set point). How α_i is modeled is essential: It is common in the (psychological) growth curve literature to model α_i as a random effect (i.e., RE model, multilevel model, latent growth curve model, etc.). Such random effects models must assume that α_i is not correlated with $Age_{n,it}$ (random effects assumption). This assumption would be met if happiness had no impact on longevity. Because happy people live longer it is violated. Therefore, random effects growth curve models will yield biased age effects and, consequently, the age-SWB trajectory will also be biased.

An alternative is to model α_i as a fixed effect (FE). Parameter estimation in a FE model relies only on within-person variation. As the random effects assumption is not needed (Author citation, 2015), the FE model is strongly advocated in the (economic) well-being literature

(Ferrer-i-Carbonell & Frijters, 2004; Frijters & Beatton, 2012). Thus, age effects are not biased in a FE estimation even if α_i is correlated with $Age_{n,it}$. In addition, a FE model controls also for all unobserved time-constant confounding variables (i.e., ascriptive characteristics and cohort) that could bias RE estimates. Finally, the FE model has another much less known property that is highly important when assessing the SWB profile: it removes endogenous selection bias related to SWB levels and thus a major threat to valid conclusions. We therefore suggest estimating equation (1) with FE models.

Estimation was done by Stata 16.1. Results are presented graphically due to the large number of estimated coefficients. Detailed regression tables corresponding to the Figures are available in the supplementary material.

How do the Four Biases Affect the Age Trajectory of Happiness?

With causal reasoning we identified four potential sources of bias. But are these biases of relevant size in real data? To investigate this question, we recreate each of the four biases step by step with our data.

To evaluate the importance of quadratic specification bias, we compare the results of a non-parametric model with dummy variables for every age (Figure 2a, grey line) with the quadratic function model (Figure 2a, blue line). Though the non-parametric profile by no means is U-shaped, a U-shape appears in the quadratic model due to the scarcity of respondents over 70 to "pull the trajectory down".

Figure 2b assesses the importance of *undercontrol bias* by successively introducing method variables (i.e., first three years in the panel, interviewer presence, first three years after interviewer change), cohort variables (i.e., dummy variables for every birth year), period variables (i.e., unemployment rate, growth rate of real disposable income, GDP growth rate, dummies for years 2004 and 2009), and individual controls (i.e., gender, nationality, and living in former East Germany). The reference trajectory that controls for nothing (Figure 2b, grey line) shows a "kink" below age 20 which is indication for a method effect bias. Controlling the method variables (Figure 2b, blue line) reveals this kink to be the result of social desirability when answering the life satisfaction question for the first/second/third time. Otherwise, the trajectory changes only marginally. When controlling for cohort variables, the decrease in SWB during adulthood becomes substantially less pronounced (Figure 2b, green line). The inclusion of period variables and individual controls changes the pattern only marginally (Figure 2b, red line).

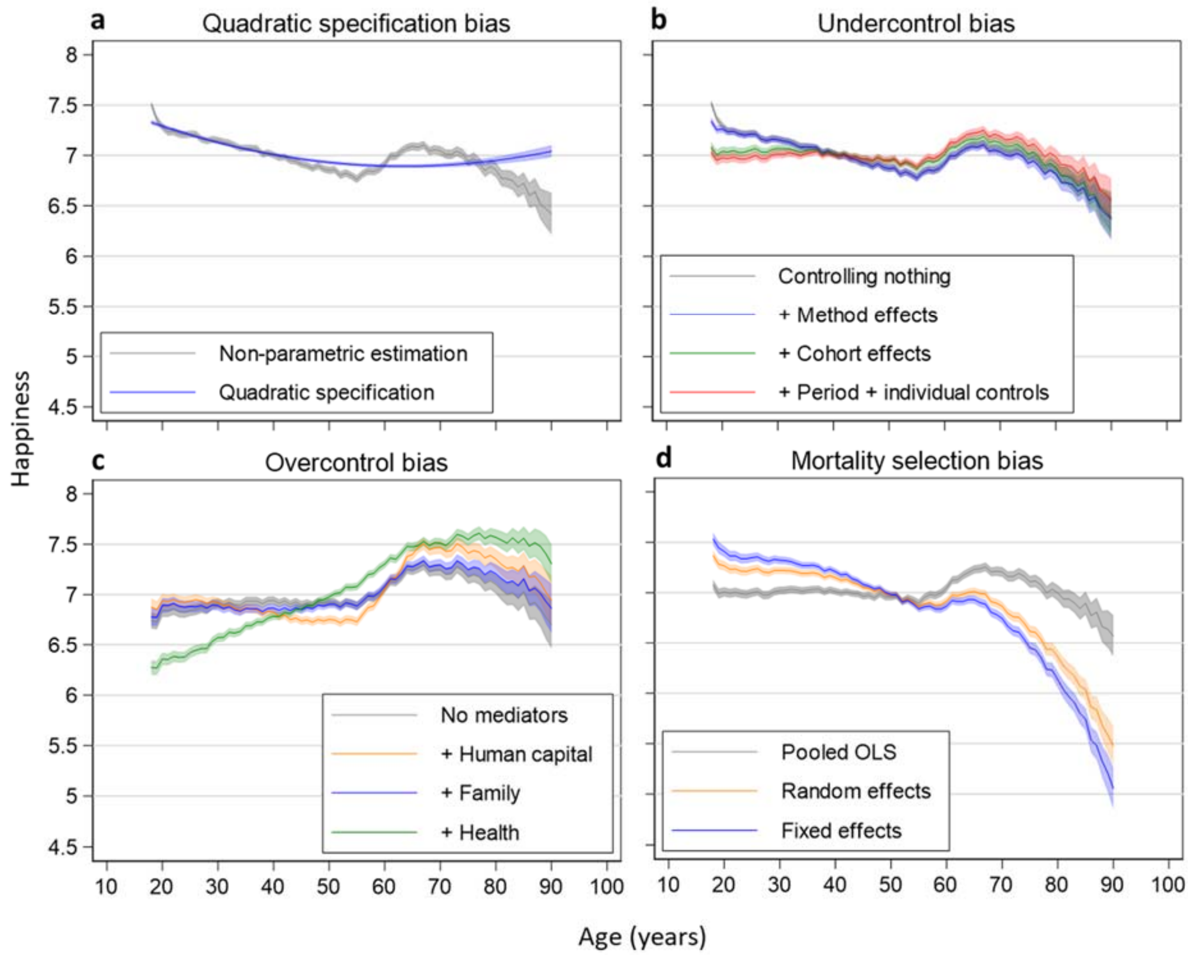


Fig. 2 | Predicted age-SWB trajectories (including 95%-CIs) resulting from different model specifications.

a, Quadratic specification bias: comparing a non-parametric pooled OLS trajectory (grey line) with a quadratic pooled OLS trajectory (blue line). Both models do not control for confounders. Supplementary Table 3 shows estimation details.

b, Undercontrol bias: comparing non-parametric pooled OLS trajectories when stepwise including confounding variables: Controlling nothing (grey line), plus method variables (blue line), plus cohort dummies (green line), plus period and individual variables (red line). Supplementary Table 4 shows estimation details.

c, Overcontrol bias: comparing non-parametric pooled OLS trajectories (controlling for confounders, reduced sample) when including mediating variables: no mediators included (grey line), only human capital variables included (orange line), only family event variables included (blue line), only health variables included (green line). Supplementary Table 5 shows estimation details.

d, Mortality selection bias: comparing non-parametric trajectories resulting from different regression models (controlling for confounders): pooled OLS model (grey line), random effects model (orange line), fixed effects model (blue line). Supplementary Table 6 shows estimation details.

The importance of overcontrol bias by erroneously including human capital variables, family events, and health is assessed in Figure 2c. Including human capital variables pulls the SWB trajectory down during adulthood, thereby increasing the slope of the downswing during adulthood until midlife. This creates the appearance of a midlife crisis (Figure 2c, orange line). Including family events affects the trajectory only marginally. Finally, including health variables results in an increasing age-SWB profile (Figure 2c, green line). These overcontrol biases are strong and affect even qualitative conclusions.

Figure 2d summarizes the consequences of mortality selection bias. The pooled OLS trajectory (grey line) is flat during adulthood, shows an upwards jump in the golden ages, and slowly starts to decline in old age. As argued previously, this is a descriptively correct representation of the age-specific SWB values in the SOEP data. However, it would be erroneous to interpret this curve causally as the effect of aging (as is done by many researchers). As a representation of the causal effect, this curve is heavily biased upwards because it is based in large parts on between-person variation that is "contaminated" by mortality selection. The more the estimation approach relies on within-person variation, the steeper declining the curve becomes. The random effects trajectory and in particular the fixed effects trajectory shows that SWB (almost) monotonically declines with age.

How Mis-Specified Models Can Produce (Almost) Any Result

To demonstrate that the conflicting findings reported in the literature indeed relate to mis-specified models and not to country, time, or data peculiarities, we show that we can reproduce almost any stylized age-happiness trajectory with our data and a combination of the four statistical pitfalls (Figure 3). In analogy to Simmons, Nelson, & Simonsohn (2011) who demonstrated in their famous paper how by (mis-) using analytical flexibility anything can be proven significant, we show here how differently specified models can "prove" any age-happiness trajectory.

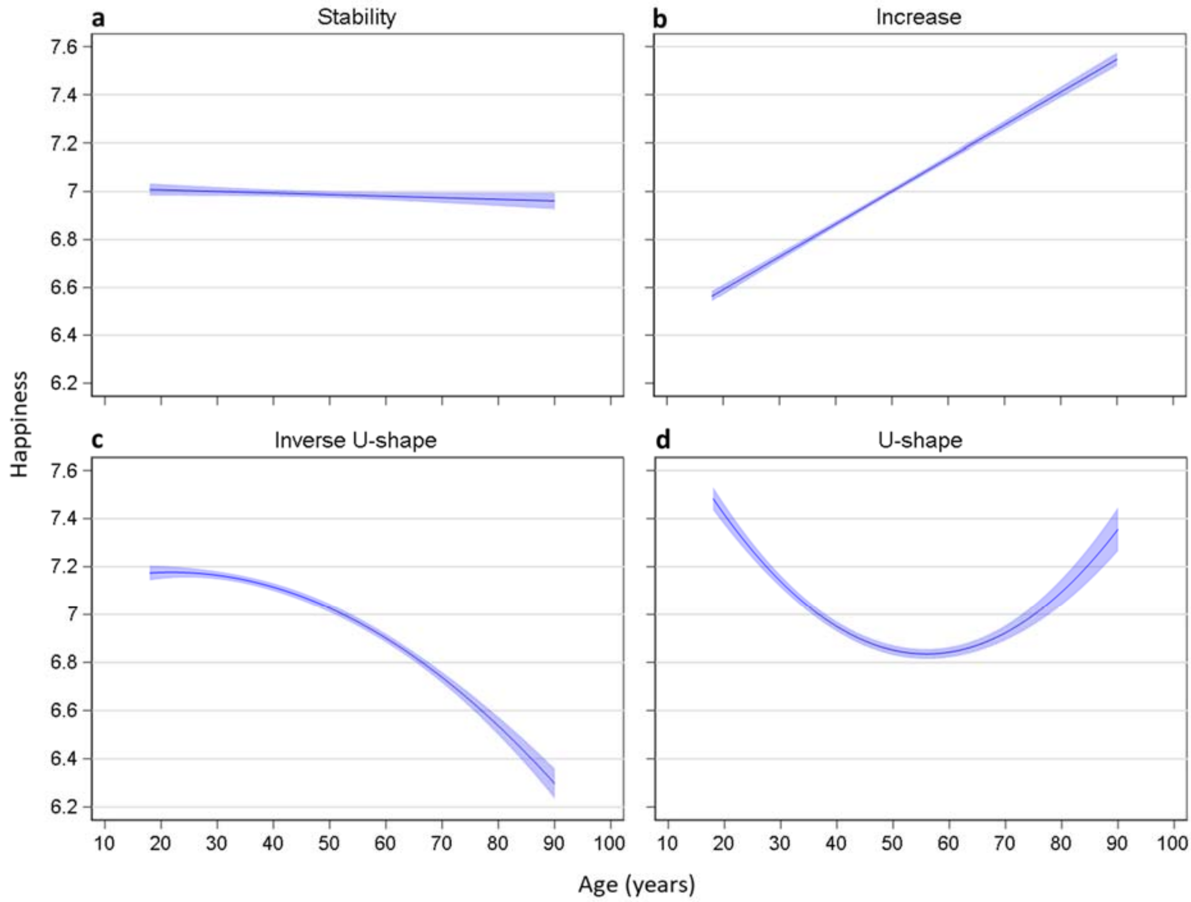


Fig. 3 | Predicted age-SWB trajectories resulting from mis-specified models.
a, Stability: pooled OLS regression (linear age specification) including confounding variables and objective health measures (i.e., disability status and hospital nights),
b, Increase: pooled OLS regression (linear age specification) including confounding variables, objective health measures (i.e., disability status and hospital nights), and self-reported health
c, Inverse U-shape: Random effects regression (quadratic age specification) controlling for confounding variables (except cohort).
d, U-shape: OLS regression (quadratic age specification) including family events, human capital, and objective health.
Supplementary Table 7 shows estimation details.

To replicate the finding of stability, we must pull the trajectory up at older ages such that the age-SWB profile flattens. Possible strategies include controlling for health (overcontrol bias) or using between-person variation such that mortality selection can pull the profile up, especially during oldest age. Figure 3a shows a combination of overcontrol (i.e., including objective health measures) and mortality selection bias (i.e., using an OLS regression).

To replicate an increase in SWB at oldest age, we additionally include self-reported health into the model to pull up the trajectory even further, according to the approach of prominent studies that found an increase (Figure 3b) (Yang, 2008).

Replicating an inverse U-shape is not entirely possible, as none of the four biases pulls the SWB curve down at young age. We therefore approximated the inverse U-shape (at least the linear age term is significantly positive and the squared age term significantly negative, see Supplementary Table 7) by using a random effects regression without controlling for cohort (Figure 3c).

In contrast, replicating a U-shape is straightforward, as each of the four biases favors a U-shape. Social desirability bias and not accounting for cohort effects increase the slope of the SWB downswing during adulthood until midlife. Similarly, erroneously including family events and human capital variables also increases the slope during this phase. In addition, erroneously including health measures pulls the curve up after age 30, and especially during the golden ages and old age. Finally, as a part of mortality selection bias, the more between-person variation used to estimate the SWB trajectory, the flatter it becomes with a more pronounced upswing after age 55. All studies that found a U-shape fall into at least one, if not several of the above outlined pitfalls (see Supplementary Table 1). We generated a U-shaped trajectory by combining all four biases (Figure 3d). To increase the downswing until midlife, we do not control for method or cohort effects (undercontrol bias) and control for family events and human capital (overcontrol bias). To pull the curve up at oldest age, we control for objective health (overcontrol bias) and use an OLS regression (mortality selection bias). This approach reflects the analytical strategy of the most famous age-SWB research: Blanchflower/Oswald 2008 (Blanchflower & Oswald, 2008) (1,800 Google Scholar citations, 22.02.2021).

Our Preferred Result on the Age-Happiness Trajectory in Germany

Now, what is our answer to our research question? The blue line in Figure 2d shows the result of our preferred specification: it uses a flexible non-parametric specification, it controls for important confounders, it does not control for mediators, and it relies on within-person variation only (by using fixed effects regression). Thus, this trajectory gives our best answer to the causal question on how aging affects happiness (in Germany 1984-2017).

For better visibility we duplicate our preferred trajectory separately in Figure 4a. Happiness declines slowly over adulthood (altogether about half a scale point). A low point is reached at the end of the 50ies.[5] Afterwards happiness increases slightly during the golden ages (about a

[5] More exactly the low is obtained at ages 55, 58, and 59. A postestimation F-test indicates that these age coefficients do not differ significantly from each other ß[55] versus ß[58]: [F(1, 70921) = 0.01, Prob > F = 0.9111]; ß[55] versus ß[59]: [F(1, 70921) = 0.01, Prob > F = 0.9061]; and ß[58] versus ß[59]: [F(1, 70921) = 0.00, Prob > F = 0.9891].

tenth of a scale point). The golden age high point is reached with age 64. Afterwards – during old age – a very steep decline in SWB sets in.[6]

Are these trends statistically significant? For a formal test we specify a spline trajectory: the first spline is from age 18 till age 24, the second from age 25 till age 34, and so on in ten-year intervals. The estimated spline trajectory is plotted in Figure 4b. Generally, it confirms well the qualitative pattern that we observe in the non-parametric trajectory in Figure 4a. Further, since all spline-slopes are statistically significant (see Supplementary Table 8) this confirms our qualitative finding: The age-happiness trajectory shows a slow decline through adulthood, a slight increase in golden ages, and a steep decline in old age. The spline-slope estimate for the golden ages is +0.0098 per year, meaning that happiness increases during the golden ages by 0.098 scale points.

Our finding is closest to the stylized finding of a "wave-like" trajectory as it has been reported in several studies (e.g., Baird et al., 2010; Bauer et al., 2017; Frijters & Beatton, 2012; McAdams et al., 2012; Wunder et al., 2013). These studies also used long-running panel data (SOEP, BHPS from the UK, HILDA from Australia) and random or fixed effects regression. Though no study used all elements of our preferred specification, they came more or less close to it. Thus, our estimated age-happiness trajectory does not appear to be peculiar.[7]

Our finding certainly does not support the notion of a "U-curve". At least not in the sense it is often conceived: There is no basis on which one could infer that people in old age become happier again. Even the small increase during the golden ages could hardly be seen as proof of the U-curve: happiness during golden ages remains well below happiness in the 20ies. And the sharp decline follows instantly! Only by superficially fitting a quadratic trajectory and by narrowly focusing on the significance of the quadratic term, one could arrive at the erroneous conclusion that happiness follows a U-curve (Jebb et al., 2020).

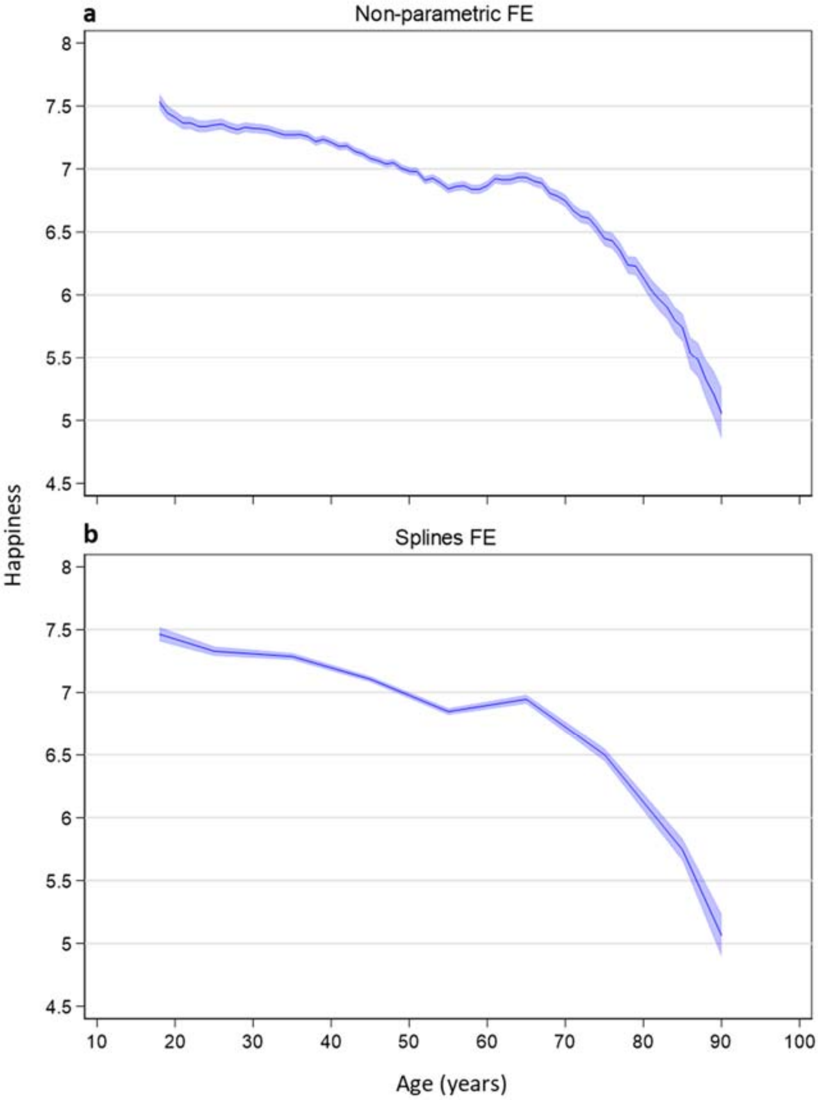
Further, we would argue that the happiness nadir around age 58 hardly can be interpreted as indication of a midlife crisis, because the decline in happiness is gradual since the 20ies. It makes more sense to interpret this pattern as a temporary halt in the gradual decline in the end of the 50ies, followed by an increase in happiness during the golden ages. We think this pattern cannot be explained reasonably by a midlife crisis. Instead, there is something going on around age 60 that (temporarily) makes people happier again. It is this increase that needs explanation.

[6] We think it is noteworthy how smooth the trajectory is: Since we simply plot the estimated age dummy effects without any smoothing, one would expect the trajectory to show a much more irregular pattern. The survey question on "life satisfaction" that certainly produces some measurement error, nevertheless seems to be able to reproduce some kind of regularity.

[7] This pattern is not really "wave-like" (see Figure 1). To us, it shows remarkable similarity with a ski-jumping hill. Thus, if one wants, one could refer to this pattern as the "ski-jump pattern".

Some scholars term this increase “retirement hump” (Wunder et al., 2013) arguing that it is (anticipation of) retirement that makes people happier again during the golden ages. This issue has not yet been settled. In the last section we will provide some more thoughts on the mechanisms that may have generated the age-happiness trajectory as reported in Figure 4.

Fig. 4 | Predicted age-happiness trajectories (including 95%-CIs) resulting from our preferred specification.



a, Predicted happiness values resulting from a fixed effects model. Supplementary Table 6 (model 3) shows estimation details.
b, Predicted happiness values resulting from a spline fixed effects model. Supplementary Table 8 shows estimation details.

Summary and Conclusions

How aging affects happiness is an important research question for the social and behavioral sciences. Our literature review demonstrates that many conflicting age trajectories have been reported in the literature. As this state of research is quite unsettling for the science of happiness, we discuss—informed by recent advances in the methodology of causal analysis—model specifications used by researchers in this field. Altogether, we identify four main biases that may distort the age trajectory of happiness. By using the German SOEP data, we show that distortions may be huge producing even qualitatively different conclusions. We demonstrate that by using different combinations of mis-specifications it is possible to generate (almost) every trajectory that has been reported in the literature. With a model specification that avoids these four biases, we find an age-happiness trajectory that declines slowly over adulthood (altogether about half a scale point). The decline comes to a halt and we observe even a small increase (about one tenth of a scale point) during the golden ages. Afterwards, in old age a very steep decline in happiness sets in.

From these results we derive several conclusions that pertain to future research on happiness. The overarching conclusion is that SWB scholars should take causal reasoning seriously in their future research. They should precisely define their research question and explicitly justify their model specification chosen according to the research question (these conclusions do not pertain to SWB research alone, but to all kind of social research as Lundberg et al. (forthcoming) argue forcefully).

Qualifying the research question before estimating age-SWB profiles is essential. Is the main research aim to describe how happy the living population is, or how SWB develops with rising age? If the aim is to answer the second research question about aging and thus to estimate a causal effect of age on SWB, scholars should not use (repeated) cross-sectional data, because these may be affected by mortality selection bias. And there is no cure for this with only cross-sectional data available. Only with panel data following the same respondents over time mortality selection bias can be fixed.

Even when using panel data, scholars must carefully consider potential sources of under- and overcontrol bias and select an estimation approach that strictly relies on within-person variation to minimize mortality selection bias. In our empirical illustration with the SOEP, the less familiar sources of bias, (i.e., overcontrol and mortality selection bias) cause more severe distortions than does undercontrol bias. We illustrated that selective mortality exhibits drastic consequences on the association between age and subjective well-being affecting even qualitative conclusions: Mortality selection systematically removes the unhappiest of the oldest old and therefore every approach that relies somehow on between persons variation underestimated the deteriorating impact of aging on subjective well-being (especially among the oldest old).

Avoiding these misspecifications is not only important for future research on the age trajectory of happiness but also for any kind of happiness research. Age is usually (and well justified) used as a control variable when investigating other determinants of happiness. Using mis-specified age trajectories can severely bias estimates of such treatment effects: the bias in the control variable age transfers to the treatment effect of interest (a formal statement of this so-called "bias transfer" can be found in Ranjbar & Sperlich, 2019). Therefore, it is important to use flexible parametrizations of the age effect in happiness research more generally. Including linear and quadratic age terms only might be problematic.

Limitations of Our Study and How We Could Move on

In this final section we want to hint towards some limitations of our study and give some suggestions on how these could be overcome by future studies. Although we used the longest-running panel survey with information on SWB, the data are limited to Germany. We can therefore not draw conclusions whether different societal settings bring about other age-SWB profiles. However, we can be certain that if scholars want to answer the causal research question of "How does aging affect happiness?", they must rely on (internationally comparable) panel data, as cross-sectional data do not allow for avoiding the many distortions that may affect qualitative conclusions.

Another limitation of our data is that it includes only respondents from private households. However, endogenous selection of the survey population may also occur because the unhappiest persons among the oldest old have higher probabilities of living in nursing homes or are in hospitals. We expect that studies excluding this population – such as ours – will underestimate the age-SWB decline in old age. Likewise, in populations with high migration rates, scholars must very carefully assess how endogenous population change due to in- or out-migration may affect age trajectories. The endogenous selection here will depend on the types of outmigration. Economic migrants are favorably self-selected with respect to happiness increasing characteristics (Author citation, 2020). Thus, the deteriorating impact of aging on subjective well-being may be overestimated in areas with high economic out-migration. It is important that future work also takes regard of such endogenous selection processes, e.g., by taking efforts to collect data from such populations.

Certainly, our list of potential confounding variables is not exhaustive. For example, if a spatial correlation of people of certain ages exists and the survey design results in a pattern of certain age groups being interviewed on certain week days (Seresinhe et al., 2019) or in certain weather conditions (Lucas & Lawless, 2013), these variables also may form part of an undercontrol bias. We invite further interdisciplinary research to complete the list of potential confounding variables.

Our preferred specification includes using fixed effects regression. Fixed effects models clearly rely on weaker assumptions than approaches using between-person variation: if low SWB levels result in premature mortality any between approach will provide biased results, whereas a fixed effects approach will work. If, however, mortality selection also operates on SWB trajectories, results from fixed effects models also will be biased. If decreasing SWB values lead to earlier mortality (beyond low SWB levels), the age-SWB profile slope will be underestimated. Thus, while we cannot be sure if the FE results are completely unbiased, we can be sure that a violation of this strict exogeneity assumption will not result in conflicting findings (i.e., stability, an increase, or a U-shape).

In defense against our argument of overcontrol bias, scholars could argue that our findings focus on the total effect of aging, whilst their results focused on the direct effect of aging net of some mediating variables. However, the identification of direct effects requires even stronger assumptions than the identification of total effects (VanderWeele, 2015), and is thus even more ambitious. To understand the mechanisms behind the total effect, one must first estimate the total effect and then separate direct from indirect effects in subsequent steps. Our paper addresses solely the first step, because first of all it is important to get the total effect right.

Starting from this, however, the next step should be studies on the mechanisms behind the age effect. As Galambos et al. (2020) argue such a research agenda could result in great scholarly benefits. Specifying and testing mechanisms could greatly contribute to a more cumulative knowledge progression in the field. Progress would then be achieved by generating an integrated web of mechanisms through interdisciplinary cooperation (Helliwell & Aknin, 2018).

In the following, we will provide some examples for such a research agenda on the age-related mechanisms that generate the SWB trajectory. To find answers on mechanisms, scholars need to draw on a potential outcome framework asking counterfactual questions. Then they should define total, indirect and direct effects. Make the assumptions for total, indirect and direct effects explicit and then start from a total effect, or even a life course specific total age effect and then explain this effect. For example, to explain the impact of aging on subjective well-being during adulthood (18-54) there may be the following mechanisms: career prospects such as employment or income increase until midlife whereas physical health decreases steadily. Thus, we can articulate the following expectations: growing (satisfaction with the) economic situation increases subjective well-being until midlife whereas decreasing health diminishes subjective well-being. Including the economic situation in the models answers the question how subjective well-being until midlife would develop if the economic situation remained constant. Thus, the age-happiness trajectory will become more negative/or less positive when we include the economic situation as control variable.

Likewise, if we ask the counterfactual research question how would aging affect subjective well-being if health remained at the level of an 18-year-old, for sure the effect of aging would be less negative/or more positive. Such clearly formulated expectations can be tested using panel data and an approach that relies solely on individual change. During golden ages (55-64) the halt/or increase in subjective well-being may for example relate to retirement or the birth of grandchildren, these mechanisms could also be tested directly. In the best-case using anticipation and adaptation effects and exploiting the fact that these critical life events happen for different individuals at different ages. During oldest age, the mechanisms that buffer against the steep decrease of subjective-well-being at the end of life should be assessed. How do different forms of care affect the subjective well-being of the oldest old that approach death? What is the role of distance to close relatives or social inclusion in networks of close friendship?

In any case, the assessment of these mechanisms must start from a (life course specific) total effect of aging followed by a clear definition of mechanisms as indirect effects and an expectation about the remaining direct effect. The research praxis of controlling for life-course specific circumstances available in the data and then speculating about which mechanisms bring about the residual age-subjective well-being profile is highly speculative and selectively pulls the age-subjective well-being profile down during midlife (because the causes that make people happy during midlife are better captured in most surveys than what makes people happy during early adulthood).

Bringing these arguments on potential mechanisms together, we suggest the following explanations for the age-happiness trajectory that we found in this study: The slow decline in happiness during adulthood may result from compensating mechanisms: favorable economic and family events might increase happiness during adulthood, but steadily declining health might have opposite effects. The health effect dominates resulting in a slowly declining trajectory. The halt/increase in SWB during golden ages may be the result of (anticipation of) retirement. Ever stronger deteriorating health may finally dominate and generate the steep decline in old age. These explanations could be tested along the lines described above.

Finally, the specification suggested in this paper aims at estimating an average SWB trajectory. We argue that this finding is important, because it informs us about a basic feature of the "conditio humana". Nevertheless, one could argue (as for instance Galambos et al. (2020) do) that this is a limited focus. Individual trajectories will deviate from the average trajectory to an unknown degree. In an extreme case, nobody will follow the average trajectory. We admit that the degree of diversity in individual SWB trajectories is an important question for future research (an early study in this direction is Baetschmann, 2013). At the same time, causal reasoning as suggested in this paper will also be of great importance to identify the diversity of happiness trajectories.

Data Note

The SOEP data can be ordered here:

https://www.diw.de/en/diw_01.c.601584.en/data_access.html

Our analysis code will be made available after publication.

References

Baetschmann, G. (2013). Heterogeneity in the Relationship between Happiness and Age: Evidence from the German Socio-Economic Panel. *German Economic Review*, *15*(3), 393–410. https://doi.org/10.1111/geer.12015

Baird, B. M., Lucas, R. E., & Donnellan, M. B. (2010). Life Satisfaction Across the Lifespan: Findings from Two Nationally Representative Panel Studies. *Social Indicators Research*, *99*(2), 183–203. https://doi.org/10.1007/s11205-010-9584-9

Bauer, J. M., Levin, V., Munoz Boudet, A. M., Nie, P., & Sousa-Poza, A. (2017). Subjective Well-being Across the Lifespan in Europe and Central Asia. *Journal of Population Ageing*, *10*(2), 125–158. https://doi.org/10.1007/s12062-016-9148-0

Bell, A., & Jones, K. (2014a). Another 'futile quest'? A simulation study of Yang and Land's Hierarchical Age-Period-Cohort model. *Demographic Research*, *30*, 333–360.

Bell, A., & Jones, K. (2014b). Current practice in the modelling of age, period and cohort effects with panel data: a commentary on Tawfik et al. (2012), Clarke et al. (2009), and McCulloch (2012). *Quality & Quantity*, *48*(4), 2089–2095. https://doi.org/10.1007/s11135-013-9881-x

Bell, A., & Jones, K. (2015). Should age-period-cohort analysts accept innovation without scrutiny? A response to Reither, Masters, Yang, Powers, Zheng and Land. *Social Science & Medicine*, *128*, 331–333. https://doi.org/10.1016/j.socscimed.2015.01.040

Berkson, J. (1946). Limitations of the Application of Fourfold Table Analysis to Hospital Data. *Biometrics Bulletin*, *2*(3), 47–53. https://doi.org/10.2307/3002000

Blanchflower, D. G. (2020). *Is Happiness U-Shaped Everywhere? Age and Subjective Well-Being In 132 Countries*. Working Paper 26641. National Bureau of Economic Research.

Blanchflower, D. G., & Graham, C. L. (2020). *The Mid-Life Dip in Well-Being: Economists (Who Find It) Versus Psychologists (Who Don't)!* (No. w26888). National Bureau of Economic Research.

Blanchflower, D. G., & Oswald, A. J. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine*, *66*(8), 1733–1749. https://doi.org/10.1016/j.socscimed.2008.01.030

Blanchflower, D. G., & Oswald, A. J. (2009). The U-shape without controls: A response to Glenn. *Social Science & Medicine*, *69*(4), 486–488. https://doi.org/10.1016/j.socscimed.2009.05.022

Chadi, A. (2013). The role of interviewer encounters in panel responses on life satisfaction. *Economics Letters*, *121*(3), 550–554. https://doi.org/10.1016/j.econlet.2013.10.024

Chauvel, L., & Schröder, M. (2015). The Impact of Cohort Membership on Disposable Incomes in West Germany, France, and the United States. *European Sociological Review*, *31*(3), 298–311. https://doi.org/10.1093/esr/jcu091

Cheng, T. C., Powdthavee, N., & Oswald, A. J. (2017). Longitudinal Evidence for a Midlife Nadir in Human Well-Being: Results from Four Data Sets. *Economic Journal*, *127*(599), 126–142. https://doi.org/10.1111/ecoj.12256

De Ree, J., & Alessie, R. (2011). Life satisfaction and age: Dealing with underidentification in age-period-cohort models. *Social Science & Medicine*, *73*(1), 177–182. https://doi.org/10.1016/j.socscimed.2011.04.008

Di Tella, R., MacCulloch, R. J., & Oswald, A. J. (2001). Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. *American Economic Review*, *91*(1), 335–341. https://doi.org/10.1257/aer.91.1.335

Di Tella, R., MacCulloch, R. J., & Oswald, A. J. (2003). The Macroeconomics of Happiness. *Review of Economics and Statistics*, *85*(4), 809–827. https://doi.org/10.1162/003465303772815745

Diener, E., & Chan, M. Y. (2011). Happy People Live Longer: Subjective Well-Being Contributes to Health and Longevity. *Applied Psychology: Health and Well-Being*, *3*(1), 1–43. https://doi.org/10.1111/j.1758-0854.2010.01045.x

Diener, E., Oishi, S., & Tay, L. (2018). Advances in subjective well-being research. *Nature Human Behaviour*, *2*(4), 253–260.

Easterlin, R. A. (2006). Life cycle happiness and its sources: Intersections of psychology, economics, and demography. *Journal of Economic Psychology*, *27*(4), 463–482. https://doi.org/10.1016/j.joep.2006.05.002

Easterlin, R. A., & Plagnol, A. C. (2008). Life satisfaction and economic conditions in East and West Germany pre- and post-unification. *Journal of Economic Behavior & Organization*, *68*(3-4), 433–444. https://doi.org/10.1016/j.jebo.2008.06.009

Elwert, F. (2013). Graphical Causal Models. In S. L. Morgan (Ed.), *Handbooks of Sociology and Social Research. Handbook of Causal Analysis for Social Research* (Vol. 51, pp. 245–273). Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_13

Elwert, F., & Winship, C. (2014). Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. *Annual Review of Sociology*, *40*, 31–53. https://doi.org/10.1146/annurev-soc-071913-043455

Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. Perspectives on Psychological Science, 7(6), 555-561. https://doi.org/10.1177/1745691612459059

Ferrer-i-Carbonell, A., & Frijters, P. (2004). How Important is Methodology for the Estimates of the Determinants of Happiness? *The Economic Journal*, *114*(497), 641–659. https://doi.org/10.1111/j.1468-0297.2004.00235.x

Fienberg, S. E. (2013). Cohort analysis' unholy quest: A discussion. *Demography*, *50*(6), 1981-4; discussion 1985-8. https://doi.org/10.1007/s13524-013-0251-z

Frijters, P., & Beatton, T. (2012). The mystery of the U-shaped relationship between happiness and age. *Journal of Economic Behavior & Organization*, *82*(2-3), 525–542. https://doi.org/10.1016/j.jebo.2012.03.008

Frijters, P., Haisken-DeNew, J. P., & Shields, M. A. (2004). Money Does Matter! Evidence from Increasing Real Income and Life Satisfaction in East Germany Following Reunification. *American Economic Review*, *94*(3), 730–740. https://doi.org/10.1257/0002828041464551

Galambos, N. L., Krahn, H. J., Johnson, M. D., & Lachman, M. E. (2020). The U Shape of Happiness Across the Life Course: Expanding the Discussion. Perspectives on Psychological Science, 15(4), 1-15. https://doi.org/10.1177/1745691620902428

Gerstorf, D., Ram, N., Mayraz, G., Hidajat, M., Lindenberger, U., Wagner, G. G., & Schupp, J. (2010). Late-life decline in well-being across adulthood in Germany, the United Kingdom, and the United States: Something is seriously wrong at the end of life. *Psychology and Aging*, *25*(2), 477–485. https://doi.org/10.1037/a0017543

Glenn, N. D. (2005). *Cohort analysis* (second edition). *Sage University papers Quantitative applications in the social sciences: Vol. 5*. Sage Publications. http://www.loc.gov/catdir/enhancements/fy0657/2004019523-d.html

Glenn, N. D. (2009). Is the apparent U-shape of well-being over the life course a result of inappropriate use of control variables? A commentary on Blanchflower and Oswald (66: 8, 2008, 1733-1749). *Social Science & Medicine*, *69*(4), 481-485. https://doi.org/10.1016/j.socscimed.2009.05.038

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The taboo against explicit causal inference in nonexperimental psychology. Perspectives on Psychological Science, 15(5), 1243-1255. https://doi.org/10.1177/1745691620921521

Heckman, J., & Robb, R. (1985). Using Longitudinal Data to Estimate Age, Period and Cohort Effects in Earnings Equations. In W. M. Mason & S. E. Fienberg (Eds.), *Cohort Analysis in Social Research* (Vol. 51, pp. 137–150). Springer New York. https://doi.org/10.1007/978-1-4613-8536-3_5

Heintzelman, S. J., Trent, J., & King, L. A. (2015). Revisiting desirable response bias in well-being reports. The Journal of Positive Psychology, 10(2), 167-178. https://doi.org/10.1080/17439760.2014.927903

Helliwell, J. F., & Aknin, L. B. (2018). Expanding the social science of happiness. *Nature Human Behaviour*, *2*(4), 248–252. https://doi.org/10.1038/s41562-018-0308-5

Jebb, A. T., Tay, L., Diener, E., & Oishi, S. (2018). Happiness, income satiation and turning points around the world. *Nature Human Behaviour*, *2*(1), 33–38. https://doi.org/10.1038/s41562-017-0277-0

Jebb, A. T., Morrison, M., Tay, L., & Diener, E. (2020). Subjective Well-Being Around the World: Trends and Predictors Across the Life Span. Psychological Science, 31(3), 293-305. https://doi.org/10.1177/0956797619898826

Kassenboehmer, S. C., & Haisken-DeNew, J. P. (2012). Heresy or enlightenment? The well-being age U-shape effect is flat. *Economics Letters*, *117*(1), 235–238. https://doi.org/10.1016/j.econlet.2012.05.013

Keyes, C. L. M., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: The empirical encounter of two traditions. *Journal of Personality and Social Psychology*, *82*(6), 1007–1022. https://doi.org/10.1037/0022-3514.82.6.1007

Kunzmann, U., Little, T. D., & Smith, J. (2000). Is age-related stability of subjective well-being a paradox? Cross-sectional and longitudinal evidence from the Berlin Aging Study. *Psychology and Aging*, *15*(3), 511–526. https://doi.org/10.1037//0882-7974.15.3.511

Lucas, R. E., & Lawless, N. M. (2013). Does life seem better on a sunny day? Examining the association between daily weather conditions and life satisfaction judgments. *Journal of Personality and Social Psychology*, *104*(5), 872–884. https://doi.org/10.1037/a0032124

Lundberg, I., R. Johnson and B.M Stewart (Forthcoming), 'What is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory', American Sociological Review. Accessed 25 February 2021 at SocArXiv, https://doi.org/10.31235/osf.io/ba67n.

Luo, L. (2013a). Assessing validity and application scope of the intrinsic estimator approach to the age-period-cohort problem. *Demography*, *50*(6), 1945–1967. https://doi.org/10.1007/s13524-013-0243-z

Luo, L. (2013b). Paradigm Shift in Age-Period-Cohort Analysis: A Response to Yang and Land, O'Brien, Held and Riebler, and Fienberg. *Demography*, *50*(6), 1985–1988. https://doi.org/10.1007/s13524-013-0263-8

Luo, L., & Hodges, J. S. (2016). Block Constraints in Age–Period–Cohort Models with Unequal-width Intervals. *Sociological Methods & Research*, *45*(4), 700–726. https://doi.org/10.1177/0049124115585359

Lykken, D., & Tellegen, A. (1996). Happiness Is a Stochastic Phenomenon. *Psychological Science*, *7*(3), 186–189. https://doi.org/10.1111/j.1467-9280.1996.tb00355.x

McAdams, K. K., Lucas, R. E., & Donnellan, M. B. (2012). The Role of Domain Satisfaction in Explaining the Paradoxical Association Between Life Satisfaction and Age. *Social Indicators Research*, *109*(2), 295–303. https://doi.org/10.1007/s11205-011-9903-9

Pelzer, B., te Grotenhuis, M., Eisinga, R., & Schmidt-Catran, A. W. (2015). The Non-uniqueness Property of the Intrinsic Estimator in APC Models. *Demography*, *52*(1), 315–327. https://doi.org/10.1007/s13524-014-0360-3

Ranjbar, S., & Sperlich, S. (2019). A Note on Empirical Studies of Life-Satisfaction: Unhappy with Semiparametrics? *Journal of Happiness Studies.* Advance online publication. https://doi.org/10.1007/s10902-019-00165-z

Realo, A., & Dobewall, H. (2011). Does life satisfaction change with age? A comparison of Estonia, Finland, Latvia, and Sweden. *Journal of Research in Personality*, *45*(3), 297–308. https://doi.org/10.1016/j.jrp.2011.03.004

Rohwer, G. (2016). A Note on the Dependence of Health on Age and Education. *Demography*, *53*(2), 325–335. https://doi.org/10.1007/s13524-016-0457-y

Safi, M. (2010). Immigrants' Life Satisfaction in Europe: Between Assimilation and Discrimination. *European Sociological Review*, *26*(2), 159–176. https://doi.org/10.1093/esr/jcp013

Schupp, J., Goebel, J., Kroh, M., & Wagner, G. G. (2013). Life Satisfaction in Germany at Highest Levels since Reunification. In S. Gerstorf & J. Schupp (Eds.), *SOEP Wave Report* (pp. 29–37). https://www.diw.de/de/diw_01.c.471201.de/sonstige_diw_publikationen/life_satisfaction_in_germany_at_highest_levels_since_reunification.html

Seresinhe, C. I., Preis, T., MacKerron, G., & Moat, H. S. (2019). Happiness is Greater in More Scenic Locations. *Scientific Reports*, *9*(1), 4498. https://doi.org/10.1038/s41598-019-40854-6

Simmons, J.P., L.D. Nelson, and U. Simonsohn (2011) False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. Psychological Science, 22(11): 1359-1366.

Socio-Economic Panel (2019), data for years 1984-2017, version 34. doi:10.5684/soep.v34.

Stone, A. A., Schwartz, J. E., Broderick, J. E., & Deaton, A. (2010). A snapshot of the age distribution of psychological well-being in the United States. *Proceedings of the National Academy of Sciences*, *107*(22), 9985–9990. https://doi.org/10.1073/pnas.1003744107

Sutin, A. R., Terracciano, A., Milaneschi, Y., An, Y., Ferrucci, L., & Zonderman, A. B. (2013). The effect of birth cohort on well-being: The legacy of economic hard times. *Psychological Science*, *24*(3), 379–385. https://doi.org/10.1177/0956797612459658

Ulloa, B. F. L., Møller, V., & Sousa-Poza, A. (2013). How does subjective well-being evolve with age? A literature review. *Journal of Population Ageing*, 6(3), 227-246. https://doi.org/10.1007/s12062-013-9085-0

VanderWeele, T. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press. https://ebookcentral.proquest.com/lib/subhh/detail.action?docID=1920742

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. In Duncker & Humblot (Ed.), *Schmollers Jahrbuch: Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften* (pp. 139–169). Duncker & Humblot.

Weiss, A., King, J. E., Inoue-Murayama, M., Matsuzawa, T., & Oswald, A. J. (2012). Evidence for a midlife crisis in great apes consistent with the U-shape in human well-being. *Proceedings of the National Academy of Sciences*, *109*(49), 19949–19952. https://doi.org/10.1073/pnas.1212592109

Wolbring, T., Keuschnigg, M., & Negele, E. (2013). Needs, Comparisons, and Adaptation: The Importance of Relative Income for Life Satisfaction. *European Sociological Review*, *29*(1), 86–104. https://doi.org/10.1093/esr/jcr042

Wunder, C., Wiencierz, A., Schwarze, J., & Küchenhoff, H. (2013). Well-Being over the Life Span: Semiparametric Evidence from British and German Longitudinal Data. *Review of Economics and Statistics*, *95*(1), 154–167. https://doi.org/10.1162/REST_a_00222

Yang, Y. (2008). Social Inequalities in Happiness in the United States, 1972 to 2004: An Age-Period-Cohort Analysis. *American Sociological Review*, *73*(2), 204–226. https://doi.org/10.1177/000312240807300202