



# To Evaluate the Age–Happiness Relationship, Look Beyond Statistical Significance

David Bartram<sup>1</sup>

Accepted: 30 December 2023 / Published online: 7 February 2024  
© The Author(s) 2024

## Abstract

The persistent contentiousness of research on the age–happiness relationship is puzzling; it should be possible to gain clarity and consensus about how to address the question effectively. In this paper I show that a key reason for the lack of clarity consists of overreliance on statistical significance as a means of evaluating empirical results. The statistical significance of a quadratic specification (age plus age-squared) is often taken as evidence in support of a ‘u-shaped’ relationship between age and happiness. But statistical significance on its own cannot tell us whether the age–happiness relationship is ‘u-shaped’ (nor indeed whether it takes any other shape). On the contrary, statistical significance can mislead us about it: a set of quadratic age coefficients can be ‘significant’ even when the relationship is obviously characterised by a different shape. When we have clarity on how to construct the analysis so that we can ‘see’ the underlying patterns in the data, it becomes obvious that the age–happiness relationship in European countries commonly shows other patterns; a u-shape is evident only in a minority of countries.

**Keywords** Happiness · Ageing · Statistical significance · Control variables · Visual analysis

## 1 Introduction

A puzzling feature of recent research on happiness is the persistence of contention (and indeed confusion) about the impact of age. There is a widespread view that the age–happiness relationship is generally ‘u-shaped’, with a decrease in happiness as people become middle-aged and subsequently an increase. This view (and in particular its alleged universality) has been challenged in a number of recent contributions (e.g. Kratz & Brüderl, 2021; Bittmann, 2021; Morgan & O’Connor, 2017; Bartram, 2021, 2023; Hellevik, 2017; see also Glenn, 2009). But other scholars have rejected key points made in these critiques (Blanchflower et al., 2023 is a recent example) and continue to offer analyses that

---

The analysis syntax for this paper is available here: [https://osf.io/zpcxj/?view\\_only=e384bd25ac6f40eaa1e273cc6417184](https://osf.io/zpcxj/?view_only=e384bd25ac6f40eaa1e273cc6417184).

---

✉ David Bartram  
d.bartram@le.ac.uk

<sup>1</sup> University of Leicester, Leicester LE1 7RH, UK

support the idea of u-shapes. Meanwhile, happiness scholars in general reinforce the idea of u-shapes when they use age as a control variable in analyses intended for other purposes: the typical practice is to use a quadratic function and then to evaluate it solely via statistical significance.

Again, the lack of clarity evident on this topic is puzzling: in principle it should be possible to achieve consensus about the right way to address the question empirically—and then to come to a settled view about what the relationship actually is. The key contribution of this paper is to offer an explanation for the persistence of the contention/confusion. The main factor leading to incorrect conclusions (and in particular the idea that the age–happiness relationship is *universally* ‘u-shaped’) is: overreliance on statistical significance as a means of evaluating results and reaching substantive conclusions. At best, statistical significance can tell us whether our results, rooted in use of sample data, are likely to be found in the corresponding population—but it cannot help us evaluate whether the results themselves are *right*. That assertion applies in particular to the specification of a particular functional form (e.g. linear vs. quadratic), which is exactly what is in play here. The key point, demonstrated below, is that a model using an incorrect functional form *can nonetheless produce statistically significant coefficients*. In particular: in many countries the age–happiness relationship consists of a decline in happiness across the entire life course—but I will show that the coefficients in a quadratic specification are sometimes statistically significant even so. This apparent disjuncture arises especially when we use a sufficiently large sample.

The main substantive finding presented below is that, when we refrain from imposing a specific functional form and instead construct a visual (non-parametric) analysis that allows us to discern the underlying patterns, the age–happiness relationship takes on a wide variety of forms (shapes), across 30 different European countries (compare Bittmann, 2021 and Morgan & O’Connor, 2017). In a few countries there are patterns reasonably described with the letter U (in particular, showing a post-middle-age increase in happiness). But the much more common pattern (especially in eastern and southern Europe) is a *life-long decline* in happiness. In a couple of countries, the opposite pattern is evident—a *life-long increase* in happiness. A very common more specific component involves the decline of happiness in old age, even in some countries where there is indeed a (temporary) post-middle-age increase. The main finding again is that there is no universal pattern; instead, there is an obvious context dependency.

## 2 Previous Research

A great many happiness researchers embrace the view that happiness is u-shaped in age. That assertion is evidenced in part by the way age is used as a control variable for analyses oriented to some other purpose: the overwhelmingly common practice is to use a quadratic specification, entering age together with age-squared in the model. This practice is typically justified with reference to the idea that the relationship is in fact u-shaped—in other words, taking for granted that the answer to the question is known.

The influential early study that underpins the widespread adoption of the u-shape position is Blanchflower and Oswald (2008).<sup>1</sup> This work draws on a range of repeated cross-sectional datasets encompassing 72 countries; the core finding is that the age–happiness relationship is u-shaped in most of these. There are exceptions, mostly among some poorer countries (e.g. Bangladesh, Jordan, and Pakistan); the authors suggest that one reason for the departure might be the small sample sizes (an idea we will explore in more detail below). The u-shapes finding is derived mainly from models that impose a quadratic function and include a range of control variables.

This early study was challenged very quickly by Glenn (2009). Glenn considered whether the appearance of the u-shape resulted from use of ‘inappropriate’ control variables (the substance of this view is described below). Similar explorations appeared in Hellevik (2017) and Bartram (2021). In certain instances, these critiques were followed by rebuttals. A key example appeared in Blanchflower (2021), offering an analysis covering an even wider range of countries and using multiple datasets—and concluding that the age–happiness relationship is u-shaped virtually ‘everywhere’. A further critique by Bartram (2023) led to another rebuttal (Blanchflower et al., 2023), a contribution notable for a sweeping statement asserting that there are no fewer than 618 published studies finding that the age–happiness relationship is characterised by a u-shape.

More broadly, we see a range of studies that produce a range of patterns characterising the age–happiness relationship.<sup>2</sup> Several studies offer support for the ‘u-shape’ finding: in addition to Blanchflower’s contributions, there is work by Beja (2018), Graham and Ruiz Pozuelo (2017), Cheng et al. (2017), and Movshuk (2011). There is a second category of studies that agree with the idea of ‘u-shape’ in the sense that happiness apparently rises after a midlife low; the difference here is that these studies discern that happiness subsequently declines as people move through older age<sup>3</sup>—such that the overall pattern is a sideways ‘s-shape’ or ‘wave pattern’ (Biermann et al., 2022; Frijters and Beaton 2012; Laaksonen, 2018). Other investigations show greater inconsistency with the ‘u-shape’: for example, Kratz and Brüderl (2021) identify a declining trend in happiness across the life course in Germany (i.e., with no post-middle-age increase). In the study of Germany by Kassenboehmer and Haisken-DeNew (2012), the life-course trend is instead found to be flat. Galambos et al. (2015) find that happiness increases among younger Canadians as they approach middle age. Another set of studies stands in direct contrast to the idea of ‘u-shapes *everywhere*’, finding instead that the age–happiness relationship takes a variety of different forms in different countries (Bartram, 2023; Becker & Trautmann, 2022; Bittmann, 2021; Morgan & O’Connor, 2017).

The current situation, then, amounts to a remarkable absence of consensus. Different researchers adopt a range of different approaches in their investigations. Key points of difference include: (1) whether to use control variables (in particular, controls that are influenced by age); (2) whether to use cross-sectional data or to insist more stringently on use of panel data; (3) whether to adopt a priori a quadratic specification (or some other functional form, e.g. cubic), as against starting with a non-parametric approach. We can then ask:

<sup>1</sup> Per Google Scholar at time of writing, this publication has more than 2400 citations (a number that increases by approximately 200 per year).

<sup>2</sup> The very large size of this literature precludes an attempt at creating a comprehensive review. The studies reviewed here are intended as representative examples.

<sup>3</sup> For research on that specific question, see Hudomiet et al. (2021) and Jivraj et al. (2014).

why is there so much contention—not only about the result itself but about the appropriate methodological way to conduct an analysis intended to give us the right result?

I propose to answer that question by describing two inter-related components characterising the studies that consistently produce the u-shape result. (1) There is, in general, an over-reliance on statistical significance as a criterion for reaching substantive conclusions. Incorrect results can be statistically significant, especially when we use large samples. (2) Insofar as these studies construct models that include control variables, the consequence is ‘overcontrol bias’ (Rohrer 2018)—and the bias in this context leads us to overestimate any tendency of happiness to increase as people move beyond middle age. The increased size of coefficients rooted in overcontrol bias exacerbates the problem arising from misuse of statistical significance.

## 2.1 Over-Reliance on Statistical Significance

In research on this topic as on many others, quantitative researchers commonly evaluate their results mainly with reference to whether they are statistically significant. The key point here is that, if we evaluate our results *only* by considering whether they are statistically significant, we run a substantial risk of drawing faulty conclusions from our analysis (compare Wasserstein et al., 2019 and Carver, 1978).

At best, statistical significance could tell us whether our results, rooted in analysis of data from a sample, are likely to characterise the population from which the sample is drawn. As conventionally understood, a  $p$  value can be used to evaluate a hypothesis of some sort. If we start with the assumption that the corresponding null hypothesis ( $H_0$ ) is true, the  $p$  value ‘is the probability ... of [getting] a test statistic value at least as contradictory to  $H_0$  as the value actually observed’ via the data we are analysing (Agresti & Finlay, 1997, p. 157). If  $p$  is small, what many researchers then conclude is that the null hypothesis can be rejected—so, our results from sample data can then tell us that an effect of some sort is likely to be found in the population.<sup>4</sup>

We can then consider the conditions that must be met for statistical significance to succeed in giving us this information. The ‘assumptions’ described in any statistics textbook include, inter alia, having a representative sample and having confidence that the ‘error term’ is not correlated with independent variables in the model. But the more important assumption in this context is as follows: if we are going to use a linear model, then we must be confident that the relationship is in fact linear. Statistical significance ( $p < 0.05$ ) does not *tell* us that the relationship is linear. Instead, it enables us to extrapolate effectively from sample to population *only if we already know that the relationship is indeed linear*.

The point is universally articulated in textbooks with reference to linear regression—but it applies just as much to situations where a different functional form is specified. The functional form relevant here is quadratic, where age plus age-squared is entered in the model. The statistical significance of these coefficients cannot be used to tell us that the relationship is in fact u-shaped. Statistical significance can be used here to extrapolate effectively from sample to population only if we *already know* that the relationship is u-shaped. Using

<sup>4</sup> Many observers (e.g. Gorard 2016) identify a number of elisions and logical fallacies in the conventional use of ‘null-hypothesis significance testing’. The main upshot of those critiques is that use of  $p$  values cannot effectively be used to tell us that a particular hypothesis is true—not least because the entire operation requires the assumption that the corresponding (contradictory) null hypothesis is true.

statistical significance as evidence about the shape itself amounts to putting the cart before the horse.

Researchers might believe that the quadratic age coefficients would be statistically significant only if the relationship is in fact u-shaped. It is no doubt jarring to imagine that the age plus age-squared coefficients could be statistically significant if the relationship is *not* u-shaped. How would this be possible? Why might statistical significance mislead us in this way?

The answer is: sample size. With a sufficiently large sample, we can get statistically significant results ( $p < 0.05$ ) from an analysis that imposes a particular functional form even when that functional form does not effectively represent the underlying social process. Here we can gain insight by revisiting a very basic question: how is  $p$  determined?  $P$  is associated with  $t$ , which results from dividing the coefficient by its standard error. The standard error is determined in part by sample size. With a larger sample, we are more likely to get  $p < 0.05$ , simply because the standard error is smaller (so,  $t$  is larger and  $p$  is smaller). This is one core reason why statistical significance is not a sufficient way of reaching substantive conclusions, certainly not when the functional form of a relationship is in question. Geerling and Diener (2020) show how use of large samples can lead to statistically significant results even when effect sizes are very small. The point here is again more jarring: using large samples, we can get statistically significant results even when those results are clearly *incorrect*. This point is demonstrated empirically in the analysis below.

## 2.2 Use of Inappropriate Control Variables

As noted above, the early study by Blanchflower and Oswald (2008) relied mainly on models that include control variables. Glenn's critique (2009) described the use of those controls as 'inappropriate'. Blanchflower and Oswald (2009, 2019) rejected this view, asserting that use of controls constitutes a 'ceteris-paribus analytical approach', ostensibly more appropriate for characterising the age–happiness relationship. This terminology needs unpacking, so that we can gain clarity about the social reality underpinning the results our analysis creates (Martin, 2018). The most effective contribution in this context is Morgan and O'Connor (2017), arguing that an analysis without controls yields results that tell us what people actually *experience* as they grow older (note their term 'experienced life-cycle satisfaction').

In many instances of quantitative research, the use of controls would of course make sense. In general, the purpose of using control variables is to mitigate the possibility of bias in our results: we want to ensure that our estimates are neither too high nor too low (as an indication of the true effect, which is unknown except via estimates using data). We might observe a correlation between height and vocabulary size, but if we conclude that getting taller *leads to* an increase in the number of words someone can use, we overlook the way age (among children) is the real cause of both processes. Once we control for age, we get the right estimate of the effect size of getting taller (i.e., zero).

That example works because the control (age) is an antecedent of both variables. In general, to estimate  $X \rightarrow Y$  (the impact of  $X$  on  $Y$ ) without bias, we need controls ( $W$ ) that are antecedents of  $X$  and  $Y$  (so,  $W \rightarrow X$  and  $W \rightarrow Y$ ) (see e.g. Pearl, 2009). A genuine problem arises when a model includes (as controls) variables that are instead *influenced by* the focal independent variable ( $X \rightarrow W$ ). If we use these 'bad controls' (Angrist & Pischke,

2009), we exacerbate bias in our results, rather than mitigating bias.<sup>5</sup> Many researchers worry about ‘omitted variable bias’, which is indeed an important issue in general. But the possibility of ‘overcontrol bias’ is no less important (e.g. Elwert & Winship, 2014; Rohrer, 2018).

The relevant point in connection with the age–happiness relationship is twofold. (1) Apart from cohort and period, there are no antecedents of our focal independent variable (X) here, age (cf. Bittmann, 2021 and Kratz & Brüderl, 2021). Until they die, everyone keeps getting older, at the same rate, no matter what their other characteristics or circumstances are (Bartram, 2021). The only relevant controls are cohort and period, to address age–period–cohort concerns (see below). Other than those, there are no needed controls, i.e., variables where  $W \rightarrow X$ . (2) The real problem is that age is likely to influence controls pertaining to individual characteristics and/or circumstances (so,  $X \rightarrow W$ ). Use of other variables as controls is very likely to lead to overcontrol bias.

We can then consider the likely direction of overcontrol bias in substantive terms. Getting older might mean loss of one’s spouse, or declining health, or reduced income. What would it mean to control for health when estimating the age–happiness relationship (as in e.g. Laaksonen, 2018)? The result for age would tell us about the way happiness changes as someone becomes one year older while health is ‘held constant’. The difficulty is that age itself does not ‘hold health constant’. Age influences people’s health; it is a ‘bad control’ ( $X \rightarrow W$ ). If we control for health, we learn about the impact of age only for people who are fortunate enough not to experience declining health (in line with the fact that health is being held constant). The result for age then does not reflect the experience of people who *do* suffer from declining health. For them, ageing means becoming less healthy, which likely also means becoming less happy, relative to the happiness one might experience if health didn’t deteriorate (Jivraj et al., 2014; Steptoe, 2019). As a representation of how age affects happiness in general (especially with reference to the idea that happiness might increase after a ‘midlife low’), results from a model that includes health as a control are very likely to be upwardly biased, giving an exaggerated impression of any tendency for happiness to increase after middle age (compare Hellevik, 2017).<sup>6</sup>

The same pattern can be anticipated in an analysis that includes marital status as a control. As people age, their likelihood of being widowed increases. Controlling for marital status would tell us about the impact of age only for people who are fortunate enough not to lose their spouses. What about those who do lose their spouses? For them, the loss of spouse that comes with getting older is likely to mean lower happiness, relative to the happiness they would experience if they didn’t lose their spouses (Clark et al., 2018). As a representation of how age affects happiness in general, results from a model that includes marital status as a control are very likely to be upwardly biased.

<sup>5</sup> The reason to designate controls with the letter W is that W should ‘come before’ X. If we include controls where X comes before W (i.e., X influences W, as in  $X \rightarrow W$ ), we are doing the analysis incorrectly (‘bad controls’), leading to biased results. That statement applies to situations where we are estimating a *total* effect for  $X \rightarrow Y$ . In a mediation analysis, we might be interested in direct effects vs. indirect effects. But when we control for mediators, we then need to avoid confusing the direct effect with the total effect. The coefficient for X when a mediator is controlled is a direct effect, not the total effect.

<sup>6</sup> This overestimation is exactly what is apparent in work by Laaksonen (2018). In his Fig. 1, we can easily see what happens when health is added as a control: the line representing how happiness changes in older age goes sharply up, relative to the other lines. As a representation of what happens to people’s happiness in general as they become old, that line is substantially biased. The articulation of this pattern as consisting of bias is demonstrated in Kratz and Brüderl (2021).

In general, getting older entails the experience of loss, for many people. If we include ‘bad controls’ (and when age is the independent variable, virtually all controls are either irrelevant or ‘bad controls’), we are likely to get upwardly biased results for post-middle-age experiences, misleading us about the impact of age on happiness in general. There might well be countervailing processes contributing also to a tendency towards increased happiness as we age. But inclusion of controls, especially for the circumstances that amount to loss, leads to overcontrol bias in a predictable direction—i.e., upwards, overstating the extent of increased happiness in later life.

Blanchflower and Oswald’s position (e.g. 2019) is that results from a model with controls give us the ‘pure’ effect of ageing. But we need clarity on what this result has been *purified of*. Another way of articulating the points made above is as follows: when we control for factors that are themselves influenced by age, our result has been ‘purified’ of part of the effects *of age itself*. If we want to know what people actually experience as they grow older, we need to omit variables that are influenced by age.

Bartram (2023) demonstrates that the inclusion of typical ‘bad controls’ in a quadratic model of the age–happiness relationship results in a doubling of the age and age-squared coefficients. The error is a consequential one. It is especially consequential when statistical significance is used as the sole means of interpreting results and drawing substantive conclusions. Coefficients that are artificially inflated away from zero are more likely to be statistically significant. Biased results are therefore more likely to be perceived as *correct* results when evaluated with reference to statistical significance. The practice of using ‘bad controls’ exacerbates the tendency to misinterpret results via a focus on statistical significance.

### 2.3 Looking Forward

To gain clarity on the age–happiness relationship—in particular, its ‘shape’—we need an analysis that has the following features: (1) omission of any control variables that are influenced by age; (2) adoption of an analytical approach that starts with agnosticism about what the shape of the relationship might be; and (3) evaluation of results via a focus on the shapes that emerge from the data (rather than via statistical significance alone). I now turn to a description of an analysis that fits these criteria.

## 3 Data and Analysis

The analysis here uses data from the European Social Survey (ESS). These are repeated cross-sectional data, conducted with a consistent format across different countries and rigorous standards for sample selection (see e.g. Jowell, 2007). I use Rounds 1 through 9, corresponding to the period 2002 to 2020 (the survey is conducted bi-annually). Having data from a sufficiently broad time range is essential for the purpose of disentangling age patterns from cohort and period effects. Broadly, the age–period–cohort (APC) dilemma is rooted in the fact that each component forms a linear combination of the other two: age is equal to the difference between current year (period) and birth year (cohort) (see e.g. Fosse & Winship, 2019). In principle, if we draw conclusions on the basis of age differences alone (especially from single-wave cross-sectional data), we risk discerning age effects that in reality reflect ‘current’ temporal changes and/or the fact that older respondents were born in an earlier era. Buecker et al. (2023) and Morgan and O’Connor (2017) show that



cohort effects do not in fact ‘confound’ the age patterns evident in life satisfaction. I will demonstrate below that the same conclusion applies to ‘period’ effects. That demonstration requires use of data covering a sufficiently extended time-frame. I therefore include countries that have participated in at least two non-adjacent rounds of the ESS. Table 1 gives information about participating countries and key variables.

Some researchers would insist more stringently on use of panel data. The position here is that repeated cross-sectional data are sufficient. The main reason to insist on panel data starts with the apparent advantages of evaluating within-person change. But we can explore more precisely what the advantages actually are. A ‘within’ (a.k.a. ‘fixed effects’) analysis is generally more robust to the possibility of omitted-variable bias: the structure of within models ‘automatically’ controls for time-constant confounders, even the unmeasured ones. But we then return to the discussion of control variables above: there *are* no confounders of the age–happiness relationship, because no other variable is an antecedent of age. We don’t need to worry about unmeasured confounders (because we don’t need to worry about confounders more broadly)—so, we don’t need a ‘within’ analysis as a means of safeguarding against omitted variable bias from unmeasured confounders. There are other potential advantages of panel data: in particular, we could consider the way attrition rooted in mortality can influence results (e.g. Kratz & Brüderl, 2021). But there is also a cost associated with restricting our analysis to panel data: we would significantly narrow the range of countries that can be investigated (because many countries do not have panel data). Instead of living with that consequence (especially in an investigation focused on discerning the prevalence of different patterns in different countries), we can (and will) reflect on the way our results might be affected by use of cross-sectional as against longitudinal data.

### 3.1 Method of Analysis

The main analysis consists of ‘local polynomial regression’ fitting, used as a foundation for construction of visualised results (via the R-package `ggplot2` and a call to the ‘`geom_smooth`’ option). The ‘local regressions’ are computed via weighted least squares, giving weight to data points in the ‘neighbourhood’ of values of the independent variable (here, age). The results of these computations are then used to construct ‘loess curves’ (i.e., locally estimated scatterplot smoothing) (Cleveland et al., 1992). The key relevant feature of this approach here is that it is non-parametric, i.e., agnostic about functional form and ‘shape’ of the relationship across the full range of the independent variable (hence the ‘local’ focus). Research on age and happiness is dominated by the idea of ‘u-shapes’. The shape of the age–happiness relationship is exactly what is in question here—so, we need an analysis that does not *presume* what the shape is. The more conventional practice of constructing models with a quadratic age specification makes exactly that assumption. (Simonsohn, 2018 shows in more general terms that a quadratic specification itself can mislead us about whether a relationship is ‘u-shaped’, via a strong possibility of false positives—i.e., instances where a quadratic specification appears to ‘fit’ the data but the relationship is not in fact u-shaped. Many researchers believe that these two ideas, quadratic and u-shape, are essentially equivalent, but Simonsohn shows that they *aren’t* equivalent.)

In contrast, a ‘smoothed’ visualisation using a flexible/local polynomial fitting allows the shape to emerge from the data themselves. If the relationship is indeed u-shaped, then that is the shape we will see in the visual results. But we might see other shapes as well, at least for some countries.



**Table 1** Data description

Country	n	ESS rounds	Happiness		Age	
			Mean	St. dev	Mean	St. dev
AT	12,655	1, 2, 3, 7, 8, 9	7.62	1.89	46.39	17.15
BE	16,377	1–9	7.73	1.52	47.49	17.56
BG	12,658	3, 4, 5, 6, 9	5.66	2.56	51.37	16.98
CH	16,071	1–9	8.09	1.49	47.74	17.18
CY	4917	3, 4, 5, 6, 9	7.44	1.87	47.31	17.52
CZ	19,107	1, 2, 4, 5, 6, 7, 8, 9	6.90	1.94	46.21	16.80
DE	24,222	1–9	7.44	1.88	48.84	17.05
DK	11,685	1, 2, 3, 4, 5, 6, 7, 9	8.33	1.42	49.01	17.03
EE	15,945	2–9	6.91	1.97	49.62	18.08
ES	16,272	1–9	7.57	1.74	47.28	17.48
FI	18,440	1–9	8.06	1.41	49.93	17.66
FR	18,128	1–9	7.30	1.80	48.46	17.34
GB	19,896	1–9	7.55	1.92	48.69	17.61
GR	12,058	1, 2, 4, 5	6.52	2.02	47.45	17.50
HR	6212	4, 5, 9	7.24	2.19	47.79	17.88
HU	15,909	1–9	6.43	2.28	48.94	17.59
IE	21,227	1–9	7.51	1.88	46.67	17.34
IS	3738	2, 6, 8, 9	8.21	1.44	48.44	17.31
IT	9568	1, 6, 8, 9	7.00	1.84	49.70	17.71
LT	10,919	5, 6, 7, 8, 9	6.59	2.15	49.04	17.46
LV	2762	4, 9	6.66	2.09	50.41	17.69
NL	17,583	1–9	7.86	1.37	48.42	16.87
NO	15,174	1–9	7.97	1.54	46.99	16.78
PL	14,570	1–9	7.05	2.11	46.08	17.73
PT	17,058	1–9	6.79	2.00	50.22	18.14
RU	11,812	3, 4, 5, 6, 8	6.18	2.21	44.83	17.79
SE	15,016	1–9	7.86	1.56	48.94	17.81
SI	12,683	1–9	7.26	1.95	48.28	17.59
SK	10,699	2, 3, 4, 5, 9	6.63	1.98	47.26	17.03
UA	9368	2, 3, 4, 5, 6	5.76	2.41	47.51	17.99

In a supplementary analysis presented below, I use regression models where age is entered as a categorical variable, denoting age ranges in 5-year intervals (18–22, 23–27, etc.). This approach likewise does not presume what the age–happiness relationship is. These models serve two purposes. First, we can create visualisations rooted in these models, to verify that they tell the same story as the smoothed results presented first. The second purpose is to consider the role of ‘period’ in our evaluation of the age–happiness relationships. In principle, if we do not include period as a control, the results for age might reflect ‘current’ changes in happiness. That possibility needs to be evaluated, but it is not inevitable that period will act as a confounder in the age–happiness relationship to any *substantial* extent. To evaluate it empirically, I construct a second set of models that include period (using the ESS round as the time variable) and then compare

the two sets of results. That comparison is then used to consider the robustness of the smoothed results (where use of controls is not possible).

The analysis thus does not focus on alleged technical ‘fixes’ to the APC dilemma (e.g. the ‘hierarchical age–period–cohort’ model, Yang, 2008). Evaluations of proposals of that sort make a compelling case that the idea of a technical ‘fix’ is a chimera (Fosse & Winship, 2019; Luo & Hodges, 2020). Instead, it is preferable to explore patterns in ways that consider the question substantively, asking e.g. whether we would have good *reasons* to expect that cohort effects would be confounded with age effects for the particular topic we are investigating (Ekstam, 2021; Voas & Chaves, 2016). A recent meta-analysis conducted by Buecker et al. (2023) offers reasons to be highly confident that the answer in regard to cohort is no: there are indeed cohort differences in subjective well-being, but these differences do not confound the effect of age in the sense of altering the patterns that appear.

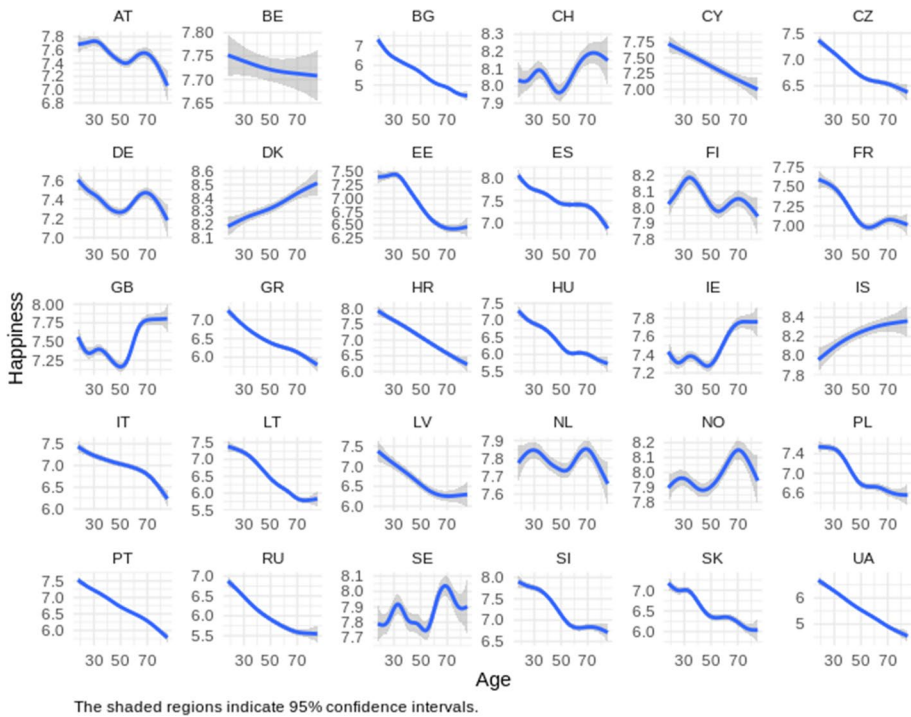
Finally, I select a smaller set of countries to explore results rooted in the conventional quadratic specification for age, with comparisons to a linear specification. This component of the analysis is focused on exploring limitations associated with using statistical significance as a way of evaluating results. Here we consider whether the results from a quadratic specification can be statistically significant even in situations where we know (from the earlier analysis) that the age–happiness relationship is not in fact u-shaped. This analysis is essential for resolving the puzzle identified above—i.e., the fact that researchers continue to lack clarity about the age–happiness relationship despite the enormous amount of attention devoted to it.

## 4 Results

The main results are presented in Fig. 1. The answer to our main question is immediately obvious: the age–happiness relationship is by no means universally u-shaped—instead, that relationship consists of a wide range of ‘shapes’. The dominant shape, evident in 17 of the 30 countries investigated here, amounts to a decline in happiness across the life course. That shape is evident in BG, CY, CZ, EE, ES, GR, HR, HU, IT, LT, LV, PL, PT, RU, SI, SK, and UA. There is some variation that could lead us to describe four of these countries in slightly different terms: for EE, LT, LV, and SI, the long-term decline in happiness does not persist into old age—instead there is a levelling off. But that levelling off does not amount to a post-middle-age *increase*; these are not u-shapes. Even for these countries (as for the broader set of 17) it is not sensible to speak in terms of a ‘midlife low’—because the decline in happiness that comes with middle age is not followed by a post-middle-age recovery.

A post-middle-age rise in happiness is evident in some countries: AT, CH, DE, FI, FR (maybe), GB, IE, NL, NO, and SE. But in almost all of these, a post-middle-age increase is followed by a decline in older age. (The exceptions are GB and IE.) In this group, the idea of a midlife low is supported. Still, the idea of a u-shape is misapplied here, insofar as what it implies about the post-middle-age life-course stage is only that happiness generally increases as people get older. It does increase—but then it decreases again as people pass through older age.

A third pattern is evident for DK and IS: in those countries happiness generally rises across the life-course. Here as well the ideas of u-shape and mid-life low are not relevant. Finally, in BE the age–happiness relationship is essentially flat. Any tendency to form an



**Fig. 1** Age and happiness

initial impression that there is a life-long decline must be tempered by closer inspection of the scale of the Y axis: the decline (from 7.75 to 7.71) is very slight.

Figure 1 is useful for presenting all 30 countries at the same time; we can easily see the variety of shapes that characterise this set of European countries. However, now that the variation in the Y-axis scales is apparent, it becomes necessary to ensure that we are not being misled in our sense of these patterns: if the age–happiness relationship in Belgium is flat (despite initial appearances), then perhaps something similar is true for some of the other countries? In the “Appendix”, I therefore present the same results, arranging the countries into three groups that facilitate use of fixed scales for the Y-axis within those groups. (The countries are grouped according to the vertical location of the lines, corresponding to average overall happiness in the countries in each group.) The bottom line is that, with use of more consistent scales, the same substantive conclusion is supported: in these European countries the age–happiness relationship takes a variety of shapes, and in the majority of instances there is no support for the idea of ‘u-shapes’. On the contrary: if anything, we see stronger support (especially in the third group) for the finding that age declines across the life course—in many instances the decline is very substantial. The decline is largest in Bulgaria, at 2.8 points; it is greater than 1.5 points in Croatia, Hungary, Lithuania, Portugal, and Ukraine.

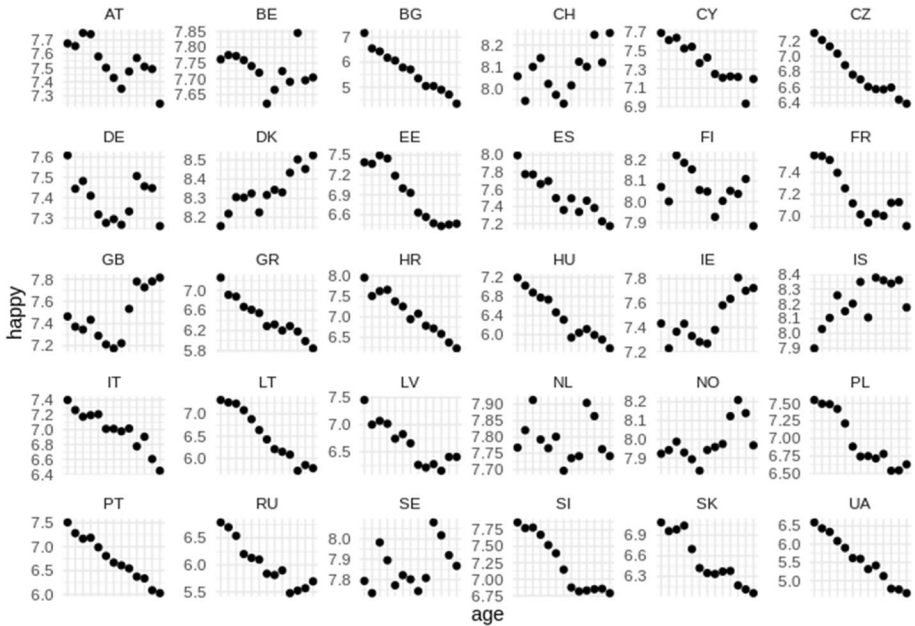


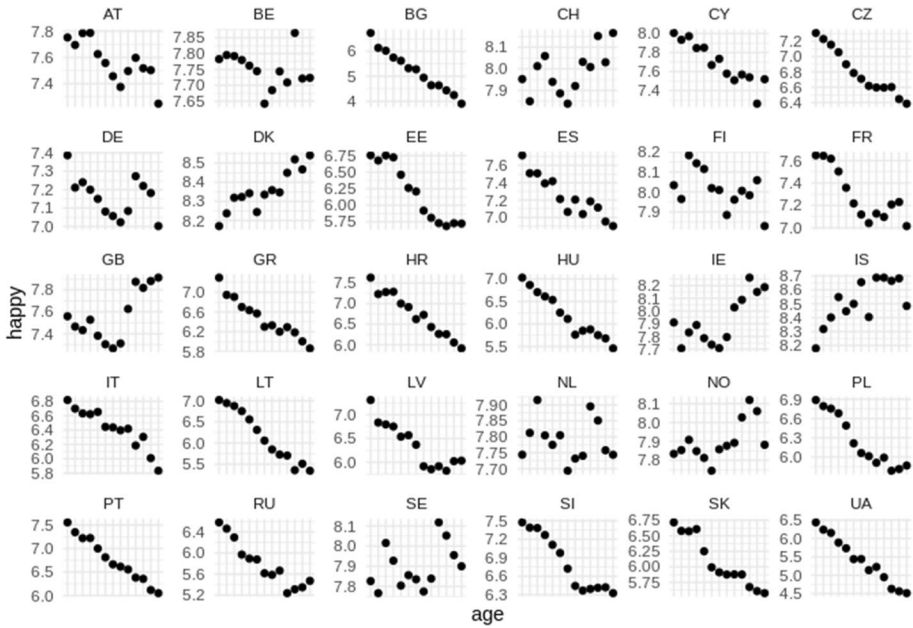
Fig. 2 Models of age & happiness, without year

### 4.1 Do We See Similar Results in a More Conventional Analysis?

For some readers, a set of results that consists entirely of graphs without underlying tables might provoke uncertainty. The choice of a visual presentation is of course deliberate: once again, we want to know what the shapes are. Also deliberate is the choice to refrain (at this stage) from presenting information about statistical significance. Still, a comparison to results from a more conventional analysis is advantageous. The results in Fig. 2 are likewise visualizations. They are rooted in more conventional regression models that use age as a ‘factor’ variable, with each category representing a 5-year age span. The model results are plotted: the intercept tells us the average level of happiness for people aged 18 to 22, and the ‘slope’ coefficients are then added to the intercept to give us the average happiness of people in the older age spans. This approach is (like the smoothed results above) agnostic about functional form. The tables themselves are available in the appendix.

The obvious conclusion here is that these more conventional results tell exactly the same story as in Fig. 1. The shapes for each country are essentially identical to the corresponding shapes in Fig. 1. Because the graphs here are not smoothed, there are some minor departures. In Switzerland and Croatia, for example, the average level of happiness for people in the second-youngest group (23–27) is lower in the non-smoothed results than it appears in the smoothed results. But these departures do not lead us to discern different shapes, relative to those in Fig. 1. In particular, there are no countries in Fig. 2 where we would discern u-shapes in a way that is not already apparent in Fig. 1.

Figure 2 is rooted in models that do not include ‘period’ as a control variable. In principle (in line with the APC dilemma), the ostensible relationship between age and happiness might partially reflect ‘current’ changes in happiness; the patterns might be different if we can disentangle the effects of period from the effects of age. That possibility can be



**Fig. 3** Models of age & happiness, controlling for year

evaluated by comparison to Fig. 3, which uses models where period is included as a control (again using a categorical/factor variable).

Here as well we see only very minor differences. In Latvia, for example, average happiness in the second-youngest group is slightly higher in the models where time is included as a control variable. Once again, however, we do not reach different conclusions for any of the countries investigated here. There are no instances where inclusion of the ESS ‘round’ variable leads to a u-shaped pattern that wasn’t already evident in Fig. 1 (and Fig. 2). Happiness levels can indeed change over time, in response to events—but for these countries during this time period this process of change is not ‘driving’ the overall age patterns. Those patterns are what they are (i.e., as apparent in analyses that rely on comparisons of people at different ages) and are not confounded by period.

### 4.2 Evaluating Results from a Quadratic Specification

For some, a presentation of visual results on their own might seem to constitute insufficient support for reaching substantive conclusions. In this section I present results from models that adopt a particular functional form (including the quadratic form). Quantitative researchers commonly set their results against a generic standard, asking whether those results meet a threshold of some sort. This is where statistical significance makes an appearance—and that appearance will allow us to see how statistical significance could mislead us if we (mis)use it to draw substantive conclusions about the ‘shape’ of the relationship.

Table 2 presents results from two parametric models for eight countries (Bulgaria, Czechia, Greece, Hungary, Latvia, Poland, Russia, and Slovakia): one model (on the left)

**Table 2** Models of happiness and age (quadratic vs. linear)

	BG	BG	CZ	CZ	GR	GR	HU	HU
Age	-0.06 (0.01)***	-0.04 (0.001)***	-0.03 (0.01)***	-0.02 (0.001)***	-0.04 (0.01)***	-0.02 (0.001)***	-0.05 (0.01)***	-0.02 (0.001)***
Age <sup>2</sup> /100	0.03 (0.01)***		0.01 (0.005)**		0.02 (0.01)***		0.03 (0.01)***	
Year?	yes	yes	yes	yes	yes	yes	yes	yes
Constant	7.78 (0.19)***	7.19 (0.09)***	7.99 (0.11)***	7.74 (0.07)***	7.92 (0.14)***	7.45 (0.06)***	8.00 (0.14)***	7.43 (0.07)***
N	12,658		19,107		12,058		15,909	
	LV	LV	PL	PL	RU	RU	SK	SK
Age	-0.05 (0.01)***	-0.02 (0.002)***	-0.04 (0.01)***	-0.02 (0.001)***	-0.04 (0.01)***	-0.02 (0.00)***	-0.04 (0.01)***	-0.02 (0.001)***
Age <sup>2</sup> /100	0.03 (0.01)*		0.02 (0.01)***		0.02 (0.01)***		0.02 (0.01)***	
Year?	yes	yes	yes	yes	yes	yes	yes	yes
Constant	8.16 (0.29)***	7.49 (0.11)***	7.72 (0.13)***	7.32 (0.06)***	7.34 (0.15)***	6.90 (0.07)***	7.50 (0.15)***	7.06 (0.07)***
N	2762		14,570		11,812		10,699	

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$ . The numbers in parentheses are standard errors.

using a quadratic specification for age (age plus age-squared/100), the other using a linear specification (on the right). All models include a categorical variable for time (using the ESS rounds). For all eight of these countries, we could find a basis for concluding that the age–happiness relationship is u-shaped, using conventional thresholds for statistical significance. But for all of these countries we also find statistically significant age coefficients using the *linear* specification. In each case the linear coefficients are negative, suggesting that happiness declines across the life course in all eight countries.

So: is the age–happiness relationship u-shaped in these countries, or is it ‘negative’ (in the sense of declining across the life course)? If we tried to decide between these two scenarios using only statistical significance as the criterion for a decision, we would struggle to make a choice, because in both specifications there are statistically significant results. (Nor does R-squared, as a measure of model ‘fit’, help us—because in each instance R-squared is identical across the pairs of models.)

But the visual results in Figs. 1, 2, and 3 have already helped us discern what the answer is. It is evident that happiness generally declines across the life course in each of these countries. If we were to start our investigation already holding the belief that the age–happiness relationship is u-shaped, we would very likely construct models with a quadratic specification—and then the fact that the age + age-squared coefficients are statistically significant in these countries would re-confirm our belief (a pattern Kratz & Brüderl, 2021 refer to as ‘quadratic specification bias’; compare Morgan & O’Connor, 2017 and Biermann et al., 2022). But this conclusion would be incorrect. We can note that this situation applies even to Bulgaria, already identified above as the country with the largest decline in happiness across the life course (2.8 points on the 11-point scale). Even here, a quadratic specification can lead to statistically significant coefficients, potentially leading to the misimpression that the age–happiness relationship is u-shaped there.

We can now appreciate why a non-parametric (e.g. visual) analysis is an essential first step. When we start with the analysis presented visually in Fig. 1, the results there lead us *away* from construction of parametric models with a quadratic function for age in most instances. The more usual practice of adopting a quadratic function by default (and then evaluating it solely on the basis of statistical significance) has strong potential to lead us astray.

Insofar as some researchers would want to evaluate results in part by considering whether they are statistically significant, it is perhaps reassuring that the negative linear coefficients in Table 2 are indeed statistically significant. (The same is true of all 17 countries identified, via Fig. 1, as showing a life-long decline in happiness.) The key argument here is that this information becomes useful only *after* we have determined what sort of functional specification to use in a regression model. Statistical significance on its own is not evidence telling us that the linear specification is the right one. Instead, it only helps us have some confidence that this particular finding, evident in the analysis of sample data, is likely to be found in the population as well.

### 4.3 Reflections on Use of Cross-Sectional Data

In general, quantitative social scientists believe that an analysis using panel data is superior to an analysis using cross-sectional data. That view is generally right. But we can be more precise about *why* a longitudinal approach is better, considering what the advantages are for specific research questions. As argued above, we do not need a longitudinal analysis for the purpose of guarding against ‘omitted-variable bias’ (because there can be no ‘confounders’ of the age–happiness relationship apart from cohort and period). The main reason to prefer a longitudinal analysis is that we can consider ‘mortality bias’. Substantively, this idea amounts to the possibility that cross-sectional results might overstate the extent of any increase in happiness as people become very old. Unhappier people are likely to die at younger ages. If we do not have data that reflect that process, we will get results that derive only from data on people who are still alive—i.e., people who are relatively happy as they age. Kratz and Brüderl (2021) demonstrate this upwards bias by comparing cross-sectional to longitudinal analysis of German panel data.

The same comparison cannot be achieved empirically in situations where we have only cross-sectional data. We can, however, engage in informed speculation about what the consequence of having only a cross-sectional analysis is likely to be. If we assume that the same pattern of bias is likely to hold generally (i.e., not just in Germany), we can conclude that the results presented in this paper are likely to show upward bias in the happiness trends among older people. The right-hand portion of the curves in Fig. 1 are therefore trending higher than would be evident in people’s experience if we somehow had direct access to that experience.

That observation reinforces the core finding of this paper: there is a diverse set of patterns characterising the age–happiness relationship. Consider the question from the reverse angle: does the likelihood of ‘mortality bias’ suggest that an analysis correcting for it would find that there are universal u-shapes after all? On the contrary: if we were able to correct for mortality bias in our cross-sectional results, we would likely find that any post-middle-age increase in happiness is smaller than what appears in the figures presented above. That correction would if anything further undermine the idea that the ‘u-shape’ is the predominant pattern in the age–happiness relationship.



Making effective decisions about whether to use cross-sectional data is important because overly stringent insistence on using panel data will severely restrict the countries we can investigate. There is a clear pattern evident in the set of countries for which panel data are available: longitudinal analysis is possible in particular for the UK, Germany, and Australia. As noted in the review of previous research above, there is support in longitudinal results for the finding of ‘u-shapes’ in the UK (Cheng et al., 2017; Movshuk, 2011) and Germany (Cheng et al., 2017). Other studies of those countries, while not identifying a u-shape, nonetheless find support for the idea that happiness rises after middle age (Biermann et al., 2022, Frijters and Beatton 2012). There is no reason to doubt those findings; among other things, the cross-sectional results presented above are consistent with them. But they cannot be used (simply on the basis of being rooted in longitudinal work) to suggest that the age–happiness relationship is *generally* u-shaped. Investigating a broader range of countries, using cross-sectional data, demonstrates that other shapes are common.

## 5 Conclusion

The relationship between age and happiness is indeed u-shaped in some countries. But the analysis here shows that it is by no means a universal pattern. Instead, a variety of patterns is evident, and a life-long decline in happiness is the more common pattern, certainly within Europe. That diversity is unsurprising if we contemplate the range of circumstances in which people live—as well as the diversity of human experience more generally (compare Galambos et al., 2020). The tendency to see instead a universal pattern is difficult to sustain as long as we explore the data in a way that goes beyond specifying a single functional form and then evaluating it via statistical significance alone. What is especially doubtful is the idea that happiness generally rises after a ‘midlife low’ (though again in some countries it does). In many cases, happiness instead continues to decline after midlife (having declined at younger ages as well). This pattern will be distorted by misuse of control variables, especially when the included controls pertain to negative aspects of people’s experience that often come with ageing in later life.

We are now in a position to evaluate the claim (Blanchflower et al., 2023) that there are a large number of studies—more than 600, a ‘vast literature’—demonstrating that the relationship between age and happiness is u-shaped. What is undeniably true is that there are hundreds of published studies containing regression models that include age and age-squared variables presented with asterisks indicating statistical significance. On the basis of the discussion above, however, we can ask two key questions: (1) do those coefficients do a good job of representing the underlying social process? Or, do they misrepresent that process, while nonetheless yielding statistically significant results simply because the sample is large enough? And (2), what other variables are present in the models? What is being controlled, and do those controls make sense in the context of considering the impact age might have on happiness?

Most research on happiness (and of course most of the 600+ studies listed by Blanchflower et al., 2023) is not constructed to evaluate the impact of age on happiness. In most of these, age is being used as a control variable in a model intended to investigate a different topic. Researchers typically use a quadratic specification for age—and when the age coefficients are then statistically significant (as indeed they typically are, in part because the sample size is large), a belief in the u-shape idea is reinforced. The fact that the age coefficients are often statistically significant comes in part from the fact that there are other

variables in the model; the resulting overcontrol bias inflates the age coefficients and the corresponding t-values, increasing the likelihood that the threshold of  $p < 0.05$  will be reached.

An effective understanding of how to select control variables can then help us see the unfortunate consequence of interpreting regression results for control variables. A control has been selected effectively if it is an antecedent of X (so,  $W \rightarrow X$ ). So, any impact of the control ( $W \rightarrow Y$ ) will be partly mediated by X ( $W \rightarrow X \rightarrow Y$ ). In the model, then, the ‘effect’ of W as represented by the coefficient partly reflects the fact that X is being held constant. In other words, the coefficient for W suffers from overcontrol bias (Rohrer 2018). In general, it does not make sense to interpret the control variables in a model that was constructed to evaluate the impact of the focal independent variable X (Keele et al., 2020).<sup>7</sup> Age is almost universally used as a control in research on the way other variables might influence happiness—but that research cannot give us useful information about the impact of age itself.

The relationship between age and happiness has evolved into a topic where many researchers’ activities amount to a clear case of ‘confirmation bias’ (Nickerson, 1998). Many people *believe* that this relationship is u-shaped, and they then construct analyses (using age as a control variable) that reflect that belief (i.e., adopting a quadratic specification for age). The subsequent evaluation of the corresponding results is often limited to the question of whether the coefficients are statistically significant—and the fact that sample sizes are generally large means that the answer to that question is typically yes. Thus is the belief confirmed. It should now be clear that this set of practices is not an effective way of evaluating whether the age–happiness relationship is in fact u-shaped.

## 5.1 Limitations

A limitation of the work presented here is again that it consists of cross-sectional analyses. A longitudinal analysis of panel data would inspire more confidence in results intended to tell us how happiness changes over the life course. It is always potentially risky to infer patterns of change using differences between people. However: it is important not to overstate the risks. In situations where it is possible to compare cross-sectional and longitudinal results for the same country, what we learn is that the cross-sectional analysis is generally effective in the sense that it gives results that are in line with what we see from a longitudinal analysis. The main departure has to do with the mortality bias evident in use of a cross-sectional approach—and once we know about that form of bias we can use it to refine our understanding of the results. This is a better perspective, relative to a view holding that cross-sectional results cannot be trusted in general. This perspective is also important in view of the limited availability of panel data. We would not want to form views about the age–happiness relationship in general solely on the basis of research using data from the few countries for which panel data are available.

<sup>7</sup> Westreich and Greenland (2013) evaluate the common practice of interpreting all the coefficients in a model and conclude that this practice is usefully described as the ‘Table 2 fallacy’.

## 5.2 Implications and Suggestions for Future Research

The main implication of the work presented here is that any future research exploring the relationship between age and happiness (as well as life satisfaction and other concepts of subjective well-being) should refrain from *assuming* that the form of the relationship is already known in any particular instance. The essential first step is a visual/non-parametric investigation. Use of a particular functional form (linear or quadratic, or perhaps cubic, etc.), if any, would *follow* from what the non-parametric analysis tells us. If we have reasons to expect to find a specific shape characterising a relationship, that shape will be evident in a non-parametric analysis; we do not risk missing it by starting with the non-parametric approach. A greater risk is taken when we start instead by assuming that a particular functional form is suitable—especially if we then evaluate that idea via statistical significance alone. These methodological arguments offered here have potential relevance to other research topics as well, especially when we do not already have good reasons to believe that the relationship in question takes a particular ‘shape’ (including a linear one).

The second suggestion has to do with use of age as a control variable, for exploring the impact of other variables on happiness (or indeed any other dependent variable). The usual practice is to use a quadratic specification. It is better to enter age as a categorical variable, e.g. using age ranges as in the non-parametric analysis for Figs. 2 and 3 above (compare Kratz & Brüderl, 2021). That practice has the merit of remaining agnostic about whether the age–happiness relationship is u-shaped. Even for control variables, the specification of any functional form should be consistent with the underlying social patterns; a misspecification in this sense can lead to biased results for the focal independent variable. In a non-parametric analysis that risk is avoided.

## Appendix

See Table 3 and 4.

**Table 3** Models of happiness

	AT	BE	BG	CH	CY	CZ	DE	DK	EE	ES	FI	FR	GB	GR	HR
Intercept	7.68	7.76	7.17	8.06	7.69	7.30	7.61	8.15	7.39	8.00	8.07	7.55	7.46	7.25	7.96
23–27	-0.02	0.01	-0.62	-0.11	-0.08	-0.09	-0.16	0.06	-0.02	-0.22	-0.07	-0.01	-0.09	-0.34	-0.45
28–32	0.08	0.01	-0.74	0.04	-0.05	-0.17	-0.13	0.15	0.11	-0.22	0.15	-0.04	-0.12	-0.38	-0.33
33–37	0.07	0.00	-0.99	0.08	-0.17	-0.27	-0.20	0.15	0.06	-0.33	0.12	-0.16	-0.03	-0.58	-0.30
38–42	-0.09	-0.02	-1.10	-0.03	-0.15	-0.42	-0.29	0.17	-0.20	-0.30	0.08	-0.30	-0.17	-0.64	-0.58
43–47	-0.17	-0.04	-1.38	-0.09	-0.32	-0.54	-0.33	0.07	-0.39	-0.50	-0.02	-0.44	-0.25	-0.70	-0.70
48–52	-0.25	-0.14	-1.46	-0.13	-0.26	-0.60	-0.31	0.16	-0.46	-0.64	-0.02	-0.54	-0.29	-0.96	-1.01
53–57	-0.33	-0.10	-1.82	-0.04	-0.44	-0.70	-0.34	0.19	-0.76	-0.50	-0.14	-0.61	-0.24	-0.93	-0.88
58–62	-0.20	-0.04	-2.12	0.07	-0.48	-0.73	-0.28	0.18	-0.82	-0.66	-0.07	-0.53	0.07	-1.05	-1.17
63–67	-0.10	-0.07	-2.12	0.04	-0.46	-0.73	-0.10	0.28	-0.92	-0.53	-0.02	-0.55	0.31	-0.96	-1.24
68–72	-0.17	0.08	-2.27	0.19	-0.47	-0.71	-0.15	0.35	-0.96	-0.61	-0.03	-0.43	0.26	-1.07	-1.37
73–77	-0.18	-0.07	-2.46	0.06	-0.75	-0.86	-0.16	0.30	-0.94	-0.77	0.04	-0.43	0.32	-1.26	-1.58
78–82	-0.43	-0.06	-2.83	0.20	-0.49	-0.92	-0.35	0.37	-0.92	-0.82	-0.19	-0.64	0.35	-1.41	-1.73
Time?	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
	HU	IE	IS	IT	LT	LV	NL	NO	PL	PT	RU	SE	SI	SK	UA
Intercept	7.19	7.43	7.90	7.40	7.31	7.45	7.77	7.92	7.55	7.51	6.79	7.79	7.87	7.09	6.60
23–27	-0.17	-0.20	0.13	-0.14	-0.06	-0.45	0.05	0.02	-0.05	-0.23	-0.09	-0.06	-0.09	-0.13	-0.17
28–32	-0.31	-0.07	0.21	-0.22	-0.08	-0.38	0.15	0.06	-0.06	-0.34	-0.25	0.19	-0.08	-0.10	-0.26
33–37	-0.42	0.00	0.36	-0.20	-0.23	-0.44	0.02	0.01	-0.13	-0.32	-0.59	0.10	-0.19	-0.05	-0.51
38–42	-0.46	-0.10	0.25	-0.19	-0.43	-0.71	0.00	-0.03	-0.33	-0.52	-0.66	-0.02	-0.35	-0.39	-0.70
43–47	-0.73	-0.15	0.30	-0.39	-0.68	-0.63	0.03	-0.09	-0.66	-0.70	-0.69	0.03	-0.47	-0.67	-0.98
48–52	-0.88	-0.17	0.45	-0.39	-0.89	-0.80	-0.07	0.02	-0.80	-0.84	-0.95	0.01	-0.72	-0.74	-1.00
53–57	-1.25	-0.05	0.21	-0.42	-1.11	-1.19	-0.03	0.04	-0.80	-0.90	-0.98	-0.05	-0.99	-0.75	-1.28
58–62	-1.15	0.15	0.48	-0.38	-1.17	-1.24	-0.03	0.05	-0.83	-0.96	-0.89	0.01	-1.05	-0.72	-1.18
63–67	-1.08	0.20	0.46	-0.62	-1.23	-1.18	0.14	0.20	-0.76	-1.14	-1.31	0.29	-1.04	-0.71	-1.47
68–72	-1.20	0.38	0.44	-0.49	-1.60	-1.30	0.10	0.29	-1.00	-1.17	-1.27	0.23	-1.02	-0.92	-1.82

Table 3 (continued)

	HU	IE	IS	IT	LT	LV	NL	NO	PL	PT	RU	SE	SI	SK	UA
73-77	-1.29	0.27	0.46	-0.79	-1.47	-1.05	0.00	0.22	-1.00	-1.42	-1.22	0.13	-1.01	-0.98	-1.84
78-82	-1.48	0.29	0.28	-0.95	-1.54	-1.05	-0.03	0.04	-0.91	-1.48	-1.10	0.07	-1.08	-1.03	-1.94
Time?	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No

**Table 4** Models of happiness

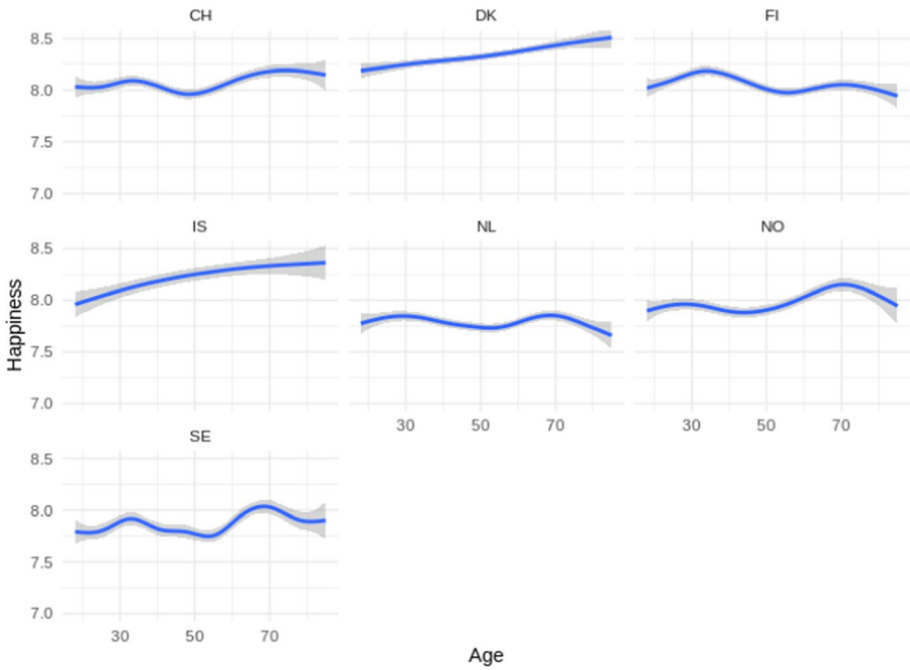
	AT	BE	BG	CH	CY	CZ	DE	DK	EE	ES	FI	FR	GB	GR	HR
Intercept	7.75	7.78	6.74	7.95	8.01	7.31	7.39	8.17	6.75	7.71	8.03	7.65	7.56	7.29	7.63
23–27	-0.06	0.01	-0.60	-0.10	-0.07	-0.08	-0.18	0.07	-0.08	-0.21	-0.07	0.00	-0.09	-0.35	-0.40
28–32	0.03	0.01	-0.71	0.06	-0.04	-0.15	-0.15	0.15	0.00	-0.21	0.15	-0.03	-0.13	-0.39	-0.35
33–37	0.03	0.00	-0.99	0.11	-0.16	-0.25	-0.19	0.15	-0.03	-0.32	0.11	-0.14	-0.03	-0.59	-0.34
38–42	-0.13	-0.02	-1.12	-0.01	-0.16	-0.41	-0.24	0.17	-0.29	-0.30	0.08	-0.29	-0.18	-0.66	-0.63
43–47	-0.19	-0.04	-1.41	-0.07	-0.34	-0.52	-0.31	0.07	-0.49	-0.50	-0.01	-0.43	-0.25	-0.73	-0.72
48–52	-0.29	-0.14	-1.46	-0.11	-0.27	-0.60	-0.33	0.16	-0.55	-0.65	-0.02	-0.53	-0.29	-0.99	-1.01
53–57	-0.38	-0.10	-1.79	-0.03	-0.43	-0.69	-0.36	0.19	-0.84	-0.51	-0.15	-0.60	-0.25	-0.97	-0.90
58–62	-0.26	-0.04	-2.10	0.08	-0.50	-0.71	-0.30	0.18	-0.94	-0.68	-0.07	-0.52	0.07	-1.09	-1.20
63–67	-0.15	-0.07	-2.10	0.06	-0.44	-0.71	-0.11	0.28	-1.02	-0.53	-0.03	-0.55	0.31	-1.00	-1.37
68–72	-0.23	0.08	-2.30	0.20	-0.47	-0.71	-0.17	0.35	-1.07	-0.60	-0.05	-0.44	0.26	-1.11	-1.38
73–77	-0.25	-0.06	-2.49	0.08	-0.74	-0.86	-0.21	0.30	-1.03	-0.76	0.02	-0.42	0.32	-1.29	-1.58
78–82	-0.50	-0.06	-2.83	0.21	-0.49	-0.93	-0.38	0.37	-1.03	-0.81	-0.20	-0.63	0.35	-1.43	-1.72
Time?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	HU	IE	IS	IT	LT	LV	NL	NO	PL	PT	RU	SE	SI	SK	UA
Intercept	7.03	7.91	8.18	6.82	7.02	7.31	7.74	7.83	6.89	7.56	6.58	7.83	7.47	6.72	6.44
23–27	-0.16	-0.20	0.14	-0.12	-0.07	-0.47	0.07	0.02	-0.10	-0.21	-0.12	-0.06	-0.08	-0.14	-0.20
28–32	-0.32	-0.08	0.22	-0.19	-0.14	-0.52	0.17	0.07	-0.13	-0.34	-0.29	0.19	-0.09	-0.14	-0.29
33–37	-0.42	-0.02	0.37	-0.20	-0.26	-0.56	0.06	0.01	-0.21	-0.33	-0.61	0.10	-0.21	-0.11	-0.55
38–42	-0.49	-0.12	0.27	-0.17	-0.46	-0.77	0.03	-0.02	-0.40	-0.56	-0.69	-0.02	-0.36	-0.47	-0.71
43–47	-0.78	-0.17	0.32	-0.37	-0.71	-0.74	0.06	-0.09	-0.67	-0.74	-0.70	0.03	-0.50	-0.73	-1.00
48–52	-0.92	-0.20	0.48	-0.38	-0.96	-0.94	-0.05	0.02	-0.83	-0.89	-0.97	0.01	-0.75	-0.81	-1.00
53–57	-1.27	-0.11	0.23	-0.42	-1.17	-1.40	-0.01	0.04	-0.87	-0.94	-1.00	-0.05	-1.03	-0.84	-1.30
58–62	-1.18	0.12	0.51	-0.40	-1.29	-1.46	0.00	0.06	-0.98	-1.00	-0.92	0.01	-1.10	-0.84	-1.21
63–67	-1.15	0.18	0.51	-0.63	-1.31	-1.40	0.15	0.19	-0.90	-1.17	-1.34	0.29	-1.08	-0.84	-1.49
68–72	-1.28	0.35	0.49	-0.51	-1.66	-1.49	0.11	0.28	-1.10	-1.20	-1.27	0.23	-1.06	-1.04	-1.81

**Table 4** (continued)

	HU	IE	IS	IT	LT	LV	NL	NO	PL	PT	RU	SE	SI	SK	UA
73-77	-1.35	0.24	0.50	-0.81	-1.51	-1.29	0.01	0.23	-1.07	-1.44	-1.23	0.13	-1.05	-1.11	-1.88
78-82	-1.57	0.28	0.31	-0.98	-1.67	-1.28	0.00	0.05	-1.02	-1.50	-1.11	0.07	-1.15	-1.14	-1.93
Time?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

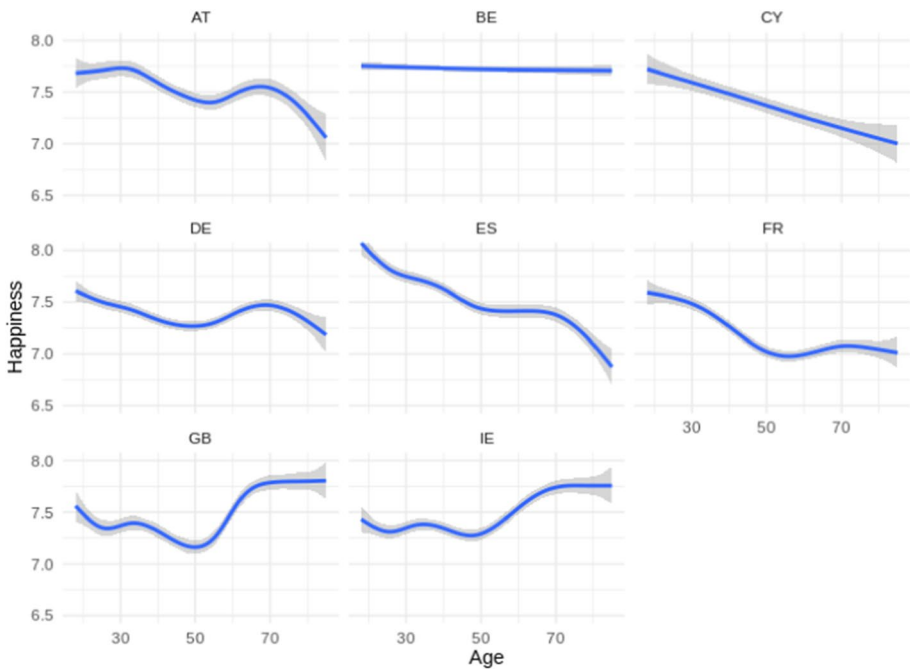


### Age and happiness (first group)

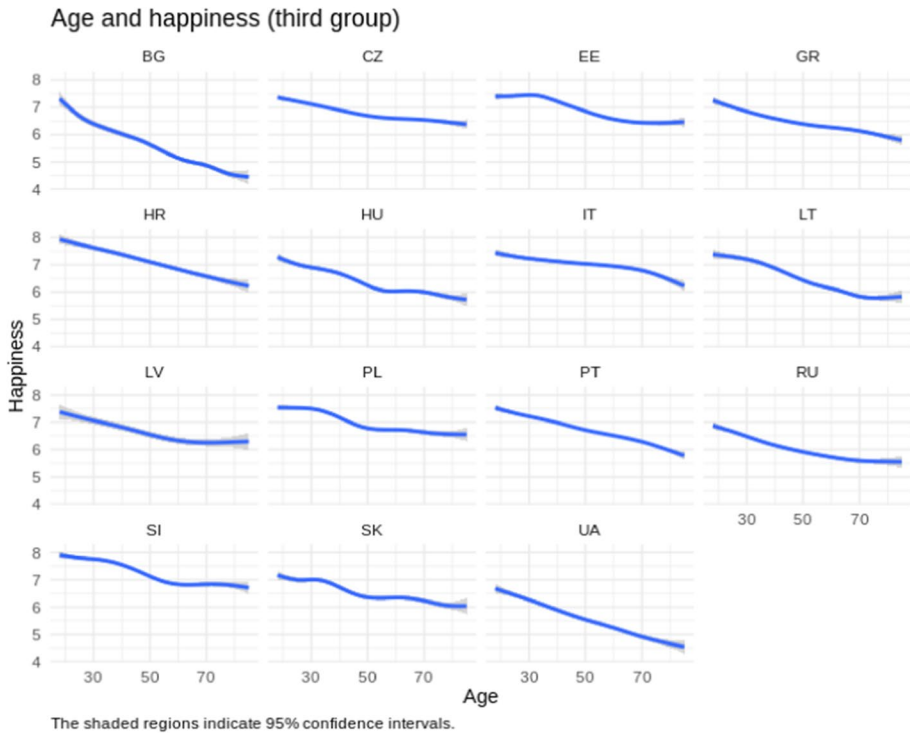


The shaded regions indicate 95% confidence intervals.

### Age and happiness (second group)



The shaded regions indicate 95% confidence intervals.



**Acknowledgements** I am grateful to Kelsey O'Connor and Matthew Tonkin for feedback on an earlier draft, and to the anonymous reviewers for their highly constructive comments on the initial submission.

**Funding** None.

**Data Availability** Data from the European Social Survey are available at [www.europeansocialsurvey.org](http://www.europeansocialsurvey.org).

**Code Availability** The R syntax/code for this article is available on request and at the following link: [https://osf.io/3gebn/?view\\_only=d088c2ef56cf4733a0db6a86be4f6a00](https://osf.io/3gebn/?view_only=d088c2ef56cf4733a0db6a86be4f6a00)

## Declarations

**Conflict of interest** This manuscript was submitted during the period while the author was a co-editor of the journal. It was therefore handled under procedures developed for these circumstances.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Agresti, A., & Finlay, B. (1997). *Statistical methods for the social sciences*. Prentice Hall.
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Bartram, D. (2021). Age and life satisfaction: Getting control variables under control. *Sociology*, 55(2), 421–437. <https://doi.org/10.1177/0038038520926871>
- Bartram, D. (2023). Is happiness u-shaped in age everywhere? A methodological reconsideration for Europe. *National Institute Economic Review*, 263, 61–75. <https://doi.org/10.1017/nic.2022.1>
- Becker, C. K., & Trautmann, S. T. (2022). Does happiness increase in old age? Longitudinal evidence from 20 European countries. *Journal of Happiness Studies*, 23, 3625–3654. <https://doi.org/10.1007/s10902-022-00569-4>
- Beja, E. L. (2018). The U-shaped relationship between happiness and age: Evidence using World Values Survey data. *Quality & Quantity*, 52(4), 1817–1829. <https://doi.org/10.1007/s11135-017-0570-z>
- Biermann, P., Bitzer, J., & Gören, E. (2022). The relationship between age and subjective well-being: Estimating within and between effects simultaneously. *The Journal of the Economics of Ageing*, 21, 100366. <https://doi.org/10.1016/j.jeoa.2021.100366>
- Bittmann, F. (2021). Beyond the U-shape: Mapping the functional form between age and life satisfaction for 81 countries utilizing a cluster procedure. *Journal of Happiness Studies*, 22, 2343–2359. <https://doi.org/10.1007/s10902-020-00316-7>
- Blanchflower, D. G. (2021). Is happiness U-shaped everywhere? Age and subjective well-being in 145 countries. *Journal of Population Economics*, 34(2), 575–624. <https://doi.org/10.1007/s00148-020-00797-z>
- Blanchflower, D. G., Graham, C., & Piper, A. (2023). Happiness and age: Resolving the debate. *National Institute Economic Review*, 263, 76–93.
- Blanchflower, D. G., & Oswald, A. J. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine*, 66, 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>
- Blanchflower, D. G., & Oswald, A. J. (2019). Do humans suffer a psychological low in midlife? Two approaches (With and Without Controls) in seven data sets. In M. Rojas (Ed.), *The economics of happiness* (pp. 439–453). Springer. [https://doi.org/10.1007/978-3-030-15835-4\\_19](https://doi.org/10.1007/978-3-030-15835-4_19)
- Buecker, S., Luhmann, M., Haehner, P., Bühler, J. L., Dapp, L. C., Luciano, E. C., & Orth, U. (2023). The development of subjective well-being across the life span: A meta-analytic review of longitudinal studies. *Psychological Bulletin*, 149(7–8), 418–446. <https://doi.org/10.1037/bul0000401>
- Carver, R. (1978). The case against statistical significance testing. *Harvard Educational Review*, 48(3), 378–399. <https://doi.org/10.17763/haer.48.3.t490261645281841>
- Cheng, T. C., Powdthavee, N., & Oswald, A. J. (2017). Longitudinal evidence for a Midlife Nadir in human well-being: Results from four data sets. *The Economic Journal*, 127(599), 126–142. <https://doi.org/10.1111/eoj.12256>
- Clark, A. E., Flèche, S., Layard, R., Powdthavee, N., & Ward, G. (2018). *The origins of happiness: The science of well-being over the life course*. Princeton University Press.
- Cleveland, W. S., Grosse, E., & Shyu, W. M. (1992). Local regression models. In T. J. Hastie (Ed.), *Statistical models in S*. Routledge.
- Ekstam, D. (2021). The liberalization of American attitudes to homosexuality and the impact of age, period, and cohort effects. *Social Forces*, 100(2), 905–929. <https://doi.org/10.1093/sf/soaa131>
- Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology*, 40(1), 31–53. <https://doi.org/10.1146/annurev-soc-071913-043455>
- Fosse, E., & Winship, C. (2019). Analyzing age–period–cohort data: A review and critique. *Annual Review of Sociology*, 45(1), 467–492. <https://doi.org/10.1146/annurev-soc-073018-022616>
- Frijters, P., & Beaton, 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>
- Galambos, N. L., Fang, S., Krahn, H. J., Johnson, M. D., & Lachman, M. E. (2015). Up, not down: The age curve in happiness from early adulthood to midlife in two longitudinal studies. *Developmental Psychology*, 51(11), 1664–1671. <https://doi.org/10.1037/dev0000052>
- 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), 898–912. <https://doi.org/10.1177/1745691620902428>
- Geerling, D. M., & Diener, E. (2020). Effect size strengths in subjective well-being research. *Applied Research in Quality of Life*, 15(1), 167–185. <https://doi.org/10.1007/s11482-018-9670-8>

- Glenn, N. (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. *Social Science & Medicine*, 69(4), 481–485. <https://doi.org/10.1016/j.socscimed.2009.05.038>
- Gorard, S. (2016). Damaging real lives through obstinacy: Re-emphasising why significance testing is wrong. *Sociological Research Online*, 21(1), 1–14. <https://doi.org/10.5153/sro.3857>
- Graham, C., & Ruiz Pozuelo, J. (2017). Happiness, stress, and age: How the U curve varies across people and places. *Journal of Population Economics*, 30(1), 225–264. <https://doi.org/10.1007/s00148-016-0611-2>
- Hellevik, O. (2017). The U-shaped age–happiness relationship: Real or methodological artifact? *Quality & Quantity*, 51, 177–197. <https://doi.org/10.1007/s1135-015-0300-3>
- Hudomiet, P., Hurd, M. D., & Rohwedder, S. (2021). The age profile of life satisfaction after age 65 in the U.S. *Journal of Economic Behavior & Organization*, 189, 431–442. <https://doi.org/10.1016/j.jebo.2021.07.002>
- Jivraj, S., Nazroo, J., Bram, B., & Chandola, T. (2014). Aging and subjective well-being in later life. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 69(6), 930–941. <https://doi.org/10.1093/geronb/gbu006>
- Jowell, R. (2007). *European Social Survey, Technical Report*. Centre for Comparative Social Surveys, City University.
- 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>
- Keele, L., Stevenson, R. T., & Elwert, F. (2020). The causal interpretation of estimated associations in regression models. *Political Science Research and Methods*, 8(1), 1–13. <https://doi.org/10.1017/psrm.2019.31>
- Kratz, F., & Brüderl, J. (2021). *The Age Trajectory of Happiness: How Lack of Causal Reasoning has Produced the Myth of a U-Shaped Age–Happiness Trajectory*. PsyArXiv preprints. <https://psyarxiv.com/d8f2z/>
- Laaksonen, S. (2018). A research note: Happiness by age is more complex than U-shaped. *Journal of Happiness Studies*, 19(2), 471–482. <https://doi.org/10.1007/s10902-016-9830-1>
- Luo, L., & Hodges, J. S. (2020). Constraints in random effects age–period–cohort models. *Sociological Methodology*, 50(1), 276–317. <https://doi.org/10.1177/0081175020903348>
- Martin, J. L. (2018). *Thinking through statistics*. The University of Chicago Press.
- Morgan, R., & O'Connor, K. J. (2017). Experienced life cycle satisfaction in Europe. *Review of Behavioral Economics*, 4(4), 371–396. <https://doi.org/10.1561/105.00000070>
- Movshuk, O. (2011). Why is life satisfaction U-shaped in age? *Journal of Behavioral Economics and Finance*, 4, 133–138. <https://doi.org/10.11167/jbef.4.133>
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, 3, 96–146. <https://doi.org/10.1214/09-SS057>
- Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, 1(1), 27–42. <https://doi.org/10.1177/2515245917745629>
- Simonsohn, U. (2018). Two lines: A valid alternative to the invalid testing of U-shaped relationships with quadratic regressions. *Advances in Methods and Practices in Psychological Science*, 1(4), 538–555. <https://doi.org/10.1177/2515245918805755>
- Stephoe, A. (2019). Happiness and health. *Annual Review of Public Health*, 40(1), 339–359. <https://doi.org/10.1146/annurev-publhealth-040218-044150>
- Voas, D., & Chaves, M. (2016). Is the United States a counterexample to the secularization thesis? *American Journal of Sociology*, 121(5), 1517–1556. <https://doi.org/10.1086/684202>
- Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a World Beyond ‘ $p < 0.05$ .’ *The American Statistician*, 73(1), 1–19. <https://doi.org/10.1080/00031305.2019.1583913>
- Westreich, D., & Greenland, S. (2013). The Table 2 Fallacy: Presenting and interpreting confounder and modifier coefficients. *American Journal of Epidemiology*, 177(4), 292–298. <https://doi.org/10.1093/aje/kws412>
- 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.