



Beyond the U-Shape: Mapping the Functional Form Between Age and Life Satisfaction for 81 Countries Utilizing a Cluster Procedure

Felix Bittmann¹

Accepted: 21 September 2020 / Published online: 22 October 2020
© The Author(s) 2020

Abstract

Classifying and explaining the causal and functional relationship between age and life satisfaction, especially in an international context, is still a major open question in demographics and happiness-research. Especially the debate whether to include sociodemographic control variables in these models has received much attention and deserves more discussion. The current contribution takes a cross-country perspective and attempts to sort countries into larger clusters, depending on their specific functional form. Using cross-sectional data from 81 countries with more than 170,000 respondents, the analyses demonstrate that there exist three larger clusters which display distinct functional relations (linear decline, U-shape, decline with a stable old-age period). Sociodemographic controls are not introduced since the total causal effect is to be estimated. Furthermore, the contribution explains cluster membership exploratively using macro indicators. While it becomes clear that countries with a linear decline are usually less developed countries, differences between the other two clusters are much less obvious.

Keywords Age · Subjective well-being · Life satisfaction · Causal relationship · Functional form · Confounding

JEL Classification C21 · D60 · I31

1 Introduction

There is a growing interest in the study of well-being, life satisfaction and happiness (Frey and Stutzer 2005). Especially the relation to age, which is not surprising given the global increase in life expectancy and standard of living, appears to be of practical relevance. Even more exciting is perhaps the quite heterogeneous picture that is found in the literature with respect to the relation between age and life satisfaction. While some studies report a clear U-shape relationship between age and well-being (Blanchflower and Oswald 2009;

✉ Felix Bittmann
felix.bittmann@uni-bamberg.de

¹ Leibniz Institute for Educational Trajectories, Wilhelmsplatz 3, 96047 Bamberg, Germany

Frijters and Beaton 2012), others come to very different conclusions (Easterlin 2006; Hellevik 2017; Mroczek and Kolarz 1998). Given the reasonable amount of high-quality data from many countries, these diverging results are puzzling and fuel the motivation to find the “correct” solution to the problem. However, as will be made clear, there is probably not *the* right answer but diverging findings can be explained by several methodological aspects and different research questions. Indeed, it seems highly relevant to spell out clearly the objectives of any research project. This not only increases the comparability of studies immensely but also allows the selection of the most appropriate research design for the given question. It is the aim of this contribution to shed some light on these critical aspects. In addition, more attention is given to cross-country comparisons and specific structural factors, that might influence the functional forms of age and life satisfaction.

Summarized, the present study is going to answer the following research questions:

1. What is the total effect of age on life satisfaction globally? What is the functional form of this relationship?
2. Are there significant differences in the functional forms between countries and can countries be classified into larger clusters which follow the similar patterns?
3. Which variables can explain why certain countries are classified into the clusters? How can clusters be characterised with respect to macro indicators?

2 Theoretical Considerations

2.1 Review of the Literature

There is a broad ongoing debate about the functional relation between age and life satisfaction. One comparative literature review from 2013 lists 20 different studies with partially quite different findings (López Ulloa et al. 2013). Regarding this summary and some more recent studies (Cheng et al. 2017; Graham and Ruiz Pozuelo 2017; Hellevik 2017; Laaksonen 2018), our impression is that especially the following five factors contribute to the diverging findings. Firstly, the selection of countries is relevant since it can be assumed that especially differences in culture, society, welfare-state and demographics are influential. Therefore, it is not surprising that different results arise when different countries or groups of countries are studied. Secondly, cross-sectional and longitudinal studies can differ since the first usually compare different countries at one point in time, while the second analyse cohort effects, usually within one or a few countries. Both approaches are valid but have quite different research interests. The first investigates how or why countries differ, while the second allows separating age effects from cohort effects and tests whether there are within-country changes over time (possibly due to shifts in macro factors, like overall economic prosperity or societal changes). Thirdly, the inclusion of sociodemographic control variables apparently has a big impact on the findings. While some studies do not include these, others add variables like gender, ethnicity, education or health. As discussed in more detail below, there are good arguments for the omission or inclusion of control variables. However, then it is not surprising that studies with different approaches reach quite different conclusions. Fourthly, results might be different due to the usage of different statistical methods and techniques. Hopefully, these differences are rather minor since most statistical tools should come to the same conclusions when underlying effects are strong and stable. Finally, another reason for differences is the usage of different constructs or

operationalizations of life satisfaction. For example, *well-being*, *happiness* and *life satisfaction*, to name only three, are all commonly used in different studies. While similar, they have different implications for the cognitive and emotional aspects of the construct and do not measure the exact same thing, which can lead to different findings (Hellevik 2017; Maddux 2018). Since the objective of the present contribution is of empirical nature, an in-depth review of the literature, regarding all factors mentioned above, is not possible.¹ Consequently, only studies and findings are discussed that relate to the research questions that were spelt out explicitly in the introduction.

There are a few studies available that focus on a similar research question and a cross-sectional/comparative analysis of many countries. Blanchflower and Oswald (2009) find that most European countries display the U-shape in analyses when omitting controls, while no such shape is found for the USA. However, since they use different constructs (*happiness* in the US and *life satisfaction* in Europe), the comparability of results is questionable. In contrast to them, Laaksonen (2018), who also omits sociodemographic controls and utilizes the *happiness*-construct, finds no U-shape for more than 25 European countries, however, the U-shape appears when controls are added. Another study, which employs even more countries and makes use of the *best possible life Cantril ladder* to measure well-being finds the U-shape for 44 of 46 countries, however, under control of several sociodemographic controls (Graham and Ruiz Pozuelo 2017). Another study with 132 countries finds that age-profiles of life satisfaction differs between countries (Deaton 2008). Summarized, while there are some relevant studies available, they often differ due to design or focus. To our knowledge, there is no study available that attempts to explain the functional form of countries systematically. While there are some good arguments why countries should differ in these forms, they are rarely stated explicitly and no well-developed theoretical frameworks are available. While the following analyses are classified as exploratively and no testable hypotheses are formulated, it should be outlined how these arguments could be structured. It is obvious that countries differ largely with respect to certain macro factors, like overall human development, social inequality, economic prosperity or demographics. Some of these factors might affect the functional forms significantly. For example, one could hypothesize that a well-developed health system and strong welfare state regimes are associated with higher life satisfaction for retired people, since they influence health and the financial means positively, which in turn relate to higher life satisfaction (Ryff et al. 2004). Consequently, the functional form between age and life satisfaction is moderated by the availability of welfare state regimes. Similar arguments could be put forward to explain why certain countries display specific functional forms, which appears to be a relevant focus of analysis.

2.2 The Function of Control Variables

As the review of the literature underlines, there is no consensus about the inclusion of control variables when the relation between age and life satisfaction is studied. In our view, the approximation of causal relationships should be the aim of most empirical studies (Hernán 2018) and consequently, a causal approach is chosen. Following Judea Pearl's framework of causality (2009a, b), to estimate the causal relationship between x (cause) and y (effect),

¹ This being said, future studies are encouraged to review the literature with respect to mentioned or other factors to create a systematic and structured overview, following the example of López Ulloa et al. (2013).

all variables have to be considered as control variables that *simultaneously* affect both cause and effect. These variables are called *confounders*. If not all relevant confounders are included in the analyses, spurious findings might arise (Shahar and Shahar 2013). The classical example is the relation between the number of storks in a region and birth rates, which is “explained away” by a confounding factor (degree of urbanization) (Matthews 2000). While it is obvious that there are many different factors influencing life satisfaction, there are no such variables that do influence age. Since (chronological) age is the time span between the date of birth and present, and time passes equally for all beings (on planet earth), there is no good justification to include any sociodemographic control variables when the causal relationship between age and life satisfaction is to be studied (Easterlin 2006; Glenn 2009).² Introducing these variables can rather lead to a biased estimate of the effect. Many studies which were presented above include, for example, marital status, health, or educational degrees as control variables. As these cannot account for spurious correlations as they do not fulfil the criterion for confounders, they are rather *mediators* which explain *how* age affects life satisfaction. When these are included, say, in an OLS regression, the regression coefficient of age is net of the indirect effect of age on life satisfaction, that is transmitted via the mediator (Gangl 2010). While studies of this kind can be elucidating for explaining mechanisms in detail (Lelkes 2008), when only the overall causal relationship or functional form is of interest, they can result in misleading findings due to *overcontrol bias* (Elwert and Winship 2014). For example, one study states that “(...) all of these studies controlled for potentially confounding factors such as income, health, employment, and gender. The intuition in this instance is to look at the ‘pure’ effects of aging, controlling for the many other confounding factors that also come with age, such as declining health and leaving the work force” (Graham and Ruiz Pozuelo 2017, p. 228). Apparently, the usage of the term “confounding factors” is incorrect here as health is a *mediator*. It is important to state that these findings are not incorrect per se, but they are probably not answering the originally posed research questions (Bartram 2020). Consequently, the inclusion of (control) variables has to be justified with the goal of the research project. As a side note, control variables and their interactions can be included when groups are to be compared. For example, when one wants to test whether the functional form is different for men and women, this approach would be justified.

As the main question of the present study is to investigate the (total) causal and functional relationship between age and life satisfaction, no sociodemographic control variables are included. However, there are two other controls to account for the survey design, which are introduced below. These were included since they actually could influence age and life satisfaction simultaneously.

² Different considerations might be necessary when the *perceived* age is to be studied. For example, when the age of the respondent is operationalized via the subjective rating of an interviewer, there might be good arguments for a different analytical framework to adjust for or explain the subjective perception of age. However, this point is not the focus of the current study.

3 Data, Sample and Methods

3.1 Data and Sample

The analyses have two data sources: the European Social Survey (ESS), round 8 (v2.1) and 9 (v2.0), and the World Values Survey (WVS), wave 6 (ERIC 2019; European Social Survey ERIC (ESS ERIC) 2017; Inglehart et al. 2014). By combining these two sources, data for 81 countries and more than 170,000 respondents are available that were surveyed between 2010 and 2019. Since the study design does not include a cross-cohort perspective, it was attempted to collect data for a rather restricted time period to avoid the influence of history- and cohort-effects. The chosen design is a compromise between data richness concerning countries and a compact time frame. Effects due to intra-country developments, for example, economic prosperity and rise in standard of living, should not be too large since only 10 years are considered. Due to the nature of the design, some countries are included repeatedly at different points in time and control variables are added to account for this issue (see next section). The only sample restriction considers age. We decided to set the cut-off point to 86 years since above the number of available cases becomes very low and no precise results can be estimated for these groups. Since even in countries with the highest life-expectancy, the overall life-expectancy is below this age, there is no loss of information for all practical purposes. Regarding missing values, listwise deletion is employed. The overall percentage of missings regarding age or life satisfaction is very low (2.8%) and ranges from zero to 7%, depending on the country.

3.2 Operationalization

Age is operationalized as time in years between date of birth and time of survey. This variable takes integer values between 18 and 86. Life satisfaction is operationalized using the following question: "All things considered, how satisfied are you with your life as a whole these days?" Minor differences arise through scaling. The ESS uses an eleven-point scale from zero to ten, while the WVS uses a ten-point scale from one to ten. To create a common index of measurement, the ESS scale was transformed to a ten-point scale by combining the two lowest categories (zero and one) into one common category. These two were chosen since they are the ones with the fewest selections, and no large bias should arise due to the merging. To account for this problem and any other differences between ESS and WVS, a binary indicator is included in the regression models to incorporate the data source. Any differences in data collection, coding or processing are captured by this variable. The second control variable is the year in which the survey took place with values from 2010 to 2019 (coded as a categorical variable). This fixed-effect accounts for any variations over time if countries are included multiple times, which are not of interest for the present research question. This approach follows the suggestion of Glenn (2009), however, only as a robustness check since the timespan of 10 years can hardly reveal any cohort differences.

3.3 Strategy of Analysis

This section outlines which steps in the process of analysis will be taken to answer the research questions. All analyses are performed in Stata 16.1, complete do-files are available on request.

1. Life satisfaction is regressed on age, age², age³ separately for each country in the dataset using a linear (OLS) regression. This design does not interpose the restriction that the shape must be (reversed) U-shaped since age³ allows for more flexible functional forms and avoids certain methodological problems (Simonsohn 2018). This is an established approach and was demonstrated in related studies (Laaksonen 2018). The regression coefficients of age, age² and age³ are stored for all countries in a new datafile.
2. To classify countries based on the functional form between age and life satisfaction, a hierarchical Ward-clustering is applied to the stored regression coefficients. Since the functional form can be described using the three coefficients, this seems straightforward. Ward's method, which is part of the family of agglomerative hierarchical clustering processes, starts with each observation being a single cluster and then combines clusters based on the result which minimizes variance (Everitt 2011). This process is repeated until all clusters are fused into one large cluster containing all observations. The final number of clusters to be generated is based on a detailed inspection of the results by the researcher and justified by a high internal consistency and a large separation to all other clusters.
3. To explain cluster membership, global macro indicators are employed. Since the case number is too low for standard techniques of analysis like multinomial logistic regressions, ANOVAS are chosen as a more basic method. By doing so we attempt to explain why there are differences in functional form between age and life satisfaction for certain clusters.

4 Results

4.1 Descriptive Statistics and Overview

Before conducting the central analyses, some descriptive statistics should give an impression of the data. In total, there are 81 countries included. Table 1 lists central descriptive statistics for age and life satisfaction, separately for all countries. The next step is to assess the relationship between age and life satisfaction for the entire sample, including all countries. This gives a general overview of the global functional form between age and life satisfaction. To do so, a linear OLS regression is estimated as described above, including country fixed-effects, year of survey and source of data as control variables. Post stratification weights are utilized in these regressions to account for minor differences between official statistics and the sampling data provided by the institutes.³ For a detailed impression, three different models are estimated: The first includes only age and controls, the second model adds age² and the third model age³. By doing so it can be checked whether adding higher ordered terms for age does improve model fit. Here the question arises whether or not to include population weights. Without population weights the relative contribution of each country is solely determined by the number of respondents within the country. Since these numbers are quite similar (see Table 1), an unweighted calculation is not representative of the "world". To correct for this, population weights can be calculated so that larger

³ Note that there are no or incomplete post stratification weights available for the following countries: CL, CO, EC, GE, HK, HT, IN, IQ, JO, JP, KG, KW, LB, MA, MX, MY, NL (partially), NZ, PS, RW, SI (partially), TN, TT, UY, UZ, YE.

Table 1 Descriptive statistics for all countries. *Source:* ESS/WVS

Country	N	Age (mean)	Age (SD)	Age (median)	Life satisf. (mean)	Life satisf. (SD)	Life satisf. (median)	Source
AM	1096	46.58	17.94	46	5.23	2.55	5	WVS
AR	1015	43.00	17.38	40	7.48	1.64	8	WVS
AT	4391	50.91	17.17	51	7.77	1.83	8	ESS
AU	1438	53.28	16.34	55	7.38	1.97	8	WVS
AZ	1002	41.13	15.23	41	6.74	2.24	7	WVS
BE	3347	48.04	17.82	48	7.49	1.62	8	ESS
BG	2083	55.25	16.81	57	4.91	2.40	5	ESS
BR	1477	42.61	16.13	41	7.85	2.20	8	WVS
BY	1531	44.43	17.08	44	5.80	2.12	6	WVS
CH	2886	48.39	17.61	48	8.16	1.65	8	ESS
CL	988	43.96	16.24	44	7.27	1.82	8	WVS
CN	2252	43.93	14.96	43	6.86	1.98	7	WVS
CO	1512	40.41	15.79	39	8.39	1.79	9	WVS
CY	1728	47.14	18.48	46	7.01	2.12	7	Both
CZ	4530	48.15	16.77	48	6.85	1.84	7	ESS
DE	6967	49.79	17.51	51	7.52	1.95	8	Both
DZ	1165	37.44	14.92	34	6.31	2.40	6	WVS
EC	1199	39.69	15.95	37	7.92	1.74	8	WVS
EE	5257	49.82	18.06	50	6.68	2.03	7	Both
EG	1520	40.51	15.07	39	5.01	2.85	5	WVS
ES	4603	48.37	17.29	48	7.21	1.88	7	Both
FI	3521	51.20	17.94	53	8.00	1.57	8	ESS
FR	3869	52.31	17.59	53	6.48	2.22	7	ESS
GB	3960	51.69	17.52	52	7.29	2.04	8	ESS
GE	1200	44.66	17.60	43	5.45	2.43	5	WVS
GH	1552	30.92	12.70	26	6.42	2.35	7	WVS
HK	982	44.62	16.42	44	6.85	1.84	7	WVS
HR	1738	51.10	17.40	53	7.00	2.41	7.5	ESS
HT	1971	33.08	13.70	29	5.58	2.60	5	WVS
HU	3145	51.20	17.74	51	6.23	2.21	7	ESS
IE	4752	50.92	17.13	51	7.25	1.91	8	ESS
IL	2417	47.35	18.62	45	7.79	1.90	8	ESS
IN	4019	41.17	14.40	39	6.51	2.27	7	WVS
IQ	1191	36.59	13.39	35	5.91	2.18	6	WVS
IS	845	49.04	17.12	49	8.02	1.65	8	ESS
IT	5017	50.43	18.07	51	6.86	2.02	7	ESS
JO	1200	39.78	15.46	38	6.61	2.24	7	WVS
JP	2381	50.61	16.30	52	6.91	1.98	7	WVS
KG	1488	38.68	14.28	37	6.96	2.17	7	WVS
KR	1189	43.18	14.94	43	6.61	1.87	7	WVS
KW	1215	36.56	11.69	34	7.23	2.29	7	WVS
KZ	1497	39.92	15.22	38	7.26	2.10	7	WVS
LB	1193	38.38	14.86	37	6.50	2.12	7	WVS

Table 1 (continued)

Country	N	Age (mean)	Age (SD)	Age (median)	Life satisf. (mean)	Life satisf. (SD)	Life satisf. (median)	Source
LT	3738	53.41	17.24	55	6.07	2.19	6	ESS
LV	890	55.52	17.10	57	6.57	2.28	7	ESS
LY	2112	38.37	13.46	37	7.26	2.54	8	WVS
MA	1173	37.08	13.36	34	5.94	2.48	5	WVS
ME	1168	48.82	16.86	49	7.21	2.18	7	ESS
MX	1996	37.37	15.02	35	8.51	1.93	9	WVS
MY	1300	40.01	13.96	41	7.13	1.80	7	WVS
NG	1758	31.19	11.58	28	6.26	2.33	7	WVS
NL	5056	51.57	17.10	53	7.70	1.41	8	Both
NO	2763	47.84	17.30	48	7.90	1.65	8	ESS
NZ	810	50.90	16.57	51	7.64	1.95	8	WVS
PE	1205	39.39	16.36	37	7.13	2.19	7	WVS
PH	1199	42.68	15.52	41	7.33	2.81	8	WVS
PK	1200	34.34	11.86	33	7.48	2.10	8	WVS
PL	4002	48.37	17.80	48.5	7.08	2.05	7	Both
PS	995	36.70	14.14	35	5.62	2.49	6	WVS
PT	2243	52.48	17.51	53	6.42	2.27	7	ESS
QA	1051	37.71	12.74	37	8.01	1.93	8	WVS
RO	1487	48.32	17.20	48	6.64	2.39	7	WVS
RS	1961	53.67	17.21	55	5.97	2.67	6	ESS
RU	4822	46.40	17.41	45	5.91	2.24	6	Both
RW	1527	33.77	11.23	31	6.47	1.61	6	WVS
SE	4153	50.68	18.55	52	7.83	1.69	8	Both
SG	1933	41.75	16.45	40	6.96	1.60	7	WVS
SI	3525	49.72	17.48	50	7.23	2.03	8	Both
SK	1048	55.25	16.60	59	6.15	2.19	6	ESS
TH	1191	45.17	12.27	46	7.57	1.93	8	WVS
TN	1197	38.67	16.11	35	5.58	2.60	6	WVS
TR	1601	38.45	14.53	36	7.27	2.07	8	WVS
TT	991	45.60	17.49	45	7.46	2.11	8	WVS
TW	1210	45.08	17.09	46	6.89	2.07	7	WVS
UA	1498	47.18	18.20	47	5.90	2.45	6	WVS
US	2205	48.74	16.71	50	7.44	1.86	8	WVS
UY	987	44.77	18.07	42	7.59	2.02	8	WVS
UZ	1492	39.32	14.84	37	7.89	1.87	8	WVS
YE	983	35.45	12.96	32	5.90	2.66	6	WVS
ZA	3433	37.15	13.93	34	6.67	2.31	7	WVS
ZW	1499	33.73	13.43	30	6.04	2.24	6	WVS

Table 2 Regression results for all countries combined, post stratification weights only. *Source: ESS/WVS*

	M1	M2	M3
Age	-0.0080*** (0.0003)	-0.0293*** (0.0018)	-0.0480*** (0.0073)
Age ²		0.0002*** (0.0000)	0.0006*** (0.0002)
Age ³			-0.0000** (0.0000)
Constant	7.8243*** (0.0691)	8.2626*** (0.0784)	8.5110*** (0.1228)
Country fixed-effects	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes
Survey fixed-effects	Yes	Yes	Yes
Observations	175,731	175,731	175,731
R ²	0.117	0.118	0.118
AIC	753,644	753,458	753,450

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ **Table 3** Regression results for all countries combined, population weights added. *Source: ESS/WVS*

	M1	M2	M3
Age	-0.0058*** (0.0010)	-0.0300*** (0.0051)	-0.0880*** (0.0203)
Age ²		0.0003*** (0.0001)	0.0016*** (0.0005)
Age ³			-0.0000** (0.0000)
Constant	7.7232*** (0.0810)	8.1960*** (0.1268)	8.9575*** (0.2855)
Country fixed-effects	Yes	Yes	Yes
Year fixed-effects	Yes	Yes	Yes
Survey fixed-effects	Yes	Yes	Yes
Observations	175,731	175,731	175,731
R ²	0.079	0.080	0.081
AIC	766,886	766,710	766,645

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

countries have a larger impact on the results, according to the number of inhabitants.⁴ While this seems straightforward, it also means that the results will be dominated by very large countries like China, India or the US. For comparison, China receives a weight of 60, while a medium sized country like the UK receives a weight of about 4. This illustrates how a few countries might drive the results and small countries have only a marginal influence. We believe that it can be illustrative to have both results available and we present two regression tables, both weighted and unweighted. Results are displayed in Tables 2 and 3. Note that population weights also incorporate post stratification weights as well.

⁴ $W_{Pop} = N_{Population} / (N_{Sample} * 10,000)$. Population figures are taken from official government statistics (UN/Worldbank).

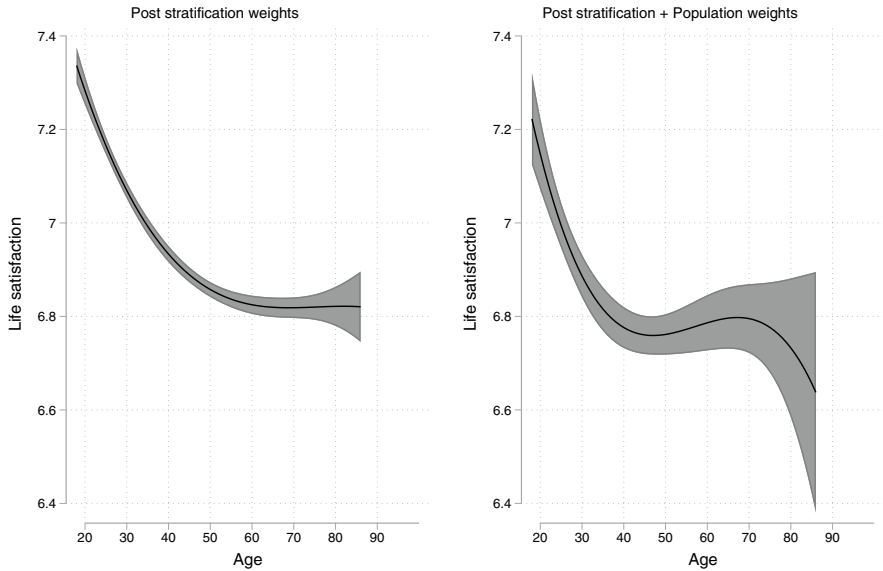


Fig. 1 Functional form between age and life satisfaction for all countries. *Source:* ESS/WVS. Predictions are based on M3

The results are similar between the weighted and the unweighted regressions. It becomes clear that adding higher-order terms for age does indeed improve the model fit. M3 with all three terms has the lowest AIC, which is an indicator of a good model fit.⁵ Consequently, all results indicate that M3 with three terms for age has the best fit. This model parametrisation will be adapted to all models in the following cross-country analyses. Before continuing with that, the global functional form should be visualized, including a 95% confidence band (Fig. 1). Here the effects of weighting become clearer. In the unweighted version we observe a linear decline which reaches a minimum at around age 60 and is stable with potentially minimally upward trends. For the weighted data, the outcomes are different and we observe a rather drastic decline after the age of 70. However, as confidence bands are very broad, the stability of the effect is unclear. The most pronounced difference between the weighted and unweighted analyses is the development of life satisfaction at old age. As explained before, it is quite obvious why differences can arise and only the weighted version can be used to predict the “average” human development of life satisfaction over the life-course.

As a further robustness check, we compare ESS-data to WWS-data using European countries that are included in both surveys (see Table 1). The predicted graphs (not depicted) indicate that the functional form is highly comparable (decline of life satisfaction over time with an upward trend from age 65) but the levels are different and the ESS-curve is about 0.5 points higher than the WWS-curve. Since we are only interested in the shape of the relationship and not the overall level, we believe that no significant bias should arise through the combination of both datasets.

⁵ Adding age⁴ to the model does, however, result in a worse fit and is therefore not regarded. Note that a LR test cannot be computed for weighted models.

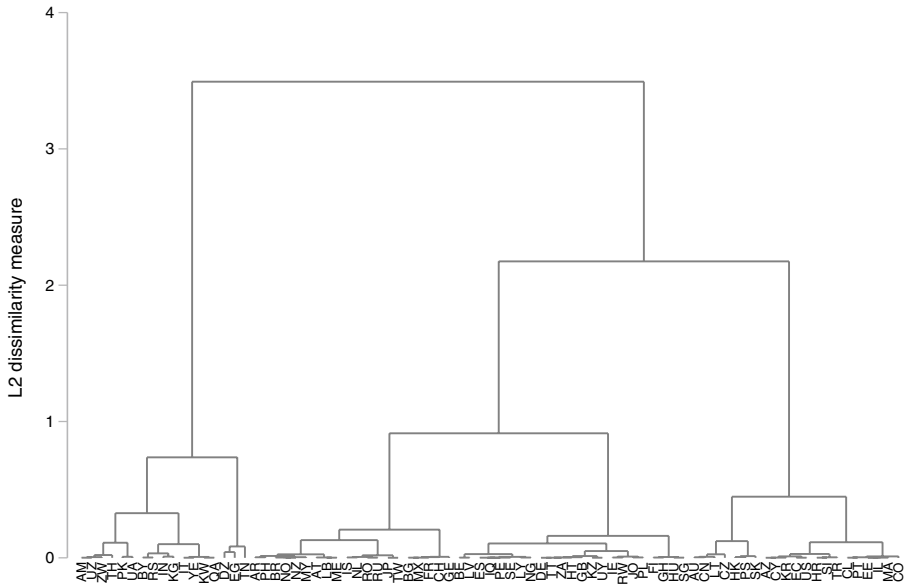


Fig. 2 Dendrogram created by the clustering procedure. *Source:* ESS/WVS

4.2 Clustering

After inspecting the global results, OLS regressions are estimated separately for all countries. The computed regression coefficients for age, age² and age³ are stored in a new file. After the regressions are estimated, this file is the base for the subsequent analyses. The next step is to estimate a hierarchical Ward-clustering using these three variables. Note that the regression intercept is not included in this clustering since only the functional *form* and not the absolute *level* of life satisfaction is of interest. Due to the nature of the procedure, as many clusters are generated as there are cases (countries) in the dataset (81). The Euclidean distance is chosen as a measurement for (dis)similarity. This procedure generates clusters which can be inspected using a dendrogram (Fig. 2).

The root of the dendrogram comprises all clusters, the leaves are clusters with only one case (country) left. Therefore, the leaves of the dendrogram are labelled with the country codes. The places where the roots split are the nodes. The longer the vertical distance between two nodes, the less similar the two resulting clusters. Based on this dendrogram and some other statistical tests we decided to create three final clusters. Choosing an appropriate number of clusters is always a compromise between the level of detail and separation. The lower the number of clusters, the lower the level of detail and the larger the separation of clusters. When many clusters are formed, the internal homogeneity is high, but it can be difficult to characterize differences between clusters. The number of three was chosen since the respective functional forms are quite different from each other. Choosing four clusters results in at least two clusters that are very similar regarding their functional form and it is unclear what characterizes them.

The countries were selected into the created clusters as follows:

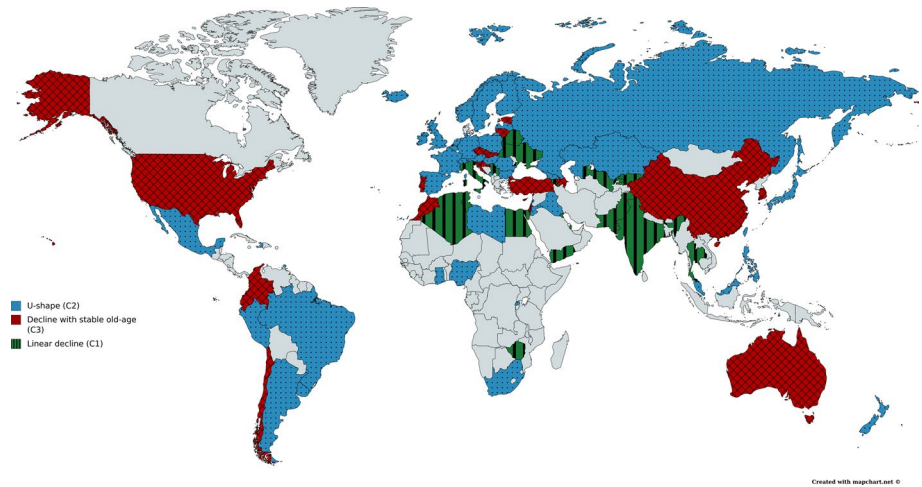


Fig. 3 Cluster results for all countries mapped. *Source:* ESS/WVS. Map created using <https://mapchart.net/>

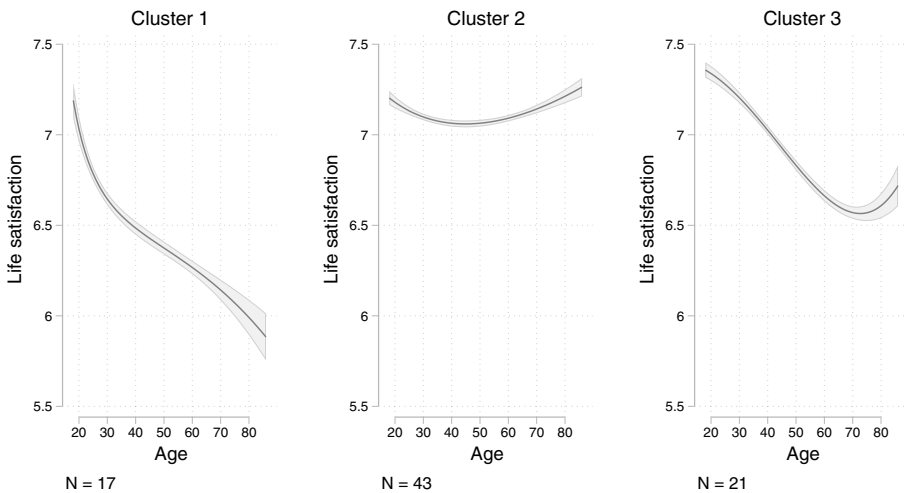


Fig. 4 Functional form between age and life satisfaction by cluster. *Source:* ESS/WVS. N refers to the number of countries within each cluster

Cluster 1 Algeria, Armenia, Belarus, Egypt, India, Italy, Kuwait, Kyrgyzstan, Pakistan, Qatar, Serbia, Thailand, Tunisia, Ukraine, Uzbekistan, Yemen, Zimbabwe.

Cluster 2 Argentina, Austria, Belgium, Brazil, Bulgaria, Finland, France, Georgia, Germany, Ghana, Haiti, Hungary, Iceland, Iraq, Ireland, Japan, Jordan, Kazakhstan, Latvia, Lebanon, Libya, Malaysia, Mexico, Montenegro, Netherlands, New Zealand, Nigeria, Norway, Peru, Philippines, Poland, Romania, Russian Federation, Rwanda, Singapore, South Africa, Spain, Sweden, Switzerland, Taiwan, Trinidad and Tobago, United Kingdom, Uruguay.

Cluster 3 Australia, Azerbaijan, Chile, China, Colombia, Croatia, Cyprus, Czechia, Ecuador, Estonia, Hong Kong, Israel, Korea (Republic of), Lithuania, Morocco, Palestine, Portugal, Slovakia, Slovenia, Turkey, United States.

For a more convenient interpretation, a global mapping is depicted in Fig. 3.

Next, the functional form between age and life satisfaction should be characterized. Here, we utilize fractional-polynomial prediction plots. The results are depicted in Fig. 4.

For cluster one, the relation is almost linear and can be described as monotonically declining. The older the individuals become, the lower their life satisfaction. No saturation-effects or plateaus are visible. The decline is substantial as the satisfaction drops from about 7.2 for young people to less than 6 for the oldest ones. For cluster two, the picture is clearly different, and a U-shape is visible. The initially high values of life satisfaction decline over the life course up to an age of around 50. From there on, the relation is inversed, and life satisfaction increases again. Interestingly, the overall variance is rather small and the curve always is between 7.0 and 7.5 over the entire age-range. Finally, for cluster three, there is a linear decline of life satisfaction up to an age of around 60 and from this point on it stays stable or even increases again slightly. For example, the USA are included in this third cluster. This result is in line with previous studies which also do not find the U-shape for this country when controls are omitted (Blanchflower and Oswald 2009; Glenn 2009). Taken together, it now becomes clearer why three clusters were chosen, since they display quite different functional forms. Two clusters would not be enough to capture all these inherent patterns, four or more clusters are too similar, and it is difficult to recognize where and why they differ. Therefore, generating three clusters appears to be an ideal solution for the data given.

A robustness check was performed regarding the choice of clustering algorithm by testing other techniques, for example, single-linkage or complete-linkage. However, none of these algorithms was able to extract as clear-cut patterns as the Ward-algorithm, resulting in clusters with sometimes very similar functional forms. As a conclusion, it seems that the chosen algorithm delivers the best results concerning the identification unique functional forms between age and life satisfaction.

4.3 Explaining Cluster Membership

The next question that arises is the following: Why do certain countries display some distinct functional form patterns between age and life satisfaction? As was described above, three major forms can be distinguished: linear decline (1), U-shape (2) and decline with stable old-age phase (3). When not countries but individuals are the subject of analysis, statistical techniques like multinomial logistic regressions or latent class analysis can be employed, which is not possible for the present dataset since only 81 countries (cases) are available, which is too low for mentioned methods. Therefore, a rather “low tech” approach is chosen which implies that means and standard deviations will be compared between clusters. Using ANOVAs as a generalization of the t test allows to test for significant group differences. It should be mentioned that this aspect of the analysis is of exploratory fashion since no strong theoretical assumptions are available. A few of them are discussed in the theoretical frameworks section. Tested are more than 100 global macro indicators which are collected and provided by the UN. These indicators are taken from 2016 if available, otherwise from adjacent years. This allows to test whether the distribution of these variables can explain group membership. Since the number of tested indicators is large, only

Table 4 Comparison of clusters regarding macro indicators. *Source:* ESS/WVS

Variable	All countries		Cluster 1		Cluster 2		Cluster 3		<i>p</i> value	N
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
Employment in services (%)	61.74	14.27	47.72	12.53	65.97	12.71	64.64	11.56	0.00	80
Internet users (%)	58.91	23.71	40.36	21.65	62.56	25.16	66.61	12.87	0.00	80
HDI	0.77	0.13	0.67	0.14	0.79	0.13	0.81	0.09	0.00	80
HDI rank	62.77	45.95	95.74	41.68	56.27	46.97	49.07	34.91	0.00	80
Expected years of schooling	14.71	2.47	12.98	2.22	15.06	2.57	15.41	1.84	0.00	80
Education Index	0.74	0.14	0.65	0.13	0.76	0.14	0.78	0.11	0.00	80
Gender Inequality Index	0.27	0.17	0.38	0.18	0.24	0.16	0.21	0.13	0.00	77
Income Index	0.78	0.14	0.70	0.16	0.80	0.13	0.81	0.09	0.01	80
Mean years of schooling (F)	9.87	2.80	8.17	3.07	10.27	2.60	10.45	2.52	0.02	80
Life expectancy at birth	72.63	6.28	69.58	5.44	72.70	7.15	74.96	3.66	0.03	80
Gender Development Index	0.95	0.07	0.91	0.11	0.97	0.05	0.96	0.05	0.03	80

p values computed using ANOVAs. Macro indicators taken from Human Development Report 2016 (<http://hdr.undp.org/>). Sorted by *p* values

those are included in the results (Table 4) which are statistically significant ($p < 0.05$) and of relevance to judge the total development of a country.

As is quite obvious, there are especially large differences between the first and the other two clusters. Clearly, the first cluster seems to include less developed countries, which is underlined by the lower average Human Development Index, the lower life expectancy at birth and the lower expected years of schooling. Characterising differences between clusters two and three is more difficult since these are usually highly similar for most statistical values.

5 Discussion

The present contribution demonstrates that the functional form between age and life satisfaction differs significantly between countries and that three larger clusters can be identified, which correspond to the following forms: linear decline (1), U-shape (2) and decline with rather stable old-age phase (3). Testing differences for macro indicators between clusters reveals that some significant effects are present. Since the linear decline of life satisfaction with age corresponds to less developed countries, one could speculate that weaker welfare-state regimes might affect life satisfaction for older people negatively. People in these countries usually cannot rely on a strong welfare-state and retirement system. Furthermore, when also the health system is not well developed, people might suffer more from the biological consequences of old age than people in other, more developed countries. This might explain the linear decline. Regarding the second cluster with the U-shape, we notice that these countries are predominantly located in Europe and Southern America. One could assume that these higher developed countries can provide more sophisticated welfare services, which might increase life satisfaction for old people. They can usually enjoy life after retirement with fewer responsibilities but with access to a modern health-care system, which might be beneficial for health and life satisfaction. However, the differences to the third cluster, with a stable old age phase without increasing life satisfaction,

is rather puzzling. Highly developed countries, like Australia, South Korea and the USA are found in this cluster, which means overall human development cannot be the reason. Another idea is to have a look at social inequality. While there are no statistically differences regarding the mean, the median of the Gini index is 32.4 in the first, 33.6 in the second and 36.2 in the third cluster, which hints for a higher inequality, albeit differences are small. Given the results of the tested indicators so far, the differences between the two more developed clusters remain rather unexplained and other theoretical approaches might be helpful to elucidate this aspect in future studies.

Finally, the limitations of the present contribution are discussed. Firstly, as the review of the literature underlines, there is no gold-standard available regarding the statistical approach and there exist several competing techniques. Instead of regressing life satisfaction on age, age^2 and age^3 , other methods identify the age-group with the lowest life satisfaction and use this value to split the regression into two parts, which are then separately analysed using a linear fit (Simonsohn 2018). While this seems valid, regarding the own questions and the clustering procedure, we chose to use regular regression coefficients instead. As the results demonstrate, this approach is also able to model functional forms flexible and generate distinct clusters. Secondly, all analyses are based on cross-sectional data. As usual, it is not possible to disentangle age-effects from cohort-effects using this kind of design. To conclude, one should also keep in mind that the entire functional form between age and life satisfaction was compressed using only three numbers per country, which means that the level of precision is not very high. While this enables us to compare a large number of countries simultaneously, this approach might be too coarse for detail analyses and different approaches might yield better results and especially inspecting each individual functional form by the researchers will produce a higher level of detail.

6 Conclusion

The study, which attempts to identify total causal effects of age on life satisfaction, finds that the functional form of the relation is dependent on the country and three global clusters can be identified. While there are strong theoretical and mathematically proven arguments for omitting mediating factors when a total effect is to be estimated, there are, of course, good reasons to include them. This should be done when mediators are at the focus of the analyses. Clearly, future studies should refer back to this question and explain in a cross-country perspective *why* the functional forms found in the present analyses arise. It would be highly interesting to check whether some mediators have a special effect for certain countries or clusters. The present contribution gives a general outline of how these analyses could be carried out. Relying on clustering approaches, especially when many countries are available, seems to be a valuable starting point to bring order to chaos and generate findings exploratively. On the other hand, when strong theoretical arguments and theories are available, they could also be tested in a causal fashion. As a conclusion, it seems central that researchers state their objectives clearly and generate their models accordingly. As there is a broad theoretical justification regarding the approximation of causal effects, this challenge should be manageable.

Acknowledgements The author wants to thank Viktoria Sophie Zorn for her helpful comments.

Funding Open Access funding enabled and organized by Projekt DEAL. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Compliance with Ethical Standards

Conflict of interest The author declares that he has no conflict of interest.

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

- Bartram, D. (2020). Age and life satisfaction: Getting control variables under control. *Sociology*. <https://doi.org/10.1177/0038038520926871>.
- Blanchflower, D. G., & Oswald, A. J. (2009). The U-shape without controls. *Economic Research Papers*. <https://doi.org/10.22004/ag.econ.271303>.
- 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>.
- Deaton, A. (2008). Income, health, and well-being around the world: Evidence from the gallup world poll. *Journal of Economic Perspectives*, 22(2), 53–72. <https://doi.org/10.1257/jep.22.2.53>.
- Easterlin, R. A. (2006). Life cycle happiness and its sources. *Journal of Economic Psychology*, 27(4), 463–482. <https://doi.org/10.1016/j.joep.2006.05.002>.
- 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>.
- (ERIC), E. S. S. E. (ESS). (2019). European Social Survey (ESS), Round 9—2018. <https://doi.org/10.21338/NSD-ESS9-2018>.
- European Social Survey ERIC (ESS ERIC). (2017). European Social Survey (ESS), Round 8—2016. <https://doi.org/10.21338/NSD-ESS8-2016>.
- Everitt, B. (Ed.). (2011). *Cluster analysis* (5th ed.). Chichester, West Sussex: Wiley.
- Frey, B. S., & Stutzer, A. (2005). Happiness research: State and prospects. *Review of Social Economy*, 63(2), 207–228. <https://doi.org/10.1080/00346760500130366>.
- 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>.
- Gangl, M. (2010). Causal inference in sociological research. *Annual Review of Sociology*, 36(1), 21–47. <https://doi.org/10.1146/annurev.soc.012809.102702>.
- 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 (66: 8, 2008, 1733–1749). *Social Science and Medicine*, 69(4), 481–485. <https://doi.org/10.1016/j.socscimed.2009.05.038>.
- 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(1), 177–197. <https://doi.org/10.1007/s11135-015-0300-3>.
- Hernán, M. A. (2018). The C-word: Scientific euphemisms do not improve causal inference from observational data. *American Journal of Public Health*, 108(5), 616–619.
- Inglehart, R., Haerfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., et al. (2014). *World values survey: Round six-country-pooled datafile version* (p. 12). Madrid: JD Systems Institute.
- Laaksonen, S. (2018). A research note: Happiness by age is more complex than U-shaped. *Journal of Happiness Studies*. <https://doi.org/10.1007/s10902-016-9830-1>.
- Lelkes, O. (2008). Happiness across the life cycle: Exploring age-specific preferences. *Policy Brief*, 2, 1–16.

- López Ulloa, B. F., 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>.
- Maddux, J. E. (2018). *Subjective well-being and life satisfaction: An introduction to conceptions, theories, and measures*. London: Routledge/Taylor & Francis Group.
- Matthews, R. (2000). Storks deliver babies ($p=0.008$). *Teaching Statistics*, 22(2), 36–38.
- Mroczek, D. K., & Kolarz, C. M. (1998). The effect of age on positive and negative affect: A developmental perspective on happiness. *Journal of Personality and Social Psychology*, 75(5), 1333.
- Pearl, J. (2009a). *Causality*. Cambridge: Cambridge University Press.
- Pearl, J. (2009b). Causal inference in statistics: An overview. *Statistics Surveys*, 3, 96–146. <https://doi.org/10.1214/09-SS057>.
- Ryff, C. D., Singer, B. H., & Dienberg Love, G. (2004). Positive health: Connecting well-being with biology. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 359(1449), 1383–1394. <https://doi.org/10.1098/rstb.2004.1521>.
- Shahar, E., & Shahar, D. J. (2013). Causal diagrams and the cross-sectional study. *Clinical Epidemiology*, 5, 57. <https://doi.org/10.2147/CLEP.S42843>.
- 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>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.