

Chapter 2

Different Versions of the Easterlin Paradox: New Evidence for European Countries



Caspar F. Kaiser and Maarten C. M. Vendrik

Abstract According to the Easterlin Paradox, richer people are happier than poorer people, but when a country becomes richer over time, its people do not become happier. There is debate on whether this paradox holds. To shed light on this controversy, we distinguish between five different versions of the paradox. They apply to either groups of countries or individual countries, and to either the long or the medium term. We argue that the long term is most appropriate for testing the paradox, and that tests of the paradox should control for an autonomous time trend. We conduct such tests by estimating country-panel equations for mean life satisfaction in 27 European countries that include trend and cyclical components of per capita GDP as regressors. Concerning groups of countries, we find a robust confirmation of the long- and medium-term versions of the paradox for a group of nine Western and Northern European countries. Moreover, we obtain a non-robust rejection of the medium-term variant of the paradox for a set of 11 Eastern European countries. Regarding individual countries, the medium-term variant of the paradox holds for the nine Western and Northern European countries, but is rejected for Greece, Ireland, Italy, Spain, Bulgaria, Lithuania, and Poland.

This is an abridged version of a longer working paper with the same title (see Kaiser and Vendrik 2018)

C. F. Kaiser
Nuffield College, Oxford University, Oxford, UK
DSPI, Oxford University, Oxford, UK

M. C. M. Vendrik (✉)
Department of Macro, International and Labour Economics (MILE), SBE, Maastricht University, Maastricht, The Netherlands
ROA, SBE, Maastricht University, Maastricht, The Netherlands

IZA, Bonn, Germany

EHERO, Erasmus University, Rotterdam, The Netherlands
e-mail: m.vendrik@maastrichtuniversity.nl

2.1 Introduction

It has been more than 40 years since Easterlin published his path-breaking study “Does economic growth improve the human lot: Some empirical evidence” (1974). In that and later papers (Easterlin 1995, 2005, 2017), he showed that while at a point in time individual happiness is positively correlated with individual income in the USA and other countries, over time average happiness in these countries does not trend upward as average income continues to grow. This seemingly contradictory pair of findings has become famous as the “Easterlin Paradox”. Although these paradoxical findings have been confirmed for several other developed countries by other happiness researchers (e.g., Layard et al. 2010; Clark et al. 2014), there are also happiness scientists (e.g., Stevenson and Wolfers 2008; Sacks et al. 2012, 2013; Veenhoven and Hagerty 2006; Veenhoven 2011; Veenhoven and Vergunst 2014; Diener et al. 2013a) who have presented counterevidence to the Easterlin Paradox. Although Easterlin (2017) has convincingly pointed out several shortcomings in the contestants’ studies, this still raises the question as to who is right.

In this study we investigate this issue both on a conceptual level and by conducting our own estimations on country panel models that are similar to those of Layard et al. (2010) and Sacks et al. (2013), using updated life satisfaction data from the Eurobarometer surveys. On a conceptual level we show that in the debate on the Easterlin Paradox at least two distinct versions of this paradox are discussed. The first version refers to individual countries and has been formulated above. The second version of the paradox extends the first part of the paradox to the positive correlation between average happiness and GDP per capita across countries (see, e.g., Deaton 2008, Easterlin 2017) and contrasts it with a zero cross-country correlation between (annual) rates of change in average happiness and GDP per capita over time. This seems like a mere cross-sectional reformulation of the paradox for groups of countries. However, there is an essential difference compared to time-series regressions that test whether individual countries with a positive rate of economic growth also experienced a positive time trend in happiness. In the cross-country regression, average annual rates of change in SWB are not only regressed on average annual rates of economic growth, but also on a constant (see, e.g., Easterlin 2017, Table 1). This constant picks up drivers of (linear) trends in SWB other than economic growth that are common to all countries (e.g., trends in marriage and divorce rates, social capital, trust, aging, and income inequality; see Angeles 2011; Bartolini and Sarracino 2014; Bartolini et al. 2013a, b; Gruen and Klasen 2013).

On the level of individual countries, this suggests that when a time-series regression of SWB of a specific country with positive economic growth reveals a significant positive time trend in SWB, this trend could be driven by trends in other determinants of SWB than economic growth (see also Clark 2011, p. 259). In such a case the positive time trend in a country’s SWB does not imply a non-spurious positive correlation between SWB and long-term economic growth in that country. Therefore, a reliable test of the paradox should in our view at least control for possible spuriousness arising from time trends in other determinants of SWB. Hence,

to reliably test the Easterlin Paradox for individual countries, one should regress SWB in a country on the long-term economic growth trend while controlling for a country-specific autonomous time trend. Unfortunately, this is not possible due to perfect collinearity of such a time trend with the time-linear long-term economic growth trend. Thus, reliable tests of the Easterlin Paradox for separate individual countries do not seem possible.

However, there are two partial ways out of this problem. First, instead of controlling for a country-specific autonomous time trend, one may control for specific other determinants of SWB. But such an approach raises the thorny question which other determinants of SWB are predetermined with respect to per capita GDP, and hence should be controlled for (“good” controls in the terminology of Angrist and Pischke 2009), and which determinants are mediating the effect of per capita GDP on SWB, and hence should not be included when wishing to estimate the total correlation of per capita GDP and SWB over time (“bad” controls). Moreover, the selected good control variables may not capture all autonomous determinants of SWB that vary in a linear-trend-like fashion. To circumvent these problems, one may adopt a country-panel approach to testing the Easterlin Paradox as introduced by Layard et al. (2010) and also used by Sacks et al. (2013). In this approach, which we follow in the present study, real GDP per capita (GDPpc) data are corrected for short-term business cycle effects¹ by means of a Hodrick-Prescott (HP) filter (Hodrick and Prescott 1997), and the resulting GDPpc trend and cyclical components are used as regressors in panel regressions for the average SWB in countries. In this study we use two variants of the HP filter: The first one sets the parameter λ of the HP filter to its conventional value 6.25 for annual data (see Ravn and Uhlig 2002), which is also used by Sacks et al. (2013).² This filters out fluctuations in GDPpc due to business cycles of up to about 8 years of length (as defined by Burns and Mitchell 1946; see, for example, Fig. 2.1 (left) for the Netherlands).³ In our second HP filter, we set $\lambda = \infty$. This filter is equivalent to the least-squares fit of a linear trend model for GDPpc with a slope coefficient given by the average growth rate of GDPpc over the whole estimation period (see Fig. 2.1 (right)). This filter also corresponds to the average growth rate used as regressor in the SWB regressions of Easterlin (2017) and Veenhoven and Vergunst (2014) and filters out all cyclical fluctuations within the estimation period. In particular, in the case of the transition of ex-communist countries from communism to capitalism, the linear trend filter filters out contraction-expansion cycles, which may take up to 20 years, and hence last

¹See Easterlin’s (2017 Sect. 2) distinction between short-term fluctuations and long-term trends in GDPpc.

²Layard et al. (2010) adopt a value of 9.5 for λ , but mention in their note of Table 6.5 that setting $\lambda = 6.25$ produces similar results.

³Hamilton (2017) criticizes the HP filter for introducing spurious dynamic relations in the cyclical component that have no basis in the underlying data-generating process. However, for our purposes of regressing SWB on primarily an appropriate GDPpc trend measure, the HP filter seems more suitable than the alternative filter that is presented by Hamilton (2017). Moreover, this alternative filter generates similar results. See Kaiser and Vendrik (2018) for further discussion of this.

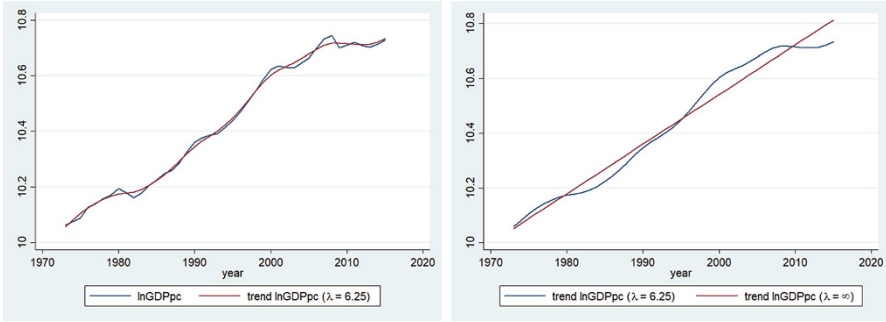


Fig. 2.1 Time paths for the Netherlands of: (left) $\ln\text{GDPpc}$ vs. trend $\ln\text{GDPpc}$ for $\lambda = 6.25$, and (right) trend $\ln\text{GDPpc}$ for $\lambda = 6.25$ vs. trend $\ln\text{GDPpc}$ for $\lambda = \infty$

much longer than the usual business cycles. Easterlin (2017) makes the point that for allowing the average growth rate GDPpc to filter out such transition cycles, the estimation period should be long enough, i.e. at least roughly 20 years for transition countries. Generally, in order to test for a long-term correlation between SWB and GDPpc in countries over time, the most appropriate filter of GDPpc is one that corrects for all cyclical fluctuations - no matter their duration. Such a filter is the linear trend filter of GDPpc with $\lambda = \infty$, which thus seems more suitable for this purpose than HP filters of GDPpc with lower values of λ as used by Layard et al. (2010) and Sacks et al. (2013).

We distinguish between tests of the paradox that apply to groups of countries and tests that apply to individual countries. Unfortunately, the linear time trend filter can only be used for testing the Easterlin Paradox for groups of countries. In the case of tests for individual countries, one cannot control for an autonomous time trend when using such a filter due to perfect collinearity of the time trend with the filtered GDPpc series. However, a HP filter of GDPpc with $\lambda = 6.25$ can be used since it is not perfectly collinear with a linear time trend. Because a HP filter of GDPpc with $\lambda = 6.25$ only corrects for business cycle fluctuations up to about 8 years of length, we refer to estimations with such a filter as tests of a medium-term version of the paradox. Thus, we conceptually distinguish between long-term and medium-term variants of the paradox.

Therefore, combining all the distinctions made above, we can distinguish the following five variants of the Easterlin Paradox:

EPgl: Whereas at a point in time happiness varies positively with income both among and within countries, over time countries with a higher long-term rate of economic growth in a certain group of countries do not exhibit a more positive change in average happiness when controlling for a common time trend (Easterlin 2017, p. 316; Veenhoven and Vergunst 2014).

EPgm: Whereas at a point in time happiness varies positively with income both among and within countries, over time countries with a higher medium-term rate of economic growth in a certain group of countries do not exhibit a more positive

change in average happiness when controlling for a common time trend (Layard et al. 2010; Sacks et al. 2013).

EPi0: Whereas at a point in time happiness varies positively with income within countries, over time average happiness in a particular individual country does not trend upward as average income trends upward (Easterlin 1974).⁴

EPil (not testable!): Whereas at a point in time happiness varies positively with income within countries, over time a higher long-term rate of economic growth in a particular individual country is not associated with a more positive change in average happiness when controlling for a country-specific time trend.

EPim: Whereas at a point in time happiness varies positively with income within countries, over time a higher medium-term rate of economic growth in a particular individual country is not associated with a more positive change in average happiness when controlling for a country-specific time trend.

We test these different versions of the Easterlin Paradox except the non-testable EPil for European countries by estimating country-panel equations for mean life satisfaction that include long- or medium-term trend and cyclical components of GDPpc and country dummies as regressors. Throughout, we take the first parts of the paradox' variants (*i.e.* correlations of happiness and income within and among countries) for granted because their validity has been confirmed in numerous empirical studies (e.g., Deaton 2008; Sacks et al. 2012, 2013). To account for heterogeneity in the correlations of mean life satisfaction and trend GDPpc between different country groups, we partition our total sample of 27 countries into subsamples consisting of Western and Northern European, Southern European, and Eastern European countries.

Our main results are as follows. Concerning groups of countries, we find a clear and robust confirmation of the paradox for the long as well as medium term for a group of nine Western and Northern European countries. Moreover, we obtain a non-robust rejection of the paradox for the medium term for a set of 11 Eastern European countries. Concerning individual countries, the medium-term version of the paradox (EPim) clearly holds for the nine Western and Northern European countries, but is significantly rejected for Greece, Ireland, Italy, and Spain. Thus, in the latter four as opposed to the former nine countries, economic growth was positively associated with changes in life satisfaction in the medium term. Regarding the individual Eastern European countries, this also holds for Bulgaria, Lithuania, and Poland, but for the other EE countries results are unreliable, partially due to the limited length of the time series (11 years).

The remainder of this paper is organized as follows. Section 2.2 reviews the state of the debate on the Easterlin Paradox in the literature. In Sect. 2.3 the estimation equations for the tests of the different versions of the paradox are explained. Section

⁴We do not follow Easterlin (2017, p. 312) in extending the formulation of variant EPi0 to also include the variation of happiness with income across countries. This is because EPi0 only concerns the variation of happiness with income *within* countries. The same holds for the individual-country variants EPil and EPim.

2.4 presents the data and descriptive statistics. Then, Sect. 2.5 discusses the estimation results for groups of countries and individual countries, respectively. Finally, Sect. 2.6 draws some general conclusions.

2.2 State of the Debate

Variant EPi0 (non-positive time trend in average happiness) of the Easterlin Paradox has been tested and confirmed by Easterlin (1974, 1995, 2017) for the USA, by Easterlin (1995, 2005) and other happiness researchers (e.g., Layard et al. 2010, and Clark et al. 2014) for many other developed countries, and by Easterlin (2009) for several transition countries. On the other hand, Veenhoven (2011) has estimated trends in mean life satisfaction for 15 developed countries over the period 1970–2010 and has found significant positive trends for seven of these countries. As GDPpc trended upwards in the period considered in all the 15 countries, Veenhoven’s results imply a rejection of EPi0 for the seven developed countries with significant positive trends in life satisfaction. Similarly, Sacks et al. (2012) report that six out of nine European countries in the period 1973–1989 show a significantly positive regression relationship between average life satisfaction and $\ln(\text{GDPpc})$ (see their Fig. 6). Because GDPpc trended upward in all the nine countries, these regressions may be interpreted as tests of EPi0, with the important limitation that these tests do not correct for business-cycle fluctuations in GDPpc. However, as argued above, EPi0 is not an appropriate version of the Easterlin Paradox and should be replaced with the country-specific medium-term variant EPim of the paradox, as this controls for an autonomous time trend, and hence for possible spuriousness driven by trends in other determinants of happiness.

Most tests of the Easterlin Paradox in the literature are tests of EPgl and EPgm on the level of groups of countries. The long-term version EPgl has been tested using cross-country regressions of average rates of change in SWB on average growth rates of GDPpc by Easterlin, Veenhoven, and their co-workers. On the one hand, Easterlin and colleagues (see, e.g., Easterlin et al. 2010; Easterlin and Sawangfa 2010; Easterlin 2015, 2017) consistently find confirmations of EPgl for groups of developed countries, developing countries, transition countries, and all countries taken together. On the other hand, Veenhoven and Vergunst (2014) find a rejection of EPgl for a large combined data set of countries and attribute the differences of their results with those of Easterlin et al. (2010) to the comparatively much larger size of their data set. Furthermore, they find that the correlation between happiness and economic growth is quite strong in the 20 lower-income nations in their data set and relatively small in the high-income nations (Table 4b). However, Veenhoven and Vergunst’s approach is extensively criticized by Easterlin (2017).

Layard et al. (2010) and Sacks et al. (2013) also test the Easterlin Paradox on the level of groups of countries using country-panel regressions. However, they test the time-series correlation of SWB with (less appropriate) medium-term rather than long-term trends in GDPpc because they use HP filters with $\lambda = 9.5$ and 6.25, respectively. Employing Eurobarometer data for average life satisfaction in a group

of 16 mainly Western European countries over the period 1973–2007, Layard et al. (2010) find insignificant coefficients of medium-term trend GDPpc in panel regressions of average life satisfaction while controlling for country-fixed effects, a time trend or year dummies, the cyclical GDPpc component, the unemployment rate, and the inflation rate. In our terminology, they thus test for and confirm EPgm for this group of Western European countries. However, the control for the unemployment rate may cause underestimation of the total effect of medium-term trend GDPpc, as parts of that effect may run via induced medium-term changes in the unemployment rate. Contrariwise, Sacks et al. (2013), using several data sets for average SWB in groups of countries all over the world and estimating country-panel regressions of average SWB on medium-term trend GDPpc similar to those of Layard et al., find significant positive correlations of SWB and trend GDPpc in most of their data sets for the world as a whole. Moreover, when using Eurobarometer data for average life satisfaction (in a group of 30 European countries over the period 1973–2009), they find a significant positive correlation of SWB and trend GDPpc as well. However, they do not find significant correlations for their Gallup World Poll data set for a “ladder-of-life” version of SWB in a world-wide group of 141 countries in the period 2005–2011 and for Latinobarometro data for average life satisfaction in 18 Latin American countries in the period 2001–2010.

An interesting study by Proto and Rustichini (2013) moves the analysis forward by analysing the relation between GDPpc and life satisfaction without imposing a functional form on the term for GDPpc. They specify the variation of GDPpc in terms of quantiles and run micro-macro-panel regressions of life satisfaction data from the World Values Survey and Eurobarometer on the GDPpc quantiles while controlling for country- and year-fixed effects, individual employment status, and personal income. These regressions reveal a non-monotonic relation between GDPpc and life satisfaction which is significantly positive for poorer countries and regions, but becomes insignificant for richer countries/regions, and even turns significantly negative for the richest countries/regions. This suggests a rejection of the medium-term variant EPgm of the Easterlin Paradox for poorer countries and regions, but not necessarily of the more appropriate, long term variant EPgl because the time series for the poorer countries and regions are too short for that. Another limitation of these tests is that the use of controls for individual employment status and personal income may either lead to an overestimation of the medium-term effects of GDPpc since effects of country-specific business cycles other than on individual employment status and personal income are not controlled for, or lead to an underestimation of the total medium-term effects of GDPpc since parts of that effect may run via induced medium-term changes in individual employment status and personal income. This ambiguity makes the use of these controls problematic.⁵

⁵Beja (2014) and Opfinger (2016) also test for the Easterlin Paradox, but in our view the dynamic model of Beja is mis-specified (missing levels of lnGDPpc), and Opfinger only uses the last two waves of the WVS, which implies an estimation period of 5–7 years that is much too short to test the Easterlin Paradox in a reliable way. Furthermore, there is no control for cyclical fluctuations in GDPpc.

2.3 Estimation Strategy

In this paper we focus on estimations that use the long-term HP filter (with $\lambda = \infty$) for groups of countries and the medium-term HP filter (with $\lambda = 6.25$) for individual countries as these most closely correspond to the two most appropriate testable versions of the paradox (EPgl and EPim; see Sect. 2.1). See Kaiser and Vendrik (2018) for extensive discussion of our estimation strategy and results when using the medium-term filter for groups of countries (corresponding to EPgm) and the long-term filter for individual countries (corresponding to EPi0).

2.3.1 Estimation Equations for Testing the Country-Group Variants of the Easterlin Paradox

We begin with our approach to testing the long-term group variant EPgl of the Easterlin Paradox. The baseline equation has the form

$$LS_{ct} = \beta \text{ trend } \ln GDPpc_{ct} + \gamma \text{ cyclical } \ln GDPpc_{ct} + \sum_{t'} \delta_{t'} d_{t'} + \sum_{c'} \alpha_{c'} d_{c'} + \varepsilon_{ct}, \quad (2.1)$$

where LS_{ct} is mean life satisfaction in country c in year t , $\text{trend } \ln GDPpc_{ct}$ and $\text{cyclical } \ln GDPpc_{ct}$ are the long-term ($\lambda = \infty$) trend and cyclical components of $\ln GDPpc_{ct}$, $d_{t'}$ and $d_{c'}$ represent year and country dummies, and ε_{ct} is the error term. The year and country dummies account for, respectively, year-specific country-invariant determinants like differences in survey design across waves and common time trends and shocks, and country-specific time-invariant determinants like institutions and cultural differences in SWB scale use. The error term is clustered over countries to account for heteroscedasticity and serial correlation, which both occur in our estimations (Angrist and Pischke 2009, Ch. 8).

To test for EPgl⁶ we then conduct two-tailed t tests of a null hypothesis of equality to zero of the parameter β of $\text{trend } \ln GDPpc_{ct}$ against the alternative hypothesis of non-equality of β to zero. If such tests fail to reject the null hypothesis or if the sign of β is negative, EPgl is confirmed. If the null hypothesis is rejected and the sign of β is positive, EPgl is rejected. Alternatively, we conduct one-tailed tests of the null hypothesis $\beta \leq 0$ against the alternative hypothesis $\beta > 0$. If such tests fail to reject the null hypothesis, EPgl is confirmed, whereas a rejection of the null hypothesis implies a rejection of EPgl. As p values in one-tailed tests are half of those in the

⁶When using an HP filter with $\lambda = \infty$ for an estimation period of much less than 20 years (see Sect. 2.1), the following applies to the medium-term version EPgm.

two-tailed tests, EPgl will more easily be rejected at conventional significance levels by the one-tailed tests than by the two-tailed tests.

The estimate of β is only driven by cross-country variation in the trend GDPpc growth rates, since trend GDPpc growth rates in individual countries are constant over time (see Fig. 2.1 (right)). These trend GDPpc growth rates correspond to the average long-term GDPpc growth rates that are used as regressor in regressions of average annual SWB changes in the methodology of Easterlin et al. and Veenhoven and Vergunst (2014). However, a difference with our approach is that Easterlin et al. and Veenhoven and Vergunst follow a two-step procedure in which they first estimate long-term average rates of changes in mean SWB and GDPpc (in percentages) and then regress these average rates of change on each other, whereas we directly regress mean SWB levels in the countries on long-term trend \ln GDPpc over time. A disadvantage of Easterlin's and Veenhoven and Vergunst's procedure is that the estimated average rates of change in mean SWB tend to be unstable (i.e. sensitive to adding or dropping observations) in samples with few observations per country (e.g. the WVS). According to a conventional rule of thumb in econometrics, stable estimates of regression coefficients require an amount of observations which is at least ten times the number of explanatory variables in the regression. Although the resulting measurement error in SWB trends may be random in large country samples, it may raise standard errors in the regression coefficients of the long-term GDPpc growth rate and therefore decrease the chances of rejecting EPgl. In the country-panel approach of Layard et al. (2010) and Sacks et al. (2013) that we follow, this complication is avoided by directly regressing SWB levels in countries on trend \ln GDPpc over time with enough panel observations to get stable, and hence reliable, estimates of coefficient β of $trend \ln GDPpc_{ct}$ in Eq. 2.1.

A concern in our country-panel approach is that with a clustered error term, the asymptotic standard errors of the regression coefficients need to be corrected for the low number of clusters, i.e. countries, in the sample and subsamples that we use (from 4 to 27; about 50 is the minimal required number of clusters, see Cameron and Miller 2015, Section VI). Therefore, we employ the command *regress y x, vce (cluster)* in Stata, which includes a finite-sample adjustment of the cluster-robust standard errors and uses a T distribution with $G-1$ degrees of freedom instead of a standard normal distribution for t -tests based on these standard errors (G denotes the number of clusters). However, even with both adjustments, Wald tests tend to over-reject (op. cit.). For us, the remaining downward bias in the cluster-robust standard errors will lead to too high a likelihood of rejection of the null hypothesis (either $h_0: \beta = 0$ or $h_0: \beta \leq 0$) of $trend \ln GDPpc_{ct}$ in Eq. 2.1, and hence rejection of the paradox. Therefore, we need a more reliable test, which we obtain by correcting for the first-order serial correlation over time more directly than by clustering standard errors over countries. Such correlation signals the joint effect on life satisfaction of lags of trend and cyclical \ln GDPpc and lags of and serial correlation in time-varying omitted variables (see Vendrik 2013, and Angrist and Pischke 2009, Sect. 8.2.2), which implies that Eq. 2.1 represents a dynamically incomplete model. The serial correlation, and hence the resulting downward bias in the standard errors of the parameter estimates, can be largely reduced by making Eq. 2.1 dynamically more

complete with the addition of one-year lagged mean life satisfaction to the right-hand side of Eq. 2.1.⁷ This yields

$$LS_{ct} = \beta \text{ trend } \ln GDPpc_{ct} + \gamma \text{ cyclical } \ln GDPpc_{ct} + \sum_t \delta_t d_t + \varphi LS_{ct-1} + \sum_{c'} \alpha_{c'} d_{c'} + \varepsilon_{ct}. \quad (2.2)$$

The lagged life satisfaction term picks up the joint effect of lags of trend and cyclical $\ln GDPpc$ and lags of and serial correlation in time-varying omitted variables. As the estimate of parameter φ turns out to be significantly positive in our estimations, the initial effect⁸ $\beta \Delta \text{trend } \ln GDPpc_{ct}$ of a change in $\text{trend } \ln GDPpc_{ct}$ in year t on life satisfaction is reinforced in year $t + 1$ by $\varphi \beta \Delta \text{trend } \ln GDPpc_{ct}$, in year $t + 2$ by $\varphi^2 \beta \Delta \text{trend } \ln GDPpc_{ct}$ etc. In the end, this reinforcement process will converge to a total long-run effect $\frac{\beta}{1-\varphi} \Delta \text{trend } \ln GDPpc_{ct}$ of the change in $\text{trend } \ln GDPpc_{ct}$ in year t on life satisfaction (see Vendrik (2013) for more complete dynamics⁹). In this case, EPgl is tested via a null-hypothesis of equality to zero or non-negativity of the long-run effect $\frac{\beta}{1-\varphi}$ of $\text{trend } \ln GDPpc_{ct}$.

The dynamic-model concept of a long-run effect should be distinguished sharply from the concept of a long-term effect in the macro-economic time series context of the analysis of the Easterlin Paradox. Whereas 90% convergence to a long-run life satisfaction equilibrium usually takes place within a range of 1–11 years¹⁰, the expression “long term” refers to time periods of at least 20 years or so. In the presence of country-fixed effects our estimate of φ in Eq. 2.2 will suffer from a downward Nickell bias. To correct for this Nickell bias, we apply a bias-corrected least squares dummy variables (BCLSDV) estimator in Stata to correct for the Nickell bias in the coefficient of lagged life satisfaction (see Bruno 2005a), for the underlying econometrics). The command for this estimator calculates bootstrap standard errors of the parameter estimates of Eq. 2.2, which are sufficiently reliable

⁷Because our regressors are likely not strictly exogenous, eliminating the serial correlation by a Prais-Winston or Cochrane-Orcutt transformation of the error term would not lead to consistent and efficient standard errors of the parameter estimates (Wooldridge 2003, Sects. 12.3 and 12.5).

⁸We here use the term “effect” rather than “correlation” because a dynamic model like Eq. 2.2 usually presupposes causality from the right-hand-side variables to the left-hand-side variable of the equation. Although testing of the Easterlin Paradox only involves correlations, dynamic-model concepts like short and long-run effects are more generally applicable to correlations as well.

⁹In that analysis adaptation of individual life satisfaction to income changes is modelled. In the simplified dynamics in the present paper such adaptation is implicitly and partially incorporated in the contemporaneous effects of the trend and cyclical $\ln GDPpc$ variables. See, however, the end of the next section for an extension that explicitly models adaptation of life satisfaction to medium-term changes in $\ln GDPpc$.

¹⁰The number of years τ within which convergence for 90% takes place can be calculated as $\varphi^\tau = 1 - 0.9 = 0.1$ or $\tau \ln \varphi = \ln 0.1$ or $\tau = \ln 0.1 / \ln \varphi$. For estimates of φ between 0.1 and 0.8 this yields $1.0 < \tau < 10.3$ (cf. Vendrik, 2013).

when the remaining serial correlation of the error term of Eq. 2.2 turns out to be weak.

In line with Easterlin (2017), we apply two criteria for including countries in our tests of the Easterlin Paradox. First, to obtain a less heterogeneous sample in terms of population size, countries must have more than one million inhabitants. Second, the available surveys for average life satisfaction in a country should minimally span 10 years and at least one complete cycle of GDPpc.

2.3.2 Estimation Equations for Testing the Individual-Country Variants of the Easterlin Paradox

Apart from our above argument that variant EPi0 of the Easterlin Paradox is not an appropriate version of this paradox, a limitation in the estimation of time trends in average happiness in individual countries as conducted in the literature (see Sect. 2.2), is that these estimations do not control for differences in survey design across waves. It is not possible to obtain reliable estimates of time trends of average happiness in individual countries from separate regressions while controlling for wave or time-fixed effects because such fixed effects then pick up part of the time trend. A partial solution to this problem is offered by Easterlin (2017, p. 319). He estimates time trends of average happiness in individual countries by adding interactions between country dummies and year to a country-panel regression of average happiness on year while controlling for country-fixed effects as well as two dummies for specific changes in survey design.

To test medium-term variant EPim¹¹ of the Easterlin Paradox for individual countries separately, we extend this approach in several directions. Firstly, we replace in Eq. 2.1 the main effects of *trend* $\ln GDPpc_{ct}$ and *cyclical* $\ln GDPpc_{ct}$ for $\lambda = 6.25$ by their interactions with country dummies, and add interactions of a time trend with the country dummies. Secondly, we drop the year-fixed effects as they would otherwise pick up part of the time trend for the reference country of the country dummies. Thirdly, since standard errors of the interaction coefficients based on clustered error terms implode as the effective number of clusters for each country-specific coefficient estimate is only one, the serial correlation in the error terms is now controlled for by adding interactions of one-year lagged life satisfaction (cf. Eq. 2.2) with the country dummies to Eq. 2.1. Fourthly, to control for different preceding questions affecting responses to the life satisfaction equation, we select waves such that the number of distinct preceding questions across time is minimised. We then include dummies for the remaining different preceding questions in our estimation equations. Because the number of these dummies is still large (ten), insignificant dummies are dropped from the regressions (see footnote 26 for more details).

¹¹See Kaiser and Vendrik (2018) for our similar approach to testing EPi0.

Implementing all these modifications results in an estimation equation of the form

$$\begin{aligned}
 LS_{ct} = & \sum_{c'} [\beta_{c'} d_{c'} trend \ln GDPpc_{c't} + \gamma_{c'} d_{c'} cyclical \ln GDPpc_{c't} + \delta_{c'} d_{c'} year \\
 & \times + \varphi_{c'} d_{c'} LS_{c't-1} + \alpha_{c'} d_{c'}] + \sum_p \delta_p d_p + \varepsilon_{ct},
 \end{aligned}
 \tag{2.3}$$

where d_p represents dummies for different preceding questions. The interaction coefficients indicate country-specific correlations of mean life satisfaction with *trend* $\ln GDPpc_{c't}$, *cyclical* $\ln GDPpc_{c't}$, *year*, and lagged mean life satisfaction, respectively. Analogously to Eq. 2.2, country-specific long-run correlations of mean life satisfaction with *trend* $\ln GDPpc_{c't}$, *cyclical* $\ln GDPpc_{c't}$, and *year* are given by $\beta_{c'}/(1 - \varphi_{c'})$, $\gamma_{c'}/(1 - \varphi_{c'})$, and $\delta_{c'}/(1 - \varphi_{c'})$, respectively. Here we have no downward Nickell bias in the country-specific estimates of $\varphi_{c'}$ as these estimates are only driven by the single cluster of observations for the specific country and Nickell bias only occurs with more than one cluster. Given the implosion of clustered standard errors, merely heteroscedasticity-robust or bootstrap standard errors can be used when serial correlation is small. We give preference to the type of standard errors which tend to be larger, as these seem to suffer less from finite-sample bias. For the bootstrap estimation of standard errors we chose to draw samples independently for each country, as this seems to be the appropriate method for the interaction coefficient estimates and since sampling across all countries broke down.

Because the number of country-specific observations may be too low for a number of countries, in some robustness regressions we replace the country-specific interactions of *cyclical* $\ln GDPpc_{c't}$ and/or $LS_{c't-1}$ with their main effects. To correct for the Nickell bias in the non-country-specific coefficient of $LS_{c't-1}$ we again apply the BCLSDV estimator of Bruno (2005a, b). However, now we do not use the bootstrap standard errors of the other coefficient estimates of Eq. 2.3 from this estimator, but calculate bootstrap-with-strata standard errors in a regression of Eq. 2.3 where the coefficient of $LS_{c't-1}$ has been fixated on the bias-corrected BCLSDV estimate. We follow this procedure because the required strata option for the interaction coefficient estimates (see above) is not available in the calculation of the bootstrap standard errors of the BCLSDV estimator. This specification is also our baseline for the Eastern European countries, where we only have 12 available observations per country. We then further run a robustness regression in which the interaction term for *year* has been replaced by its main effect. However, the latter regression is not very reliable as a test of the Easterlin Paradox for individual countries because country-specific correlations of mean life satisfaction with trend GDPpc are then only controlled for by a common time trend. Because of this concern, the much shorter time series, and the different levels and development of GDPpc of East as compared with Western European countries, we estimate the various variants of Eq. 2.3 for subgroups of Western and Eastern European countries separately.

A final concern is that country-specific estimates of φ_c and standard errors of all coefficient estimates are still biased for countries for which, after the addition of the interactions of one-year lagged life satisfaction, significant serial correlation in the error term continues to exist. In robustness checks and to diminish this serial correlation, we add country-specific interactions of one-year lagged *trend* $\ln GDPpc_{c,t-1}$ to the above variants of Eq. 3.3 for the Western European countries. For countries for which the estimate of the interaction coefficient β_c of *trend* $\ln GDPpc_{c,t}$ is significant and positive and the estimate of the interaction coefficient β_{-1c} of *trend* $\ln GDPpc_{c,t-1}$ is significant and negative, the latter coefficient estimate can be interpreted as modelling adaptation of life satisfaction to medium-term changes in GDPpc. For all countries, long-run correlations of mean life satisfaction with *trend* $\ln GDPpc_{c,t}$ are given by $(\beta_c + \beta_{-1c})/(1 - \varphi_c)$.¹²

2.4 Data and Descriptive Statistics

We use data from the nationally representative Eurobarometer surveys, ranging from 1973 to 2015. To elicit responses on life satisfaction, respondents are typically asked the following question: “*On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with the life you lead?*” with response options: “*Very satisfied (1), fairly satisfied (2), not very satisfied (3), not at all satisfied (4)*”. In most years more than one EB survey took place. In order to obtain country-year averages of life satisfaction, we take the mean of all responses in a given year and country.

For our estimations concerning groups of countries to test EPgl, we include all waves apart from those in which the set of response options or question format deviate from the format given above. We exclude these waves because such framing effects can have substantial effects on response patterns (Diener et al. 2013b). Henceforth, we will refer to this set of waves as “EB Standard”. Since we cannot use year-fixed effects in our country-specific estimations that test EPim (see Sect. 2.3.2), it is even more crucial for our purposes that country-year means of life satisfaction remain comparable over time. However, questions that immediately precede the life satisfaction question may impact answers to the life satisfaction question (see, e.g., Easterlin 2017). For our estimations in Sect. 2.5.2, we therefore select waves such that the number of distinct preceding questions across time is minimised, while continuing to have at least one EB wave available per year. This allows us to use dummies for preceding questions without them being collinear with the time trend of the reference country. We will call this set of waves “EB Restricted”. In total, “EB Standard” and “EB Restricted” cover 35 countries for the years 1973 to 2015. Of these we exclude Cyprus, Luxembourg, and Malta

¹²This expression follows from noting that in the long-run equilibrium current and past values of all variables are equal to each other.

because their populations do not exceed our threshold of one million inhabitants. We additionally exclude Albania, Iceland, Macedonia, Montenegro, Norway, and Serbia because they are observed for fewer than 10 years (see Sect. 2.3.1). This leaves us with 27 countries in total.

We use real and PPP-adjusted data on GDP per capita (GDPpc) for all estimations (in constant 2010 international \$). We primarily rely on data from the OECD (2017). Since not all European countries and years are covered by this data set, we supplement it with various other sources. We thus mainly use constant GDPpc data from the World Bank (2017) for Bulgaria, Croatia, and Romania. We also use this data for Ireland in 2015 because the OECD data for Ireland shows an implausible growth rate of 22% in that year.¹³ The OECD does not provide data on GDPpc for West and East Germany separately. For all years prior to 1991, we therefore use UNCTAD (2017) data for West Germany and data from Henske (2009) for East Germany. For years since 1991 we use data from Destatis (2017a). In cases where the OECD data does not extend far enough into the past, we use data from Penn World Tables (expenditure-side real GDP) (Feenstra et al. 2015). Finally, to minimize end-point problems in the estimation of the Hodrick-Prescott filter with $\lambda = 6.25$, we use GDPpc projections by the IMF (2016) for the years 2016–2021. As this series is expressed in current prices, we convert this series into constant prices using the inflation projections from the IMF for these years. For some robustness tests we include the unemployment and the inflation rate to our estimations. We source this data from the OECD and secondarily the World Bank. We additionally use data from the German Bundesagentur für Arbeit (2017) and DeStatis (2017b) to have distinct series for West and East Germany.

In our analyses we distinguish between Eastern and Western European countries because of their very different levels of GDPpc, the fact that most Eastern European countries went through an economic transition from communism to capitalism, and the much different observation windows we have available for each group.¹⁴ Mean levels of life satisfaction and GDPpc in the period 2004–2015 are clearly higher amongst Western than amongst Eastern European countries (3.07 vs. 2.68 and \$38,017 vs. \$21,762, respectively). However, the subset of Southern European countries (Spain, Greece, Italy, Portugal) falls short of that tendency and has a mean LS (= 2.61) and a mean GDPpc (= \$30,386) closer to the Eastern European countries.

¹³This extreme growth was largely driven by an accounting trick of a number of multinational companies (Inman, 2016). Therefore, this change in GDPpc is unlikely to have had an impact on living standards. The World Bank data set records a growth of only 7%.

¹⁴Mean T is 35 for the Western European countries and 13 for the Eastern European countries.

2.5 Results

2.5.1 Results for Groups of Countries

In this section we present the results for groups of countries¹⁵ when using the HP filter with $\lambda = \infty$, which is suitable for testing the long-term version EPgl of the paradox. However, the estimation period is too short for the Eastern European countries to actually allow for tests of the long-term version of the paradox. When we present results for this group alone, we therefore label these results as results about the medium-term variant EPgm of the paradox.¹⁶

We begin with presenting the estimation results for Eq. 2.1. First, we estimate this equation for the group of all 27 European countries selected in Sect. 2.4. Figure 2.2 (left) presents a scatterplot for this country group in which residuals from regressing Eq. 2.1 without trend $\ln GDP_{pc}$ are plotted against residuals from regressing $trend \ln GDP_{pc_{ct}}$ on the country and year dummies. The linear regression fit of this cloud of data points is rising, but only slightly and the slope as given by the coefficient estimate 0.10 of trend $\ln GDP_{pc}$ in column (1) of Table 2.1, turns out to be strongly insignificant. However, a striking feature in the scatter diagram in Fig. 2.2 (left) is that the data points for Ireland (as indicated by red dots) are outliers with extremely low and high values of the residual of trend $\ln GDP_{pc}$ (which represents the double difference of trend $\ln GDP_{pc}$ with respect to its time and country means). This raises the question on the impact of these outliers, which becomes visible in Fig. 2.2 (right) where we drop Ireland. This leads to a remarkably strong rise in the slope of the regression line, which is reflected in a marginally ($p = 0.10$) significant¹⁷ and much larger coefficient estimate of 0.62 for trend $\ln GDP_{pc}$ in column (2) of Table 2.1. The result of column (1) was hence largely driven by the outlier Ireland. Therefore, we drop Ireland from the subsequent regressions in this section.

Thus, for our sample of 26 European countries without Ireland the long-term variant of the Easterlin Paradox is marginally rejected. However, Proto and Rustichini (2013) found a non-monotonic relation between GDP_{pc} and life satisfaction, which is significantly positive for poorer countries/regions, but insignificant or significantly negative for richer countries/regions. This suggests that our rejection of the paradox may be driven by the subgroup of the 13 less developed Eastern European countries with their lower mean GDP_{pc} . Therefore, in column (3) we drop these countries from the regression, leaving us with 13 mainly Western

¹⁵Here we use “EB-Standard” data.

¹⁶In Kaiser and Vendrik (2018) we extensively discuss the results of our tests when setting $\lambda = 6.25$ which more closely corresponds to tests of EPgm. These results are very similar to those for $\lambda = \infty$, the main difference being slightly more reliably significant coefficients for sets of Eastern European countries.

¹⁷In this study we call an estimate (strongly) significant when its p-value in a two-tailed t test is below 0.05 (0.01), and marginally significant when its p value in a two- or one-tailed t test is higher than 0.05, but lower than 0.10. In the latter case we mention the p value in parentheses, which refers to a two-tailed t test unless it is explicitly stated that it refers to a one-tailed t test.

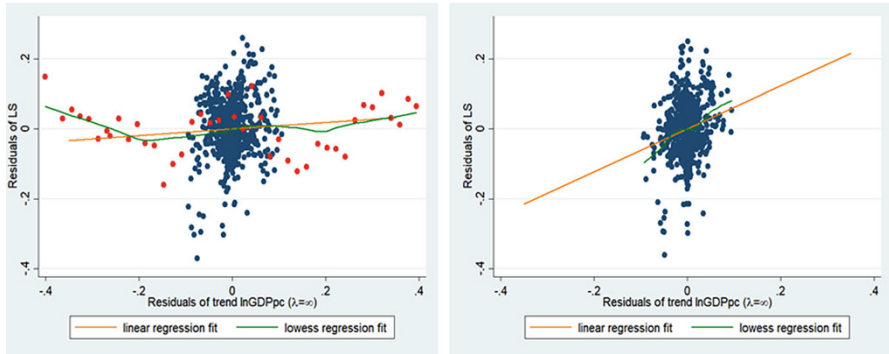


Fig. 2.2 Scatterplots of residuals of regression of Eq. 2.1 for the long term without *trend ln GDPpc_{ct}* against residuals of regression of *trend ln GDPpc_{ct}* on year and country dummies for all countries with Ireland marked in red (left) and when omitting Ireland (right)

Table 2.1 Baseline Results for Eq. 2.1 for the Long Term

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All, no IE	EU-13	EU-9	SE	EE
Trend LnGDPpc	0.095 (0.125)	0.616* (0.356)	0.783 (0.762)	0.014 (0.556)	1.425 (1.250)	0.477* (0.228)
Cyclical LnGDPpc	0.765*** (0.137)	0.782*** (0.150)	0.872*** (0.176)	0.159 (0.324)	0.527 (0.413)	0.832** (0.275)
R-squared	0.943	0.945	0.931	0.934	0.885	0.938
Number of Countries	27	26	13	9	4	13
Number of Observations	666	624	454	315	139	170

Note: Estimated with Stata’s ‘regress’ command. Country and year dummies included. Country-clustered standard errors in parentheses. Significance levels: for two-tailed t test: +: $p < 0.20$, *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$; for one-tailed t test: +: $p < 0.10$, *: $p < 0.05$, **: $p < 0.025$, ***: $p < 0.005$

European countries without Ireland (EU-13). For this EU-13 the coefficient estimate is insignificant, but surprisingly it is even larger in size than for the total group of 26 European countries without Ireland (0.78 vs. 0.62). The large standard error (0.76) of this estimate may be due to strong heterogeneity in the effects of differences in long-term economic growth on life satisfaction across different (groups of) EU-13 countries. Given the strong sensitiveness of mean life satisfaction in the Southern European (SE) countries Greece, Italy, Spain, and Portugal to the recent Euro crisis and their lower mean GDPpc, the large size of the coefficient for the EU-13 may be driven by this group of four SE countries. This is also suggested by the scatter diagram for the EU-13 in Fig. 2.3 (left) in which the data points for the four SE countries are indicated by red dots. Dropping these data points from the regression, we obtain Fig. 2.3 (right) with a slope that is virtually flat. This is reflected by the strongly insignificant and very small coefficient 0.01 of *trend ln GDPpc* in the regression for the nine remaining Western and Northern European countries in

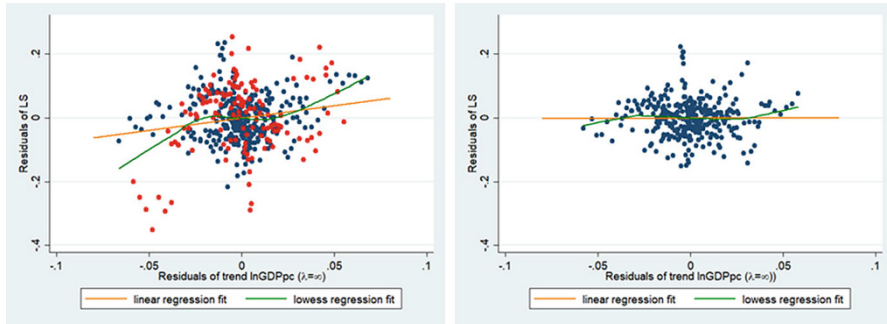


Fig. 2.3 Scatterplots of residuals of regression of Eq. 2.1 for the long term without *trend* $\ln GDPpc_{ct}$ against residuals of regression of *trend* $\ln GDPpc_{ct}$ on year and country dummies for the group of EU-13 countries with Southern European countries marked in red (left) and when omitting Southern European countries (right)

column (4) of Table 2.1. Thus, in this subgroup of highly developed countries (EU-9) a higher long-term growth of GDP per capita was not associated with a more positive change in average life satisfaction. So, the group of these nine Western and Northern European countries clearly satisfies the long-term-variant EPgl of the Easterlin Paradox.¹⁸

Figure 2.3 also suggests that when we restrict the regression to the four Southern European countries, the coefficient of *trend* $\ln GDPpc$ will be significant, positive, and large. However, column (5) of Table 2.1 shows that although this coefficient is indeed large and positive, it is not statistically significant ($p = 0.34$). The large standard error that drives this (1.25) seems to be due to the coefficient of *trend* $\ln GDPpc$ being identified by only three¹⁹ differences in country-specific observations for the average growth rate of GDPpc. Finally, column (6) shows that for the group of 13 Eastern European (EE) countries the coefficient of *trend* $\ln GDPpc$ is marginally ($p = 0.06$) significant, positive, and large. Thus, in this group of countries (with their short estimation period) a higher *medium*-term growth of GDP per capita was associated with a more positive change in average life satisfaction. This implies a marginal rejection of the medium-term-variant EPgm of the Easterlin Paradox for this group of countries. Because these countries had a lower mean GDP per capita than the Western and Northern European countries, this is in line with the significantly positive relation between GDPpc and life satisfaction for poorer countries and European regions as found by Proto and Rustichini (2013).

However, especially the last result may be biased due to the small number (13) of country clusters. Clustered standard errors of the parameter estimates then tend to be

¹⁸The coefficient of cyclical $\ln GDPpc$ is insignificant for this group of countries as well, so even cyclical fluctuations in GDPpc were not associated with changes in average life satisfaction in these countries.

¹⁹Given this extremely low number of effective observations, we also run a robustness regression of Eq. 2.1 for this group of countries without cyclical $\ln GDPpc$. This yields a coefficient of 1.66 with a standard error of 1.11 ($p = 0.23$), which is close to marginally significant in a one-tailed t test.

underestimated (see Sect. 2.3.1). In our case this downward bias in the standard errors is likely to be especially strong as tests for first and second-order serial correlation of the error term (see Wooldridge 2003, pp. 399–402) in Eq. 2.1 show strong positive first-order serial correlation (in the order of 0.50–0.70).²⁰ We therefore reduce this first-order serial correlation by adding one-year lagged mean life satisfaction to Eq. 2.1, yielding Eq. 2.2. Table 2.2 presents estimation results for Eq. 2.2 for the same groups of countries as those distinguished in Table 2.1 for Eq. 2.1. Now, the long-run effects of trend $\ln\text{GDPpc}$ and cyclical $\ln\text{GDPpc}$ are the relevant estimates that can be compared with the coefficient estimates in Table 2.1.²¹

For the total sample of all 27 European countries, column (1) of Table 2.2 shows a strongly significant and large bias-corrected coefficient of 0.81 for lagged life satisfaction, which implies a strong persistence of mean life satisfaction. This persistence does not only reflect a possible direct effect of lagged life satisfaction on current life satisfaction, but also reinforcement of the effects of trend and cyclical $\ln\text{GDPpc}$, and those of omitted variables (*e.g.*, health)²² on life satisfaction. A serial correlation test for Eq. 2.2 shows that, as a result of the addition of the lagged life satisfaction term, all first-order and second-order serial correlation is eliminated (*i.e.* becomes insignificant) except for marginally ($p = 0.08$) significant, negative, and small (-0.05) first-order serial correlation for the EU-13 countries. Hence, the bootstrap standard errors of the parameter estimates calculated by the BCLSDV estimator in Stata (see Sect. 2.3.1) are more reliable than those obtained from the estimation of Eq. 2.1. The coefficient estimates for trend $\ln\text{GDPpc}$ and cyclical $\ln\text{GDPpc}$ in column (1) of Table 2.2 can be interpreted as short-run effects of these variables (see footnote 8 in Sect. 2.3.1), which is insignificant for trend $\ln\text{GDPpc}$ but significant for cyclical $\ln\text{GDPpc}$. The reinforcement of these effects results in much larger long-run (LR) effects, which are nevertheless insignificant for trend $\ln\text{GDPpc}$, but significant for cyclical $\ln\text{GDPpc}$. The size and standard error of the long-run effect of trend $\ln\text{GDPpc}$ in Table 2.2 are both about twice as large as their equivalents in Table 2.1. In general, the much larger standard errors of the long-run effects in Table 2.2 not only reflect the downward bias of the standard errors estimates in Table 2.1 due to the low number of clusters (13), but also the partial control for serially correlated and time-varying omitted variables via the added lagged life satisfaction term. Therefore, the estimates for Eq. 2.2 in Table 2.2 seem more reliable than those for Eq. 2.1 in Table 2.1.

For the sample of 26 European countries without Ireland, column (2) of Table 2.2 shows a long-run effect of trend $\ln\text{GDPpc}$, which is again somewhat larger than the

²⁰The second-order serial correlation is significant for most groups, but relatively small (at most 0.16). We do not explicitly correct for that in the following.

²¹In time-series analysis a static equation like Eq. 2.1 is interpreted as the long-run-equilibrium equation that corresponds to a dynamic equation like Eq. 2.2 (*cf.* Vendrik, 2013).

²²In Eq. 2.2 such omitted variables work via changes in the error term which in the next year are reinforced via the lagged life satisfaction term. This reinforcement also picks up the effects of positive serial correlation in time-varying omitted variables. See Vendrik (2013) for a deeper dynamic analysis.

Table 2.2 Baseline Results for Eq. 2.2 for the Long Term

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	All, no IE	EU-13	EU-9	SE	EE	EE-11
L-Life Satisfaction	0.805*** (0.033)	0.806*** (0.035)	0.850*** (0.037)	0.716*** (0.049)	0.795*** (0.091)	0.668*** (0.080)	0.527*** (0.102)
Trend LnGDPpc	0.039 (0.045)	0.170+ (0.111)	0.250+ (0.184)	-0.092 (0.173)	0.484 (0.511)	0.097 (0.171)	0.225+ (0.162)
Cyclical LnGDPpc	0.146** (0.060)	0.170** (0.068)	0.079 (0.098)	0.053 (0.111)	0.049 (0.205)	0.595** (0.211)	0.614** (0.216)
LR effect of Trend LnGDPpc	0.198 (0.235)	0.877+ (0.568)	1.665+ (1.199)	-0.323 (0.618)	2.363 (2.344)	0.291 (0.504)	0.476+ (0.319)
LR effect of Cyclical LnGDPpc	0.748** (0.310)	0.875** (0.364)	0.528 (0.659)	0.188 (0.386)	0.238 (1.012)	1.790** (0.802)	1.297** (0.556)
Number of Countries	27	26	13	9	4	13	11
Number of Observations	631	591	434	300	134	157	121

Note: Estimated with 'xtlsdvc' command, written by Bruno (2005b). Country and year dummies included. Bootstrapped (400 replications) standard errors in parentheses. Significance levels: for two-tailed t test: +: p < 0.20, *: p < 0.10, **: p < 0.05, ***: p < 0.01; for one-tailed t test: +: p < 0.10, *: p < 0.05, **: p < 0.025, ***: p < 0.005. R-squareds are not reported by the 'xtlsdvc' command

coefficient in column (2) of Table 2.1, but which is now only marginally ($p = 0.06$) significant in a one-tailed test. Hence, the marginal rejection of the Easterlin Paradox for this group of countries in Table 2.1 is now ambiguous in Table 2.2. We also find such weak evidence for a rejection of EPgl for the group of 13 mainly Western European countries without Ireland (EU-13) in column (3) of Table 2.2 (one-tailed $p = 0.08$), which is in contrast with the insignificant result in Table 2.1 and which is due to the much larger size of the long-run effect of trend $\ln\text{GDPpc}$. However, when we drop the four Southern European countries in column (4) of Table 2.2, the long-run effect of trend $\ln\text{GDPpc}$ is again strongly insignificant and even negative, implying a clear confirmation of the long-term variant EPgl of the Easterlin Paradox for this group of nine highly-developed Western and Northern-European countries (EU-9).

For the group of four Southern European countries column (5) of Table 2.2 shows an insignificant long-run effect of trend $\ln\text{GDPpc}$ as well, which is consistent with the result in column (5) of Table 2.1. However, for the group of 13 Eastern European countries the strongly insignificant long-run effect of trend $\ln\text{GDPpc}$ in column (6) is inconsistent with the marginally significant effect of trend $\ln\text{GDPpc}$ in column (6) of Table 2.1. This is due to a much lower size as well as much larger standard error of the estimate in Table 2.2. Especially the much smaller size of the latter estimate is puzzling and may be driven by one or more outlier countries. Such outliers may be Turkey because it is not an ex-communist country like the other EE countries, and East Germany because it has been integrated with highly developed West Germany since 1990 and has a much longer time series for life satisfaction in our dataset than the other EE countries (since 1990 vs. 2004). When we drop these two countries from the group of EE countries, the long-run effect of trend $\ln\text{GDPpc}$ as shown in column (7) of Table 2.2 becomes much larger and marginally significant in a one-tailed t test ($p = 0.07$). In addition, an estimation of Eq. 2.1 for the remaining subgroup of 11 EE countries yields a significant coefficient of trend $\ln\text{GDPpc}$ (with size 0.568). We thus obtain weak evidence of a rejection of the medium-term variant EPgm of the Easterlin Paradox for this group of 11 Eastern European countries (EE-11).

We performed a number of checks to assess how robust these results are to dropping or adding relevant control variables (particularly the unemployment and inflation rate) and to restricting the estimation period. See Kaiser and Vendrik (2018) for an extended discussion. Our estimates are qualitatively robust to most of these checks. However, the weak rejection of the paradox for the group of EE-11 countries turns out to be driven by an associated decline in the inflation rate and its positive effect. Since long-term changes in the inflation rate should be considered a good control in the sense of Angrist and Pischke (2009),²³ the rejection of the Easterlin paradox in our baseline regressions for this set of countries should be viewed as possibly spurious.

²³For example, economic reforms in a country and globalization may lead to both higher long-term economic growth and lower inflation.

In sum, for the group of Northern and Western European countries (EU-9) we have obtained a clear and unambiguous affirmation of the long-term version of the Easterlin Paradox. Moreover, we have obtained weak and non-robust evidence for a rejection of EPgm for the set of Eastern European countries without Turkey and East Germany. As explained earlier, the estimation period for the 11 Eastern European countries is only 11 years (2004–2015), which includes short-term, but not medium-term cycles of GDPpc that tend to last between roughly 15 and 30 years (see Sect. 2.1). Hence in the case of the EE-11 countries, the HP filter with $\lambda = \infty$ may only provide us with tests of the medium-term version of the paradox.²⁴ For a genuine test of the more appropriate long-term variant EPgl for this country group longer time series are needed.

2.5.2 Results for Individual Countries

In this section we present the results of testing the medium-term variant EPim²⁵ of the Easterlin Paradox for individual countries. Here we use “EB-restricted” data. Our discussion starts with the groups of Western European countries and then moves on to the group of Eastern European countries.

We estimate Eq. 2.3²⁶ for the group of 14 WE countries plus East Germany (WE⁺).²⁷ Column (1) of Table 2.3 presents the long-run effects of trend lnGDPpc and cyclical lnGDPpc for all individual WE⁺ countries. The long-run effect of trend lnGDPpc is (marginally) significant and positive for Greece (1.14), Ireland (0.26; $p = 0.07$), Italy (1.49), and Spain (0.86) in a two-tailed t test, and marginally significant and positive for Portugal (0.58) only in a one-tailed t test ($p = 0.10$).

²⁴When using an HP filter with $\lambda = 6.25$ we obtain somewhat larger and slightly more significant estimates for this group of countries. Our rejection of EPgm is nevertheless non-robust in this case, too. See Kaiser and Vendrik (2018) for an extended discussion.

²⁵In Kaiser and Vendrik (2018) we also present tests of EPI0 (non-positive long-term time trend in happiness) for Western European countries (plus East Germany). There we show that EPI0 is robustly confirmed for Austria, Belgium, Greece, the Netherlands, Portugal, and West Germany, but robustly violated for Denmark, Finland, France, Great Britain, Italy, and Sweden. For the other countries the results are not robust. In the case of the Eastern European countries the limited observation period does not allow for reliable tests of a long-term non-positive time trend in happiness. Among these countries, we obtain (marginally) significant and positive medium-term time trends for Bulgaria, Hungary, Lithuania, Latvia, Poland, and Romania.

²⁶These regressions include only two significant dummies for the preceding questions. The first dummy controls for a question on which political party the respondent supports (1979 and 1983 waves). The second dummy concerns a question on the share of friends appreciating talk about politics (1998 wave). These dummies have long-run effects averaged across countries of -0.06 and -0.07 , respectively.

²⁷We include East Germany here because East Germany has an estimation period that is much closer to the Western European country group. See Kaiser and Vendrik (2018) for more explanation.

Table 2.3 Baseline Results for Eq. 2.3 for WE⁺ Countries and the Medium Term

		(1)		(2)		(3)	
		Standard		Main effect of Cyclical LnGDPpc		Main effects of Cycl. LnGDPpc and L.Cycl. LnGDPpc, interactions of L.Trend LnGDPpc	
		Effect	SE	Effect	SE	Effect	SE
Austria	LR effect of Trend LnGDPpc	-1.362***	(0.490)	-1.427***	(0.468)	-1.450***	(0.515)
(T = 20)	LR effect of Year	0.013**	(0.006)	0.014**	(0.006)	0.015+	(0.009)
Belgium	LR effect of Trend LnGDPpc	0.442	(1.339)	-0.016	(1.029)	-0.146	(1.307)
(T = 40)	LR effect of Year	-0.007	(0.022)	0.000	(0.016)	0.003	(0.024)
Denmark	LR effect of Trend LnGDPpc	-0.120	(0.211)	-0.174	(0.196)	-0.240+	(0.175)
(T = 40)	LR effect of Year	0.006*	(0.003)	0.007**	(0.003)	0.009***	(0.002)
East Germany	LR effect of Trend LnGDPpc	-2.373**	(1.062)	-2.475***	(0.951)	-2.357**	(1.108)
(T = 25)	LR effect of Year	0.057***	(0.020)	0.058***	(0.019)	0.056***	(0.021)
Finland	LR effect of Trend LnGDPpc	0.403	(0.378)	0.311	(0.325)	0.285	(0.348)
(T = 20)	LR effect of Year	0.001	(0.004)	0.003	(0.004)	-0.000	(0.011)
France	LR effect of Trend LnGDPpc	-0.154	(0.396)	-0.153	(0.364)	-0.264	(0.368)
(T = 40)	LR effect of Year	0.008+	(0.005)	0.008+	(0.005)	0.011*	(0.006)
Great Britain	LR effect of Trend LnGDPpc	-0.585***	(0.157)	-0.580***	(0.154)	-0.547***	(0.156)
(T = 40)	LR effect of Year	0.016***	(0.003)	0.015***	(0.003)	0.015***	(0.003)
Greece	LR effect of Trend LnGDPpc	1.141***	(0.349)	1.154***	(0.351)	0.853***	(0.221)
(T = 34)	LR effect of Year	-0.030***	(0.006)	-0.030***	(0.006)	-0.023***	(0.003)

(continued)

Table 2.3 (continued)

		(1)		(2)		(3)	
		Standard		Main effect of Cyclical LnGDPpc		Main effects of Cycl. LnGDPpc and L.Cycl. LnGDPpc, interactions of L.Trend LnGDPpc	
		Effect	SE	Effect	SE	Effect	SE
Ireland	LR effect of Trend LnGDPpc	0.255*	(0.139)	0.310**	(0.152)	0.269 ⁺	(0.166)
(T = 40)	LR effect of year	-0.007 ⁺	(0.005)	-0.010 ⁺	(0.006)	-0.008	(0.006)
Italy	LR effect of Trend LnGDPpc	1.490***	(0.141)	1.479***	(0.133)	1.455***	(0.171)
(T = 40)	LR effect of Year	-0.018***	(0.002)	-0.018***	(0.002)	-0.017***	(0.004)
Netherlands	LR effect of Trend lnGDPpc	-0.445 ⁺	(0.338)	-0.351	(0.367)	-0.302	(0.383)
(T = 40)	LR effect of year	0.009 ⁺	(0.006)	0.007	(0.006)	0.007	(0.007)
Portugal	LR effect of Trend LnGDPpc	0.577 ⁺	(0.449)	0.556	(0.512)	0.727 ⁺	(0.499)
(T = 29)	LR effect of Year	-0.023**	(0.009)	-0.023**	(0.010)	-0.011	(0.025)
Spain	LR effect of Trend LnGDPpc	0.863**	(0.420)	0.996**	(0.428)	0.805**	(0.379)
(T = 29)	LR effect of Year	-0.015*	(0.007)	-0.018**	(0.008)	-0.006	(0.008)
Sweden	LR effect of Trend LnGDPpc	0.189	(0.439)	0.137	(0.383)	0.065	(0.373)
(T = 20)	LR effect of Year	0.004	(0.006)	0.005	(0.005)	0.004	(0.004)
West Germany	LR effect of Trend LnGDPpc	-0.861	(0.932)	-0.851	(0.901)	-0.438	(0.697)
(T = 40)	LR effect of year	0.014	(0.013)	0.014	(0.013)	0.011	(0.010)
Number of Observations:		497		497		497	

Note: Estimated with Stata's 'regress' command, controlling for country-specific or main effects of cyclical lnGDPpc, PQDs, and country dummies. T now indicates the number of effective observations that identifies the interaction coefficient for a particular country. Strata bootstrapped (400 replications) standard errors in parentheses. Significance levels: for two-tailed t test: +: $p < 0.20$, *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$; for one-tailed t test: +: $p < 0.10$, *: $p < 0.05$, **: $p < 0.025$, ***: $p < 0.005$

Interestingly, these are precisely the countries that suffered most from the recent Eurocrisis. Thus, for these countries the medium-term variant EPim of the Easterlin Paradox for individual countries is violated. Note that the positive long-run effects of trend $\ln \text{GDPpc}$ for these countries go together with (marginally) significant negative time trends. For the other countries the long-run effect of trend $\ln \text{GDPpc}$ is either insignificant or (marginally) significantly negative (for Austria, East Germany, Great Britain, the Netherlands, one-tailed $p = 0.09$). Thus, for these individual countries the medium-term variant EPim of the Easterlin Paradox is confirmed.

For countries with less than 40 observations (Austria, East Germany, Finland, Greece, Portugal, Spain, and Sweden) the number of observations may be too low to lead to stable, and hence reliable, estimates of the four interaction terms for each country in Eq. 2.3 (see Sect. 2.3.2). In column (2) of Table 2.3 we therefore replace the interaction terms for cyclical $\ln \text{GDPpc}$ with its main effect. This yields qualitatively robust results except for insignificant effects of trend $\ln \text{GDPpc}$ for Portugal and the Netherlands.

We observe (marginally) significant first-order serial correlation in the error terms of the regressions in columns (1) (Greece, Spain, West Germany) and (2) (additionally Denmark and Ireland). As explained above, this may result in a downward bias to our standard errors. To reduce this serial correlation, we add country-specific interaction terms for one-year-lagged $\text{trend} \ln \text{GDPpc}_{ct-1}$ and a main effect of one-year-lagged $\text{cyclical} \ln \text{GDPpc}_{ct-1}$ to the regression in column (2). This yields the estimates in column (3) of Table 2.3. These estimates are similar to those in column (2), with again a rejection of the Easterlin paradox for Greece, Ireland, Italy, and Spain, but now also marginally for Portugal. We further observe (marginally) significant partial adaptation to medium-term changes in GDPpc for Greece, Ireland, Portugal, and Spain, as well as full adaptation for West Germany.

The estimates of column (3) may again not be robust due to too few observations for some countries.²⁸ We therefore conducted a robustness test in which we used the main effect of lagged life satisfaction instead of country-specific interactions. Our results are robust to this test, except that we no longer find a significant rejection for Ireland (this is likely driven by Ireland's strong deviation from the uniform coefficient of lagged life satisfaction). We also tried dropping the interactions of lagged $\text{trend} \ln \text{GDPpc}_{ct-1}$ and the main effect of $\text{cyclical} \ln \text{GDPpc}_{ct-1}$ from the previous regression, which showed no qualitative change in the estimates. Finally, we also checked for robustness against the impact of the recent Great Recession by restricting the estimation period to the period before 2008, which again yielded rejections of EPim for the same set of countries.

Overall, we conclude that the medium-term variant EPim of the Easterlin Paradox is robustly violated for Greece, Ireland, Italy, and Spain, but robustly confirmed for the nine Western and Northern European countries (EU-9). The latter confirmation is consistent with the confirmation of EPgl for the EU-9 as a group of countries.

We now turn to Eastern Europe. For these countries, column (1) of Table 2.4 shows the long-run effects of trend $\ln \text{GDPpc}$ and year from a regression of Eq. 2.3,

²⁸See Kaiser and Vendrik (2018) for a more extensive discussion of these robustness results.

Table 2.4 Baseline Results for Eq. 2.3 for Eastern European Countries and the Medium Term

		(1)		(2)	
		Standard		Main effect of Year	
		Effect	SE	Effect	SE
Bulgaria	LR effect of Trend LnGDPpc	1.516	(2.155)	0.957**	(0.412)
(T = 11)	LR effect of Year	-0.018	(0.051)	-0.003	(0.006)
Croatia	LR effect of Trend LnGDPpc	1.011	(2.108)	1.335	(1.708)
(T = 11)	LR effect of Year	0.004	(0.012)	-0.003	(0.006)
Czech Republic	LR effect of Trend LnGDPpc	-1.253	(4.130)	0.421	(0.993)
(T = 11)	LR effect of Year	0.020	(0.035)	-0.003	(0.006)
Estonia	LR effect of Trend LnGDPpc	1.787	(1.740)	0.828	(0.684)
(T = 11)	LR effect of Year	-0.018	(0.029)	-0.003	(0.006)
Hungary	LR effect of Trend LnGDPpc	3.527	(3.997)	3.530	(2.853)
(T = 11)	LR effect of Year	-0.000	(0.027)	-0.003	(0.006)
Latvia	LR effect of Trend LnGDPpc	-0.855	(1.531)	1.045*	(0.582)
(T = 11)	LR effect of Year	0.042 ⁺	(0.032)	-0.003	(0.006)
Lithuania	LR effect of Trend LnGDPpc	1.540	(1.931)	0.895***	(0.324)
(T = 11)	LR effect of Year	-0.024	(0.063)	-0.003	(0.006)
Poland	LR effect of Trend LnGDPpc	1.037	(2.987)	0.434*	(0.244)
(T = 11)	LR effect of Year	-0.024	(0.101)	-0.003	(0.006)
Romania	LR effect of Trend LnGDPpc	0.246	(3.138)	0.922 ⁺	(0.691)
(T = 11)	LR effect of Year	0.019	(0.080)	-0.003	(0.006)
Slovakia	LR effect of Trend LnGDPpc	1.124	(4.495)	0.220	(0.315)
(T = 11)	LR effect of Year	-0.033	(0.116)	-0.003	(0.006)
Slovenia	LR effect of Trend LnGDPpc	-0.170	(1.194)	-0.483	(0.979)
(T = 11)	LR effect of Year	-0.009 ⁺	(0.006)	-0.003	(0.006)
Turkey	LR effect of Trend LnGDPpc	2.823	(7.908)	-0.595	(0.760)
(T = 11)	LR effect of Year	-0.086	(0.201)	-0.003	(0.006)
Number of Observations:		132		132	

Note: Estimated with Stata's 'regress' and Bruno's 'xtlsdvc' command, controlling for main effect of cyclical lnGDPpc and country dummies. T now indicates the number of effective observations that identifies the interaction coefficient for a particular country. Strata bootstrapped (400 replications) standard errors in parentheses. Significance levels: for two-tailed t test: +: $p < 0.20$, *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$; for one-tailed t test: +: $p < 0.10$, *: $p < 0.05$, **: $p < 0.025$, ***: $p < 0.005$

using main effects instead of interactions for both cyclical lnGDPpc and lagged life satisfaction. In this regression all dummies for different preceding questions have been dropped because they were jointly insignificant. Strikingly, the long-run effects of trend lnGDPpc are (strongly) insignificant for all EE countries. This is unexpected in view of the marginally significant long-run effect of trend lnGDPpc for the group of EE countries without East Germany and Turkey in column (7) of Table 2.2. Our result may be due to too little variation in medium-term economic growth rates of the EE countries over the short estimation period of 11 years. Therefore, column (2) shows the long-run effects of trend lnGDPpc while controlling for a common time trend instead of country-specific time trends. Now for five out of 12 countries these

long-run effects are (marginally) significant and positive, namely for Bulgaria, Lithuania, Latvia ($p = 0.08$), Poland ($p = 0.06$), and Romania (one-tailed $p = 0.09$). However, the control for a common instead of country-specific time trends makes the results of this test dubious for Latvia and Romania as the marginally significant and positive long-run effects of trend $\ln\text{GDPpc}$ for these countries in column (2) of Table 2.4 apparently pick up the positive country-specific long-run time trends found in column (1) (see also the discussion in Sect. 2.1). For the other three countries, i.e. Bulgaria, Lithuania, and Poland, the country-specific long-run time trends in column (1) are more negative than the common long-run time trend, and hence cannot account for the (marginally) significantly positive long-run effects of (mainly positive) changes in trend $\ln\text{GDPpc}$ in column (2). Thus, we conclude that the medium-term variant EPim of the Easterlin Paradox for individual EE countries is only rejected for Bulgaria, Lithuania, and Poland. For a reliable test of whether this variant of the Easterlin Paradox is also rejected for other EE countries, longer time series than those currently available are needed.

2.6 Conclusions

Our starting point was the argument that reliable tests of the Easterlin Paradox should control for the possibility of spuriousness of the correlation between average happiness and long-term economic growth by means of common or country-specific time trends. This led to a distinction between five variants of the paradox along the two dimensions of groups of countries *versus* individual countries and the long *versus* medium-term. We further argued that the long-term version of the paradox for groups of countries and the medium-term version for individual countries are the most appropriate testable versions of the paradox. We found a clear and robust confirmation of the long-term version (as well as the medium-term version) of the paradox for a group of nine Western and Northern European countries. Moreover, we obtained a non-robust rejection of the medium-term variant of the paradox for a set of 11 Eastern European countries. On the level of individual countries, the Easterlin Paradox for the medium term turned out to clearly hold for the nine Western and Northern European countries, but to be consistently rejected for Spain, Greece, Ireland, and Italy. Thus, in the latter four as opposed to the former nine countries, economic growth was positively associated with the development of life satisfaction in the medium term. In the case of the individual Eastern European countries, this was also found to hold for Bulgaria, Lithuania, and Poland, but for the other EE countries the test results are unreliable, partially due to the limited length of the time series (only 11 years).²⁹ Note that our results for individual European

²⁹Note that this limitation also holds for other international data sets for life satisfaction and happiness. Examples are the World Values Survey (up to six observations) and the Gallup World Poll (13 observations for 2005–2017).

countries in the medium term are largely consistent with our findings for the groups of countries to which the individual countries belong.

We thus give a nuanced picture of the empirical validity of the Easterlin Paradox. On the one hand, we show that the paradox is confirmed for Western and Northern European countries, both as a group and individually. On the other hand, our results imply a rejection of the medium-term version of the paradox for three individual Southern European countries and Ireland, and at least suggest a rejection of the paradox for Eastern European countries in the medium term. Because the Western and Northern European countries have a high per capita GDP as compared to that of Southern and Eastern European countries and (initially) Ireland, our results are in line with those of Proto and Rustichini (2013), who find a non-monotonic relation between per capita GDP and life satisfaction over time. Thus, on the one hand and in line with Proto and Rustichini and Veenhoven and Vergunst (2014), but contrary to Easterlin (2017), we have obtained evidence that suggests that, at least in the (less appropriate) medium term, the Easterlin Paradox does not hold for lower-income European countries. On the other hand, and in line with Proto and Rustichini and Easterlin (2017), but contrary to Sacks et al. (2013) and Veenhoven and Vergunst (2014), we have found evidence that strongly suggests that, over the last 40 years, economic growth did not raise average life satisfaction in the long and medium term in higher-income European countries. Thus, in response to the title of Easterlin's 2016 paper: although the "blissful paradise" of universal validity of the paradox may have been lost, the paradox itself is not!

Acknowledgements We thank Erik de Regt for econometric advice, Marta Golin for advice on using Stata, Guy Mayraz and Ruut Veenhoven for sending us their data sets, Daniel Sacks and Justin Wolfers for information about their research, and John Barkley Rosser Jr., Robert Dur, Richard Easterlin, Ari Kapteyn, Maurizio Pugno, Mariano Rojas, Francesco Sarracino, and other participants of the International Conference on Policies for Happiness and Health in Siena, March 19–21, 2018, and the USC Happiness Conference at the occasion of the retirement of Richard Easterlin in Los Angeles, April 6, 2018, for helpful comments.

References

- Angeles, L. (2011). A closer look at the Easterlin Paradox. *The Journal of Socio-Economics*, 40, 67–73.
- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press.
- Bartolini, S., & Sarracino, F. (2014). Happy for how long? How social capital and economic growth relate to happiness over time. *Ecological Economics*, 108(Suppl. C), 242–256.
- Bartolini, S., Bilancini, E., & Pugno, M. (2013a). Did the decline in social connections depress Americans' happiness? *Social Indicators Research*, 110, 1033–1059.
- Bartolini, S., Bilancini, E., & Sarracino, F. (2013b). Predicting the trend of well-being in Germany: How much do comparisons, adaptation and sociability matter? *Social Indicators Research*, 114 (2), 169–191.
- Beja, E. L. (2014). Income growth and happiness: reassessment of the Easterlin Paradox. *International Review of Economics*, 61(4), 329–346.

- Bruno, G. S. (2005a). Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models. *Economics Letters*, 87(3), 361–366.
- Bruno, G. S. F. (2005b). *XTLSDVC: Stata module to estimate bias corrected LSDV dynamic panel data models*. *Statistical Software Components S450101*, revised 08 Sep. 2005. Boston: Boston College Department of Economics.
- Bundesagentur für Arbeit. (2017). *Arbeitslose und Unterbeschäftigung*. Retrieved November 10, 2017, from https://statistik.arbeitsagentur.de/nn_217700/Statischer-Content/Rubriken/Arbeitslose-und-gemeldetes-Stellenangebot/Arbeitslose/Arbeitslosigkeit-in-Deutschland-seit-1950-Monats-Jahreszahlen.html
- Burns, A. F., & Mitchell, W. C. (1946). *Measuring business cycles*. New York: National Bureau of Economic Research.
- Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *The Journal of Human Resources*, 50(2), 317–372.
- Clark, A. E. (2011). Income and happiness: Getting the debate straight. *Applied Research in Quality of Life*, 6, 253–263.
- Clark, A. E., Flèche, S., & Senik, C. (2014). The great happiness moderation. In A. E. Clark & C. Senik (Eds.), *Happiness and economic growth: Lessons from developing countries* (pp. 32–139). Oxford: Oxford University Press.
- 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.
- Destatis. (2017a). *Bruttoinlandsprodukt, Bruttowertschöpfung in den Ländern*. Retrieved July 2, 2017, from https://www.destatis.de/DE/Publikationen/Thematisch/Volkswirtschaftliche-Gesamtrechnungen/VGRderLaender/VGR_LaenderergebnisseBand1.html
- Destatis. (2017b). *Publikation – Preise – Verbraucherpreisindex für Deutschland – Lange Reihen ab 1948*. Retrieved November 10, 2017, from <https://www.destatis.de/DE/Publikationen/Thematisch/Preise/Verbraucherpreise/VerbraucherpreisindexLangeReihen.html>
- Diener, E., Inglehart, R., & Tay, L. (2013a). Theory and validity of life satisfaction scales. *Social Indicators Research*, 112(3), 497–527.
- Diener, E., Tay, L., & Oishi, S. (2013b). Rising income and the subjective well-being of nations. *Journal of Personality and Social Psychology*, 104(2), 267–276.
- Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. *Nations and Households in Economic Growth*, 89, 89–125.
- Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization*, 27(1), 35–47.
- Easterlin, R. A. (2005). Feeding the illusion of growth and happiness: A reply to Hagerty and Veenhoven. *Social Indicators Research*, 74(3), 429–443.
- Easterlin, R. A. (2009). Lost in transition: Life satisfaction on the road to capitalism. *Journal of Economic Behavior & Organization*, 71(2), 130–145.
- Easterlin, R. A. (2015). Happiness and economic growth: The evidence. In W. Glatzer, L. Camfield, V. Möller, & M. Rojas (Eds.), *Global handbook of quality of life* (pp. 283–299). International Handbooks of Quality-of-Life. Dordrecht: Springer.
- Easterlin, R. A. (2017). Paradox lost? *Review of Behavioral Economics*, 4(4), 311–339.
- Easterlin, R. A., & Sawangfa, O. (2010). Happiness and economic growth: Does the cross section predict time trends? Evidence from developing countries. In E. Diener & D. Kahneman (Eds.), *International differences in well-being* (pp. 166–216).
- Easterlin, R. A., McVey, L. A., Switek, M., Sawangfa, O., & Zweig, J. S. (2010). The happiness-income paradox revisited. *Proceedings of the National Academy of Sciences*, 107(52), 22463–22468.
- Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2015). The next generation of the Penn World Table. *The American Economic Review*, 105(10), 3150–3182.
- Gruen, C., & Klasen, S. (2013). Income, inequality, and subjective well-being: an international and intertemporal perspective using panel data. *Jahrbuch Für Wirtschaftsgeschichte / Economic History Yearbook*, 54(1), 15–35.

- Hamilton, J. D. (2017). Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics*, Just accepted MS, available online, https://doi.org/10.1162/REST_a_00706, 1–25.
- Heske, G. (2009). Volkswirtschaftliche Gesamtrechnung DDR 1950–1989. Daten, Methoden, Vergleiche. *Historical Social Research/Historische Sozialforschung, Supplement, 21*, 1–356.
- Hodrick, R. J., & Prescott, E. C. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking, 29*(1), 1–16.
- IMF. (2016). *World economic outlook database*. Retrieved July 2, 2017, from <http://www.imf.org/external/pubs/ft/weo/2016/02/weodata/index.aspx>
- Inman, P. (2016, July 12). Irish economy surges 26% as revised figures take in foreign investment. *The Guardian, Business*. Retrieved June 29, 2017, from <https://www.theguardian.com/business/2016/jul/12/irish-economic-growth-revised-figures-foreign-investment-aircraft>
- Kaiser, C. W., & Vendrik, M. C. M. (2018). *Different versions of the Easterlin Paradox: New evidence for European countries*. IZA DP No. 11994. Bonn: IZA.
- Layard, R., Mayraz, G., & Nickell, S. J. (2010). Does relative income matter? Are the critics right? In E. Diener, J. Helliwell, & D. Kahneman (Eds.), *International differences in wellbeing* (pp. 139–165). Oxford: Oxford University Press.
- OECD. (2017). *Level of GDP per capita and productivity*. Retrieved July 2, 2017, from https://stats.oecd.org/Index.aspx?DataSetCode=PDB_LV#
- Opfinger, M. (2016). The Easterlin paradox worldwide. *Applied Economics Letters, 23*(2), 85–88.
- Proto, E., & Rustichini, A. (2013). A reassessment of the relationship between GDP and life satisfaction. *PLoS One, 8*(11), 1–10.
- Ravn, M. O., & Uhlig, H. (2002). On adjusting the Hodrick-Prescott filter for the frequency of observations. *The Review of Economics and Statistics, 84*(2), 371–376.
- Sacks, D. W., Stevenson, B., & Wolfers, J. (2012). Subjective well-being, income, economic development and growth. In P. Booth (Ed.), *The pursuit of happiness: Well-being and the role of government* (pp. 59–97). London: The Institute of Economic Affairs.
- Sacks, D. W., Stevenson, B., & Wolfers, J. (2013). *Growth in income and subjective well-being over time*. Mimeo.
- Stevenson, B., & Wolfers, J. (2008). Economic growth and subjective well-being: Reassessing the Easterlin Paradox. *Brooking Papers on Economic Activity, 2008*(1), 1–87.
- UNCTAD. (2017). *Gross domestic product: Total and per capita, current and constant (2005 prices, annual, 1970–2015)*. Retrieved July 2, 2017, from <http://unctadstat.unctad.org/wds/TableViewer/tableView.aspx?ReportId=96>
- Veenhoven, R. (2011). Trend average happiness in nations 1946–2010: How much people like the life they live. *Trend Report Average Happiness*. World Database of Happiness, Happiness in Nations. Rotterdam: Erasmus University.
- Veenhoven, R., & Hagerty, M. (2006). Rising happiness in nations 1946–2004: A reply to Easterlin. *Social Indicators Research, 79*(3), 421–436.
- Veenhoven, R., & Vergunst, F. (2014). The Easterlin illusion: economic growth does go with greater happiness. *International Journal of Happiness and Development, 1*(4), 311–343.
- Vendrik, M. C. M. (2013). Adaptation, anticipation and social interaction in happiness: An integrated error-correction approach. *Journal of Public Economics, 105*, 131–149.
- Wooldridge, J. M. (2003). *Introductory econometrics; A modern approach* (2nd ed.). Mason: South-Western Thomson Learning.
- World Bank. (2017). *World Development Indicators*. Retrieved July 2, 2017, from <http://databank.worldbank.org/data/reports.aspx?source=2>