WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER

CIW CENTRUM FÜR INTERDISZIPLINÄRE WIRTSCHAFTSFORSCHUNG

Centrum für Interdisziplinäre Wirtschaftsforschung
Diskussionspapier

4/2013

Testing the Easterlin Hypothesis with Panel Data: The Dynamic Relationship Between Life Satisfaction and Economic Growth in Germany and in the UK

Tobias Pfaff, Johannes Hirata

Center for Interdisciplinary Economics
Discussion Paper

Centrum für Interdisziplinäre Wirtschaftsforschung
Diskussionspapier
4/2013

February 2013

ISSN 2191-4419

Testing the Easterlin Hypothesis with Panel Data: The Dynamic Relationship Between Life Satisfaction and Economic Growth in Germany and in the UK

Tobias Pfaff and Johannes Hirata

Abstract

Recent studies focused on testing the Easterlin hypothesis (happiness and national income correlate in the cross-section but not over time) on a global level. We make a case for testing the Easterlin hypothesis at the country level where individual panel data allow exploiting important methodological advantages. Novelties of our test of the Easterlin hypothesis are a) long-term panel data and estimation with individual fixed effects, b) regional GDP per capita with a higher variation than national figures, c) accounting for potentially biased clustered standard errors when the number of clusters is small. Using long-term panel data for Germany and the United Kingdom, we do not find robust evidence for a relationship between GDP per capita and life satisfaction in either country (controlling for a variety of variables). Together with the evidence from previous research, we now count three countries for which Easterlin's happiness-income hypothesis cannot be rejected: the United States, Germany, and the United Kingdom.

Keywords: Subjective well-being, economic growth, income, Easterlin hypothesis

Zusammenfassung

Die neuere Forschung hat sich darauf konzentriert die Easterlin-Hypothese (Wohlbefinden und Volkseinkommen korrelieren im Querschnitt, aber nicht in der Zeitreihe) auf globaler Ebene zu testen. Unser Artikel liefert Argumente für das Testen der Easterlin-Hypothese auf nationaler Ebene, wo individuelle Paneldaten das Ausschöpfen wichtiger methodologischer Vorteile ermöglichen. Wir erweitern die bisherige Literatur zur Easterlin-Hypothese durch a) Schätzungen mit individuellen fixen Effekten anhand von längerfristigen Paneldaten, b) Verwendung von regionalem BIP pro Kopf mit einer höheren Varianz als nationale BIP-Daten und c) Berücksichtigung von potentiell verzerrten Cluster-Standardfehlern im Fall von wenigen Clustern. Wir verwenden längerfristige Paneldaten für Deutschland und Großbritannien und finden keine robuste Evidenz für einen Zusammenhang zwischen BIP pro Kopf und Lebenszufriedenheit in den beiden Ländern (unter Verwendung von einer Reihe von Kontrollvariablen). Zusammen mit früheren Forschungsergebnissen zählen wir drei Länder in denen die Easterlin-Hypothese nicht verworfen werden kann: die Vereinigten Staaten, Deutschland und Großbritannien.

Schlagwörter: Subjektives Wohlbefinden, Wirtschaftswachstum, Einkommen, Easterlin-Hypothese

Im Internet unter:

http://www.wiwi.uni-muenster.de/ciw/forschen/downloads/DP-CIW_04_2013.pdf

Westfälische Wilhelms-Universität Münster
CIW – Centrum für Interdisziplinäre Wirtschaftsforschung
Scharnhorststraße 100
D-48151 Münster

Tel: +49-251/83-25329 (Sekretariat)
E-Mail: clementine.kessler@uni-muenster.de
Internet: www.wiwi.uni-muenster.de/ciw

1. Introduction

Does economic growth improve the human lot? Since Richard Easterlin's seminal 1974 paper, the question of how exactly economic growth affects subjective well-being has given rise to a lively and controversial debate.[1] Over the years, a series of empirical studies has tried to test the famous happiness-income paradox (better known as the Easterlin paradox or Easterlin hypothesis), i.e., the hypothesis that "at a point in time both among and within nations, happiness varies directly with income, but over time, happiness does not increase when a country's income increases" (Easterlin et al., 2010, p. 1).[2] Easterlin stresses the long-term perspective of the hypothesis, i.e., 10 years or more.

Easterlin has long recognized the strong positive cross-sectional relationship between income and subjective well-being within countries (Easterlin, 1974) as well as across countries (Easterlin, 1995). However, some authors look at the cross-sectional evidence of the relationship between national income and subjective well-being and then go on to draw unwarranted conclusions for the relationship over time (e.g., Arrow and Dasgupta, 2009; Guriev and Zhuravskaya, 2009). On the other hand, new studies rely on time series data of countries and indeed find a positive relationship between national income and happiness over time for several countries, contradicting the Easterlin hypothesis (e.g., Sacks et al., 2010, 2011; Stevenson and Wolfers, 2008). In short, there is no consensus yet on the dynamic relationship between economic growth and subjective well-being.[3] This study addresses the question of how individuals' subjective well-being is affected over time by, on the one hand, the growth of Gross Domestic Product (GDP) and, on the other hand, by the growth of their own income, controlling for a number of other potential influences. As a novelty, we use individual panel data, which allows us to control for individual fixed effects.

Using individual fixed effects has several important methodological advantages (cf. Vendrik and Woltjer, 2007). Fixed-effects estimation enables us to isolate the *dynamic* relationship between subjective well-being and national income, stripped of any potentially confounding static patterns (using only the within-variation, while disregarding the between-variation).[4] With fixed effects, we can also rule out potential disturbances by time-invariant unobserved heterogeneity, such as birth cohort, family background or even, for some individuals, neighborhood, permanent health conditions, etc. In particular, fixed effects eliminate the influence of stable personality traits, some of which are well-known to correlate strongly with subjective well-being (Diener and Lucas, 1999). By stripping the error term of any time-constant factors which could be potentially correlated with the regressors, fixed-effects es-

[1] In this paper, we follow the simple definition of economic growth as increase in (real) GDP per capita.
[2] The definition of the Easterlin hypothesis appears in different versions in the literature. We propose that tests of the Easterlin hypothesis should refer to this definition, which is clearly stated by Easterlin himself.
[3] In the remainder of this article, we will adopt Alan Krueger's terminology using "Easterlin hypothesis" instead of "Easterlin paradox" in order to reflect this lack of consensus (Stevenson and Wolfers, 2008, p. 96).
[4] The dynamic relationship has been analyzed before in studies using *macro* data and *country* fixed effects (e.g., Hagerty, 2000; Sacks et al., 2010). See Section 3 for a detailed overview of previous studies.

timation also reduces a potential endogeneity bias that could not be ruled out by previous studies on the Easterlin hypothesis.

It is plausible that singular events happening within a nation in a specific year affect the life satisfaction of individuals. We want to make sure that our estimates are not tainted by such events, and we achieve this by controlling for year fixed effects. To circumvent the problem of perfect collinearity between a full set of year dummies and national GDP data, we use regional GDP data with the positive side effect of increased statistical power of the tests thanks to larger variance.

Panel surveys that include subjective well-being questions and cover at least 10 years are scarce. The two longest running panel data sets with questions on subjective well-being match our criteria: the German Socio-Economic Panel (SOEP) and the British Household Panel Survey (BHPS). Fortunately, for both of these countries, regional GDP data are available. We will analyze both of these datasets in turn.

Our analysis proceeds as follows. After discussing theoretical considerations regarding the mechanics of GDP, income, and subjective well-being in Section 2, we zoom in on the core of the dispute around the Easterlin hypothesis by means of a systematic comparison of relevant studies in Section 3. We explain our empirical identification strategy in Section 4. Descriptive and analytical results are presented in Section 5. A series of robustness checks is presented in Section 6, followed by a brief discussion of our results in Section 7. Section 8 concludes.

2. A theory of the mechanics of GDP, income, and subjective well-being

GDP is a measure of the total monetary value of the economic output of a geographical entity within a given period of time, usually calculated at the national or regional level. Setting the measure in relation to the size of the underlying population provides information on the average economic output per person (GDP per capita).

[Figure 1 about here]

Fig. 1 shows the channels through which GDP growth may influence subjective well-being. Under normal circumstances, steady economic growth may be a favorable condition for political stability, a more effective civil society, better education, better health care, better infrastructure, etc. (Friedman, 2005). Empirical evidence shows that most of these aspects are in fact positively correlated with subjective well-being (Dolan et al., 2008). On the other hand, an increase in GDP per capita can also give rise to negative externalities such as environmental degradation or erosion of social capital (Fleurbaey, 2009; Putnam, 2000; van den Bergh, 2009), which tend to reduce subjective well-being.

Conventionally, the primary channel through which economic growth is thought to affect subjective well-being is an increase in consumption possibilities. We use the term “absolute income effect” to describe this effect. Economic growth can also lead to a "relative income effect", i.e., a change in one person's subjective well-being induced by the change of others' income, holding own income constant. The relative income effect can be split into a positive "information effect" (ambition) and a negative "comparison effect" (jealousy), as argued by Senik (2004, 2008) following the work of Hirshman and Rothschild (1973). Senik (2008) proposes that these two partial effects always coexist but that "the degree of mobility and uncertainty in the economic environment" (p. 496) determines which of the two is dominant. In societies with high (perceived) socio-economic mobility, e.g., transition countries in Eastern Europe, a rise in others' income is more likely to induce positive feelings such as optimism and ambition because individuals tend to interpret this as a precursor of a better future for themselves. However, in countries with lower (perceived) socio-economic mobility, a rise in others' income is more likely to reduce a person's subjective well-being due to, e.g., a loss of socio-economic status.

While an income shock might have a sizable absolute income effect on people's subjective well-being in the short run, individuals may adapt—fully or partially—to income changes in the long run. In other words, individual well-being could gradually revert to the ex-ante level over time.[5] Early theoretical work has been done by economists Pollak (1970) and van Praag (1971), the latter of which refers to a "preference drift" over time. Psychologists Brickman and Campbell (1971) coined the term "hedonic treadmill" for this phenomenon.[6]

The bottom line is that theory alone cannot predict whether a rise in GDP per capita leads to an increase in subjective well-being. It is even conceivable that a rise in GDP brings about negative effects on such a scale that well-being is actually diminished.[7] This fundamental ambiguity seems to be at the heart of the divergent empirical findings of the dynamic relationship between subjective well-being and GDP per capita as discussed in Section 3.

In the light of the various channels through which GDP per capita may affect well-being as sketched in Fig. 1, we are rather pessimistic that empirical studies of the effect of GDP per capita will ever lead to unambiguous results valid for contexts as diverse as high-income and low-income countries. Therefore, we prefer to focus on individual countries.

Easterlin's paradoxical findings of flat curves of subjective well-being over long periods of remarkable economic growth are usually explained with relative income effects and adaptation to rising levels of income. However, tests of relative-income effects are faced with the difficulty of constructing plausible proxies for reference income. The results shown in Pfaff (2013b) cast doubt on some common methods

[5] Adaptation (or habituation) has also been discussed in relation to other life events. See Frederick and Loewenstein (1999) and Clark et al. (2008a) for reviews.
[6] Clark et al. (2008b) provide an excellent overview of theoretical and empirical studies of relative income and adaption effects.
[7] The theoretical ambiguity is also a strong theoretical case for seeking better measures of societal welfare instead of gauging welfare with GDP per capita (Stiglitz et al., 2010).

for measuring reference-income effects. He also does not find robust evidence for adaptation to income after four years in Germany with samples that are similar to the ones used in this study. Therefore, our data do not allow us to disentangle all of the mechanisms depicted in Fig. 1. We reemphasize, therefore, that our objective is not to identify all possible causes for flat curves of subjective well-being, but to separately quantify the respective effects of GDP per capita and individual income on subjective well-being.

3. Previous studies on the relationship between GDP per capita and subjective well-being

We present a comprehensive overview of the ambiguous findings on the relationship between GDP per capita and subjective well-being in Table 1.[8] Building on Clark and Senik's (2010b, pp. 161–162) classification, we group models by their focus on the static or dynamic relationship (i.e., cross-sectional or time-series data) and by usage of macro or micro data (i.e., average or individual subjective well-being).

[Table 1 about here]

In our overview, Easterlin's regressions are the only ones restricted to a specific country, focusing on the United States (Easterlin, 2005b) and on Japan (Easterlin, 2005a).[9] All other regressions are based on multi-country analyses, with the Gallup World Poll as the most comprehensive, or "first representative sample of planet Earth" (Diener et al., 2010, p. 52). The time span for analyses of the dynamic relation ranges from 18 to 35 years.[10] The number of observations ranges from 24 (macro data) to 850,153 (micro data). The specific subjective well-being question of the survey determines the dependent variable and ranges from a 3-point scale happiness question in the General Social Survey to an 11-point scale life evaluation question (Cantril's ladder) in the Gallup World Poll. GDP per capita is our primary variable of interest. The standard method is to take the (natural) logarithm of real GDP per capita because of the assumption of decreasing marginal utility of income (Layard et al., 2008).[11] However, some models deviate from the standard and do not use logarithms, or they use some other specification (as explained in the notes of Table 1).

Deaton (2008) and Sacks et al. (2010) are the latest of prominent cross-section studies on the static relationship between GDP per capita and average subjective well-being. They confirm, once again, the

[8] Some studies have more models with GDP per capita than are shown in Table 1. In these cases, we have picked the models that we deemed most relevant while attempting to avoid misrepresenting the range of sizes and significance levels of the coefficients. We also left out models/studies that analyzed financial satisfaction or change in life satisfaction as dependent variable. We neither consider studies analyzing GDP instead of GDP per capita.
[9] Stevenson and Wolfers (2008) argue that Easterlin's (2005a) results for Japan are flawed because of series breaks in the wording of the survey questions.
[10] Note that the number of years may differ from the number of waves.
[11] Note that using the logarithm of income does not imply that the effect of income on subjective well-being becomes nil for high income levels.

earlier results of Easterlin (1974): richer countries enjoy higher levels of average well-being. The significantly positive static relationship also holds when micro data are used (Diener et al., 2010; Sacks et al., 2010). In the analysis of Diener et al. (2010), GDP per capita even has the largest standardized coefficient among the predictors of life evaluation used.

However, it is the divergent findings on the dynamic relationship between national income and subjective well-being which keep the debate on the Easterlin hypothesis alive. In most models and using either macro or micro data, GDP per capita enters positively, at least significant at the five percent level. Exceptions are the one-country regressions by Easterlin (Easterlin, 2005a, b) with coefficients insignificantly different from zero (and partly negative).[12] Another non-significant coefficient appears in Inglehart et al. (2008) using a 4-point scale happiness question, and Sacks et al. (2011) find insignificant coefficients in 4 out of 7 panel regressions.[13] For the 10-point scale life satisfaction question the coefficient becomes significant. The only insignificant micro-data result we were able to find appears in Di Tella et al. (2003) once they add two lags of GDP per capita.[14]

From the results showing a significantly positive relationship, it is interesting to observe in Stevenson and Wolfers' (2008) micro-data analysis that the coefficient drops sharply from .737 to .192 once country fixed effects are introduced. Moreover, once country and year fixed effects are added, the coefficient for GDP per capita increases slightly to .208, while losing some of its significance. This confirms the importance of adding year dummies, which obliges us to use regional GDP per capita in our empirical strategy with single-country data sets.

The true dynamic relationship is revealed when only the within-variation is used, which can be achieved with macro data by adding country dummies to the model. Such models are estimated by Hagerty (2000) and Sacks et al. (2011), producing diverging results.

The recent analysis of Diener et al. (2013) applies a hierarchical linear model to macro data from the Gallup World Poll for 135 countries and the period 2005–2011. Although important to the field, we do not include this study in Table 1 because coefficients cannot be readily compared. Diener et al. (2013) conclude that changes in GDP per capita significantly predict changes in life evaluation, while Sacks et

[12] The research group around Richard Easterlin has found more negative coefficients with larger *t*-values, but then for the change of GDP per capita and with the change in life satisfaction as the dependent variable (Easterlin, 2009; Easterlin and Angelescu, 2009; Easterlin and Sawangfa, 2010). For comparability reasons, these studies are not shown in Table 1.

[13] We did not count the significant result in their somewhat daring "panel of panels", where several data sets with different questions on subjective well-being are combined into one sample.

[14] Surprisingly, the coefficient more than doubles after adding five lags of GDP per capita in a later study by Di Tella and MacCulloch (2010) where they use the same data set, but for another time period.

al. (2011) show an insignificant coefficient for GDP per capita applying a different estimation approach to the same data set.[15]

While the results of previous studies presented in Table 1 point to a positive dynamic relationship between national income and subjective well-being, one should take note of two issues before taking these results as a falsification of the Easterlin hypothesis.[16] The first issue concerns the need for clustering of standard errors when observations are grouped in clusters (Cameron and Miller, 2011). Without clustering, standard errors can be biased downwards and statistical significance would thus be overstated (see Section 4.2.). Some of the studies in Table 1 use multi-country data sets, but apparently account neither for possible within-cluster correlation nor for serial correlation (e.g., Di Tella and MacCulloch, 2010; Di Tella et al., 2003; Diener et al., 2010; Hagerty, 2000; Inglehart et al., 2008). Some of the studies appropriately use clustered standard errors, but do not account for potential bias if the number of clusters is small (e.g., Di Tella and MacCulloch, 2008; Sacks et al., 2010, 2011; Stevenson and Wolfers, 2008). When reading the results, one should be aware that there could be either form of potential bias, whereas both can lead to underestimation of standard errors and overstatement of the significance of the statistics.

The second issue preventing a general falsification of the Easterlin hypothesis is the fact that the comprehensive study by Stevenson and Wolfers (2008) shows one important exception: the United States. The authors acknowledge that "there is a clear evidence of the absence of a time-series happiness-income relationship". They conclude that "[a]lthough the U.S. time series is thus a data point supporting the Easterlin paradox, it should be regarded as an interesting exception warranting further scrutiny." (2008, p. 58). While most of the evidence based on multi-country data suggests a positive dynamic influence of GDP per capita on subjective well-being, we are aware of the U.S. exception and again make a case for the importance of scrutinizing the Easterlin hypothesis on the country level.

[15] In contrast to the results for live evaluation, Diener et al. (2013) suggest that changes in GDP per capita do not significantly predict *emotional* well-being. Kahneman and Deaton (2010) have stressed earlier that analyses of income and well-being should distinguish between life evaluation and emotional well-being. However, we have doubts that the Easterlin hypothesis with its focus on a long-term relationship can be tested with data from questions with a short-term focus, such as yesterday's emotional feelings.

[16] Moreover, our Table 1 shows that Sacks et al. (2012, p. 1185) are not correct in concluding that all data sets they have studied show significant evidence "that those countries which enjoyed faster economic growth, on average experienced greater growth in well-being".

4. Empirical strategy for testing the Easterlin hypothesis in Germany and the United Kingdom

4.1. From macro models to micro models with individual fixed effects

Our aim is to test the validity of the Easterlin hypothesis. In other words, we use models that allow us to test if the dynamic, long-term relationship between subjective well-being and economic growth is nil.[17] Among measures of subjective well-being, we choose life satisfaction. Life satisfaction has a broader scope than, e.g., happiness, which is considered to reflect a more momentary evaluation of well-being. This broader scope conforms with our test of long-term effects.

We begin our empirical strategy with mimicking macro and micro models of previous studies, before introducing individual fixed effects. The macro model has the form

$$LS_t = \alpha + \beta \ln(GDP\_PC_{t-1}) + \varepsilon_t, \tag{1}$$

where LS_t is average life satisfaction in year t, GDP_PC_{t-1} is national GDP per capita of the previous year, and ε_t is a random error term. The error term reflects the fact that in reality many factors other than GDP per capita have an influence on life satisfaction. We focus on the preceding year's GDP per capita because of the fieldwork periods of the surveys, but we will also test current GDP per capita in the micro models.[18] The micro models without individual fixed effects have the general form

$$LS_{ijt} = \alpha + \beta \ln\big(GDP\_PC_{j,t-1}\big) + \lambda_t + \varphi_j + u_{jt} + \varepsilon_{ijt}, \tag{2}$$

where LS_{ijt} is life satisfaction of individual i in region j in year t, $GDP\_PC_{j,\,t-1}$ is GDP per capita in region j of the previous year, λ_t refers to year fixed effects, φ_j refers to region fixed effects, u_{jt} is a region-year error component, and ε_{ijt} is an individual error term. We begin the micro data analysis without individual fixed effects followed by a stepwise introduction of region and year fixed effects in order to compare our results with the results of Stevenson and Wolfers (2008).

The main part of our analysis is dedicated to micro models with individual fixed effects of the general form

[17] Sacks et al. (2011) argue that a test of the Easterlin hypothesis should rather focus on the similarity of the coefficients from the within-country cross-section, the between-country cross-section, and the national time-series. Given the empirical evidence for a positive relation of (national) income and well-being in the cross-section, we conclude that a simple and straightforward test of the time-series relation is not inferior to their approach.

[18] Usually more than 90 percent of the SOEP interviews are conducted in the first half of the year. BHPS interviews are usually conducted from September until May. GDP represents economic transactions of the full year. It is not very plausible to analyze the influence of GDP on life satisfaction values stated far from the end of the year. It is for our data thus more straightforward to take GDP of the previous year. We presume that the relation between last year's GDP per capita and current life satisfaction is stronger than the relation between current GDP per capita and current life satisfaction.

$$LS_{ijt} = \alpha_i + \beta \ln\left(GDP\_PC_{j,t-1}\right) + \delta \ln\left(INC_{ijt}\right) + \gamma' X_{ijt} + \lambda_t + \varphi_j + u_{jt} + \varepsilon_{ijt}, \tag{3}$$

where α_i refers to individual fixed effects, INC_{ijt} is individual income, and the vector X_{ijt} refers to a set of further micro control variables. All control variables are described in Appendix A.[19] OLS is performed on the mean-differenced data to obtain the within estimator.[20]

At first, we estimate an individual fixed effects model without the set of control variables so that we can isolate the impact of controlling for unobserved heterogeneity. We then add the set of micro control variables. In a further step, we add individual income and expect the coefficient for regional GDP per capita to shrink because the coefficient should now be net of a likely positive individual income effect, consistent with the theoretical model described in Section 2. In a subsequent model we slightly alter equation (3) and use current regional GDP per capita (rather than that of the previous year).

It can be argued that people compare their income to the average income in one's region. In order to avoid that our coefficient of regional GDP per capita partly reflects such a regional relative-income effect, we also estimate a model augmented by a term for average regional income:

$$LS_{ijt} = \alpha_i + \beta \ln\left(GDP_{PC\,j,t-1}\right) + \delta \ln\left(INC_{ijt}\right) + \mu \ln\left(\overline{INC}_{jt}\right) + \gamma' X_{ijt} + \lambda_t + \varphi_j + u_{jt} + \varepsilon_{ijt} \tag{4}$$

$\overline{INC}_{jt}$ is average income of region j in year t. We expect the coefficient for average regional income to be insignificant, because we know from the literature that average regional income is not a likely yardstick for comparisons because "people compare to the groups with whom they interact more frequently" (Clark and Senik, 2010a, p. 585), that means neighbors, friends, and foremost colleagues. Note, however, that equations like (4) with one regressor being the average of another regressor potentially bear identification problems (Angrist and Pischke, 2009, pp. 192–197). Results of this model should be interpreted with caution.

OLS requires stationary data to work properly. Otherwise, results might be biased (Granger and Newbold, 1974). This bias problem is rarely addressed in the literature, with the exceptions of Di Tella et al. (2003) and Sacks et al. (2011). To mitigate the problem of using potentially trended variables such as levels of GDP per capita with OLS, Di Tella et al. (2003) propose using GDP growth rates or other variables measured relative to trend. In our case, GDP per capita and individual income may, in principle, be trended. We therefore estimate a model similar to equation (3) where possible trends are

[19] With individual fixed effects, we do not use education as a control variable, in contrast with many other studies (e.g., Di Tella et al., 2010; Layard et al., 2010). As Dolan et al. (2008) convincingly argue, "most adult survey respondents are unlikely to change their education level during their time in a panel survey, and consequently fixed effects models are unlikely to find any significant effect for education" (p. 100). One example in which the effect of education vanishes when using individual fixed effects is Oswald and Powdthavee (2008).

[20] We prefer OLS because results can be easily interpreted. Also, results are usually not qualitatively different if ordinality of the dependent variable is assumed (Ferrer-i-Carbonell and Frijters, 2004). We nevertheless describe a robustness check with an estimation method that takes the ordinal character of our dependent variable into account (see Section 6).

removed by replacing levels of GDP per capita with the growth rate of GDP per capita from *t*-2 to *t*-1, and another model with the growth rate of GDP per capita from *t*-1 to *t*. In these models, levels of *individual* income are replaced with the respective growth rate from *t*-1 to *t*.

4.2. Econometric treatment of cluster correlation

The assumption of independent disturbances is usually not valid for regressions of a micro variable on an aggregate regressor (Moulton, 1990). If the group structure of the errors remains unaccounted for, OLS standard errors can be severely biased downwards, with the consequence of over-rejecting *t*-tests. The comfortable solution to account for the group structure is to use cluster-robust standard errors as proposed by Liang and Zeger (1986). However, the asymptotic theory behind the calculation of cluster-robust standard errors requires a large number of clusters (Wooldridge, 2003). An insufficient number of clusters (approximately less than 50) can once again lead to drastic overstatement of the significance of statistics (Donald and Lang, 2007). For such cases, the literature proposes several methods for adjusting standard errors or *t*-statistics (Angrist and Pischke, 2009; Cameron and Miller, 2011; Pfaff, 2013a). Alas, the bottom-line is that no perfect solution has yet been found to correctly adjust standard errors if the number of clusters is small.

In our setting, we have a micro independent variable and an aggregate key regressor (namely regional GDP per capita), while the number of clusters is small (between 6 and 12 regions). The adjustment method that is feasible and seems most promising for our setting is wild cluster bootstrap.[21] Cameron et al. (2008) find that wild cluster bootstrap performs well in cases with few clusters. For aggregate key regressors, we therefore estimate *p*-values with the wild cluster bootstrap-*t* procedure in order to re-assess the significance of the statistics.[22] Wild bootstrap requires an additively separable error term and therefore does not work with ordered probit. For ordered probit regressions with few clusters we derive *p*-values from the pairs cluster bootstrap-*t* procedure.[23] Note that we also take account of serial correlation by clustering on the region level while assuming that the regions are independent (Angrist and Pischke, 2009, p. 319).

The calculation of cluster-robust standard errors works only with nested data. However, panel data as ours are typically non-nested in regions because some individuals move between regions. An approach for clustering standard errors with non-nested data is two-way clustering (Cameron et al., 2011;

[21] Both, bias-reduced linearization (Bell and McCaffrey, 2002) and its modified version proposed by Imbens and Kolesar (2012) seem not to work with large samples like ours. The between-group estimator proposed by Donald and Lang (2007) only works for regressors that are fixed within groups, which is not the case for regional GDP per capita that varies over time. The approach of Ibragimov and Müller (2010) is not feasible in our models with time dummies. Finally, the parametric correction with the Moulton factor could suffer from poor estimation of the intraclass correlation coefficient if the number of groups is small (Feng et al., 2001).

[22] In contrast to bootstrap-se procedures, bootstrap-*t* procedures have the advantage of providing asymptotic refinement (Cameron et al., 2008).

[23] The bootstrap-*t* procedures for pairs cluster bootstrap and wild cluster bootstrap are explained in Appendix B of Cameron et al. (2008).

Thompson, 2011). Again, two-way clustering assumes that the number of clusters in each cluster dimension is sufficiently large. A solution for the problem of biased two-way clustered standard errors in settings with few clusters is yet to be developed. As a consequence, we prefer to present potentially unbiased (or at least less biased) inference results using the wild bootstrap method, even if this means that we can only use data which are nested within regions.

We produce a nested data set from the originally non-nested data by keeping only the region in which the individual stayed the longest in our period of analysis. If we cannot identify a main region of residence for an individual (e.g., a person lives four years each in two different regions), we drop all observations for this individual.[24] Obviously, we need to make sure that this selection process does not influence our results. We address this problem in Section 6.

4.3. Addressing endogeneity

Our principal concern with our identification strategy is that we cannot rule out endogeneity bias. Endogeneity bias is caused by violating the assumption that regressors are uncorrelated with the error term (Antonakis et al., 2010).[25] In our setting, we suspect that endogeneity could be an issue due to measurement error and due to omitted variable bias. Considering measurement error, we suspect life satisfaction, GDP per capita, household income, and health satisfaction as primary candidates. Measurement error of the dependent variable still leads to unbiased estimators if we assume that the error of measurement in life satisfaction is uncorrelated both with the regressors and with the error term (Gujarati and Porter, 2009, p. 483). Measurement error of the dependent variable would then lead to larger standard errors. If the regressor is measured correctly, Greene (2008, p. 326) argues that one can ignore the measurement error on the dependent variable because it can be absorbed in the error term of the regression. Thus, we are not particularly concerned with potential measurement error for our dependent variable. Modeling measurement error for the independent variables would be possible with some reliability measure, which our data do not provide.

Omitted variable bias seems more problematic in our setting. Our concern is somewhat mitigated by the fact that mean-differentiation applies to our fixed-effects models, whereby α_i is eliminated. Eliminating α_i allows for consistent estimation of endogenous regressors, provided that the endogenous regressors are only correlated with the time-constant component of the error, α_i, and uncorrelated with the time-varying component ε_{ijt} (Cameron and Trivedi, 2010, p. 257). However, it is still conceivable that regional GDP per capita is correlated with other region-year effects represented by the error component u_{jt} of equation (3). Our results are therefore somewhat vulnerable. We would appreciate if future research identifies solid instruments for GDP per capita, notwithstanding the fact that finding such in-

[24] The alternative would have been to drop all individuals which move between regions with an even greater loss of observations. We cannot think of a reason why dropping all movers would be superior to our method.
[25] We check further OLS assumptions for our results in Appendix C.

struments with life satisfaction as the dependent variable is an arduous endeavor. Nonetheless, we believe that our fixed-effects results are less vulnerable than the results of previous studies, most of which did not address endogeneity at all.[26]

4.4. Brief description of the panel data sets SOEP and BHPS

Our first data set is the German Socio-Economic Panel (SOEP, 2011), the world's longest-running socio-economic panel study with the first wave in 1984 (Wagner et al., 2007). The primary question of interest is: *"How satisfied are you with your life, all things considered?"*, and the answers range from 0 (*"completely dissatisfied"*) to 10 (*"completely satisfied"*). The German re-unification in 1990 had a strong impact on the lives and satisfaction levels of East Germans (Frijters et al., 2004). We want to avoid confounding our results by effects of the re-unification and divide the sample by Western and Eastern Germany. The Western German sample consists of 27 waves covering the period of 1984–2010. For Eastern Germany, we use 19 waves (1992–2010).[27]

The second data set is the British Household Panel Survey (BHPS, 2012), which was started in 1991. The BHPS asks for life satisfaction on a 7-point scale: *"How dissatisfied or satisfied are you with your life as a whole?"*. The question was introduced in wave 6, but not asked in wave 11. This allows us to use waves 6–10 and 12–18, covering 12 waves or the years 1996–2008 (without 2001). Because the UK Office for National Statistics does not provide regional GDP, we use regional Gross Value Added (GVA).[28] Although we refer to the United Kingdom in this paper, note that the BHPS was extended to Northern Ireland only in wave 11. Data on regional GDP/GVA per capita and price levels are from the German Federal Statistical Office and the UK Office for National Statistics.

We restrict our samples to adults (> 18 years), but we do not truncate age upwardly because we want to analyze the effects of GDP per capita independent of age group or working status. Our proxy for individual income is net household income in real terms and equivalized according to the modified-OECD scale (De Vos and Zaidi, 1997).[29] Regarding outlier treatment, we exclude the first percentile of real net equivalized household income because some values are implausibly low.[30]

For our clustering purposes, we require data nested in regions and lose 2.2 percent of observations by keeping only the main region of residence for an individual in the Western German sample. With the

[26] The only exception we could find is Di Tella et al. (2003) who briefly discuss endogeneity problems, but they do not present a solid quasi-experimental approach to overcome potential bias either.
[27] The SOEP sample was extended to Eastern Germany by 1990, but regional GDP per capita is only available for Eastern Germany beginning in 1991. Because we use GDP per capita (*t*-1), we can begin our analysis in 1992.
[28] GVA is GDP minus taxes on products plus subsidies on products. We treat the two concepts equally in our analysis, and refer only to GDP per capita in the text when all three samples are meant.
[29] For nominal to real transformations, we attempt to use price index data at the smallest geographical level possible. Details are explained in Appendix D.
[30] Other studies, e.g., Clark et al. (2005), exclude the first and last percentile of household income. However, in our sample, values in the last percentile still seem plausible.

same operation, we lose 2.3 percent in the Eastern German sample, and 2.3 percent in the UK sample. The Easterlin hypothesis refers to the long-term relationship between subjective well-being and economic growth, i.e., 10 years or more. The average number of years covered by an individual is 8.8 years in the Western German sample, 8.5 years in the Eastern German sample, and 6.6 years in the UK sample. The percentage of individuals covering at least 10 years is 38.7 percent, 42.6 percent, and 42.0 percent, respectively. Because an individual fixed-effects regression requires at least two interviews per individual as well as some variation of the life-satisfaction variable, we initially exclude all individuals who do not match either of these criteria.

5. Results

5.1. Descriptive statistics and preliminary analysis

For the primary variables of interest, Table 2 gives an overview of basic descriptive statistics.

[Table 2 about here]

By using regional GDP per capita, we obtain a higher variation than would be possible with national data, which increases statistical power. The standard deviation in Western Germany is close to what some studies show only for international comparisons in the cross-section (Hagerty and Veenhoven, 2003, p. 5). We will focus on the within estimator in individual fixed-effects regressions with the drawback that variables which vary relatively little over time are estimated rather imprecisely. The decomposition into overall, between, and within variation is shown in Tables B.1a–c in Appendix B. The within variation of regional GDP per capita in levels and log form is always smaller than the between variation. This means that the within estimation in the fixed-effects models leads to an efficiency loss compared to alternative estimators. However, the within variation of the growth rate of regional GDP/GVA per capita is always larger than the between variation. Besides the advantage that our growth rate variables can be regarded as stationary (see Section 4.1.), we acknowledge as a second advantage that the efficiency loss for growth rates using the within estimator is negligible.

At the macro level, the validity of the Easterlin hypothesis is often supported by graphs of aggregate time-series. The empirical analysis therefore begins with graphs of GDP per capita, household income, and life satisfaction in Western Germany, Eastern Germany, and the UK.

[Figure 2 about here]

[Figure 3 about here]

Visual inspection of Figures 2 and 3 – as well as the underlying data – show that life satisfaction in Western Germany and the United Kingdom exhibits a slightly negative trend, while the curve in Eastern

Germany shows no obvious trend.[31] At the same time, GDP per capita and household income show an upward trend over the whole period for all three samples.[32] This picture of a rise in national income coinciding with constant average life satisfaction is clearly consistent with the Easterlin hypothesis. However, only multiple regression analysis can uncover the hidden dynamics of income and subjective well-being.

Before we begin with the analytical section, we discuss the objection that an upwardly limited measure of life satisfaction is valid for the cross-section but not over time (e.g., Deaton, 2008, p. 70). The argument is as follows: if a person lives under rather miserable circumstances in time *t*, this person has a certain definition of a particular category of a fixed-scale life-satisfaction measure. When the same person is asked, for example, 20 years later, life circumstances might be much better, hence the definition of this particular category has changed, but the numerical value the person chooses could well be the same, given the upper limit of the rather narrow scale. This is why limited measures might not be able to reflect betterment in life. But consider Fig. 3 which compares average life satisfaction with the average GHQ-12 score in the UK from 1996–2008 (without 2001). The GHQ-12 (General Health Questionnaire) is a 12-item measure of psychological well-being (Vieweg and Hedlund, 1983). Each item has four categories that represent evaluations relative to a subjective anchor (e.g., *"more so than usual"*, *"same as usual"*, *"less so than usual"*, *"much less than usual"*), mapped to the values 0–3. The twelve responses are recoded in the BHPS so that the scale of the GHQ-12 goes from 0 (the least distressed) to 36 (the most distressed). The yearly weighted average of the GHQ-12 score ranges from 11.02 to 11.47 in our sample from 1996–2008. This means that the 12 questions were answered on average slightly below the neutral category *"same as usual"* in each year, which implies that there has been no improvement in average psychological well-being in the UK in the respective time period. Given the purely relative nature of the GHQ questions, the above argument against the validity of a fixed-scale life-satisfaction measure does not hold for the GHQ measure because subjective improvements over time should be reflected by GHQ-12 scores larger than 12. This finding suggests that subjective well-being in the UK was indeed rather constant for the respective time period, and that the reason for the flatness of the life-satisfaction curve in Fig. 3 is not the limited scale of the life-satisfaction question. Although we do not have similar data for Germany, the finding gives us some confidence that limited measures of life satisfaction are indeed suitable instruments for our time-series analyses, at least as long as the scores do not scratch the upper limit of the scale.

[31] Coefficients of OLS-fitted trend lines are -0.01 ($p < 0.01$) for Western Germany, 0.01 ($p = 0.09$) for Eastern Germany, and -0.01 ($p < 0.02$) for the UK.

[32] It is remarkable that the peak of Eastern German wealth in terms of average real equivalized net household income occurs in 2003 (Western Germany: 2010). In Western Germany, the slowly widening gap between average and median equivalized household income is apparent, increasing from 9 percent in 1984 to 13 percent in 2010 (measured in terms of average equivalized household income). The Eastern German gap is smaller at 7 percent in 1992 and almost 10 percent in 2010. In the UK we do not see a clear trend of a widening gap between average and median income.

5.2. Macro and micro estimates without individual fixed effects

We begin the regression analysis with a macro model and with micro models without individual fixed effects.[33] Results are shown in Table 3. Using OLS for macro data, we find highly significant negative coefficients for GDP per capita (*t*-1) for Western Germany and the UK, and an equally significant positive relationship for Eastern Germany. The result of a negative relationship in both Western Germany and in the UK qualitatively coincides with the result of Easterlin (2005b) for the U.S. The result of a positive relationship in Eastern Germany coincides with other macro regressions that show a significant positive relationship for a number of countries (e.g., Sacks et al., 2010). However, we agree with Clark and Senik (2010b, p. 99) that "cross-country time-series analyses are based on aggregate measures, which are less reliable than those at the individual level". Thus, we endeavor to create more reliable estimates from individual (micro) data.

[Table 3 about here]

The micro models without individual fixed effects are estimated with ordered probit. Standard errors are robust to cluster correlation at the regional level. The number of regions in our samples is small and cluster-robust standard errors are potentially biased downwards, as explained in Section 4.2.[34] In order to re-assess inference, we present *p*-values obtained with pairs cluster bootstrap (999 replications) for the micro models in Table 3.

The first micro specification is without year and region fixed effects. Results in Table 3 show that the magnitude of the coefficient for regional GDP per capita (*t*-1) is reduced drastically compared to the macro model, while the signs do not change. The bootstrap *p*-values suggest that significance levels should be adapted in Western and Eastern Germany, while the coefficient in the UK remains highly significant. We now add region fixed effects. The magnitude of the coefficients increases in all three samples. The bootstrap *p*-values suggest that significance levels for the German samples increase compared to the model without region fixed effects, and slightly decrease in the UK sample. For European data, Stevenson and Wolfers (2008, p. 47) show results where the size of the GDP per capita coefficient is reduced by more than two thirds once they introduce country fixed effects.

The next specification is with year fixed effects. We expect the coefficients to change in an unpredictable direction, because the GDP coefficient is then net of the effects of singular events occurring

[33] We only estimate unweighted regressions in this paper under the assumption that we sufficiently control for the determinants of the sampling frame so that $E(u_i|x_i) = 0$. The assumption seems specifically realistic for the individual fixed-effects regressions controlling for all time-invariant characteristics. Such time-invariant characteristics include the SOEP sampling criteria West German, East German, foreigner, and immigrant. The SOEP also contains a high-income sample, which is a time-variant criterion, but this should not cause problems for our main results because we control for household income in most of our specifications with individual fixed-effects.

[34] The Western German sample has 11 regions, the Eastern German sample has 6 regions, and the UK sample has 12 regions. The regions in Germany correspond to the 16 federal states, but Berlin appears in both of the German samples because the SOEP allows differentiating between West and East Berlin.

within a country in a specific year. As it turns out, the coefficients in the German samples lose magnitude and significance. The sign in Western Germany even changes. The coefficient for the UK changes by very little, but significance is reduced somewhat, compared to the model without fixed effects. We now add year and region fixed effects simultaneously and consider this as the most meaningful of our models without individual fixed effects. The coefficients for GDP per capita (t-1) are now positive but statistically insignificant for all three samples. A loss of significance can also be seen for the micro regressions with European data in Stevenson and Wolfers (2008), where significance is reduced from the 1 percent level to the 5 percent level after adding country and year fixed effects.

The result of an insignificant relationship is robust to using regional GDP per capita of the current year (last model in Table 3) . Note that the model with current regional GDP per capita yields coefficients with larger magnitudes for Western Germany and the UK, compared to the model with GDP per capita (t-1). This is contrary to our earlier expectations (see footnote 18).

5.3. Micro estimates with individual fixed effects

In the main part of our analysis, we use models with individual fixed effects due to the models' desirable features as described in Section 1. We use OLS with standard errors adjusted for clustering on region.[35] Results are shown in Tables 4a–c. Note that region is a time-invariant variable in our nested data sets and is automatically controlled for in the fixed-effects models. In order to re-assess inference in our case with a small number of clusters and potentially downward biased cluster-robust standard errors, we present p-values obtained with wild cluster bootstrap (999 replications, null hypothesis imposed, Rademacher weights) for the key regressor GDP per capita.[36]

[Tables 4a–c about here]

The basic individual fixed-effects model without the set of micro control variables (column 1) results in positive and, according to p-values derived from wild bootstrap, insignificant coefficients for regional GDP per capita (t-1) in all three samples. This result is qualitatively identical to what we observed from the previous estimations for the model without individual fixed effects and with region and year fixed effects (see Table 3). After adding a set of micro control variables, we observe in column 2 of Tables 4a–c that coefficients for GDP per capita (t-1) lose magnitude for all three samples. Following the theoretical perspective presented in Section 2, we would conjecture that controlling for household income decreases the effect of GDP per capita on life satisfaction, now net of the individual income effect. Column 3 reveals that coefficients of GDP per capita (t-1) indeed decrease in size when we add

[35] We use the Stata command -xtivreg2- (Schaffer, 2010) because singletons are not included in the estimation, while the standard command -xtreg- includes singletons, which is odd considering the within transformation. We do not show the coefficient for the constant in our tables because it is not reported by -xtivreg2-.

[36] Compared to Mammen weights, Rademacher weights have the advantage that they work for both symmetric and asymmetric distribution of the errors (Davidson and Flachaire, 2008).

household income to the equation, but only marginally. In line with previous studies, the dynamic effect of equivalized net household income on life satisfaction is positive and highly significant for all samples, and sizable especially in the German samples.

The model in column 4 tests current GDP per capita. We observe for all three samples that the magnitude of the coefficients for GDP per capita (*t*) is smaller than the coefficients for GDP per capita (*t*-1). In contrast to the micro models without individual fixed effects, this result is now consistent with our earlier conjecture that using GDP per capita of the previous year is more plausible given the interview periods of the surveys.

In column 5 we introduce a term for average regional income. As noted in Section 4.1., the model in column 5 is problematic from an econometrics perspective. We show column 5 for illustrative purposes and underline that results should be interpreted with caution. If the unlikely case is true that GDP per capita is considered as reference income, we should see some changes in the coefficients. Coefficients for average regional income are negative for all three samples (and probably significant in the UK), while coefficients for GDP per capita (*t*-1) increase compared to column 3.[37] The behavior of the GDP per capita coefficients suggests that the effect of GDP per capita on life satisfaction could include a reference income effect, but this result would have to be re-analyzed with a different econometric methodology and additional data sets, an endeavor which is beyond the scope of this study.

In order to avoid spurious relationships caused by trended variables, we replace levels of GDP per capita and household income with the respective growth rates in columns 6 and 7. Results across the samples are not uniform. In Western Germany, growth of regional GDP per capita has a negative coefficient. For GDP per capita growth from *t*-2 to *t*-1, the negative coefficient is insignificant, and from *t*-1 to *t*, the negative relation is weakly significant according to the *p*-value obtained from the wild cluster bootstrap. Deaton (2008) also finds negative coefficients for GDP growth in the global cross-section using macro data; to him "one of the most surprising results" (p. 61), and certainly contrary to the usual expectations.[38] For Eastern Germany, we see a significant positive relation for GDP per capita growth from *t*-2 to *t*-1 with life satisfaction, and a significant negative relation for economic growth from *t*-1 to *t*. The relation between economic growth and life satisfaction in the UK seems to be insignificant, while the sign of the coefficients is positive. For the above mentioned reasons due to interview periods of the surveys, we have a preference for the results of GDP per capita growth *t*-2 to *t*-1. Surprisingly, the relationship between household income growth and life satisfaction is close to zero and not significant for the UK.

---

[37] We use weights for calculating average regional income, and weights for Northern Irish observations are all zero in the BHPS data set. Therefore, the number of observations in column 5 is smaller than in the previous columns.
[38] Graham (2010) gives some explanations for what she calls the "unhappy growth effect".

For theoretical reasons (see footnote 18), we consider columns 3 and 6 to be our benchmark models. For the sake of brevity, we present robustness checks only for the benchmark models in the following section.

6. Robustness checks

We divided the German sample in order to avoid that results are affected by the German re-unification process. Any analysis from before 1996 might be biased by the unusually great increases in the GDP of Eastern Germany during the turbulent re-unification period (see Fig. 2).

[Table 5 about here]

Column 1 of Table 5 shows results for the benchmark models with a combined sample of Western and Eastern Germany for the period of 1996–2010. Wild cluster bootstrap p-values indicate that both the levels and the growth coefficient for regional GDP per capita are insignificant in the combined sample.

Table 5 also shows robustness checks concerning the effect of restricting the sample so that the data are nested in region, which is necessary for the estimation of one-way clustered standard errors. We do not observe notable differences between estimated coefficients and standard errors from the restricted, nested sample (columns 2, 5, and 8) and estimated coefficients and standard errors from the unrestricted, non-nested sample (columns 3, 6, and 9).[39] The only exception is the growth rate model in Eastern Germany where the coefficient diminishes from .254 in the nested sample without interstate movers to .125 in the non-nested sample with interstate movers, while significance is lost. Note here, that we use standard errors which are robust against clustering in two dimensions (individual over time and region) for estimations with non-nested data.[40] Although there is no methodology available to correct for potential downward bias in the case of two-way clustering if the number of clusters is small, we assume that the coefficient for GDP per capita growth in the non-nested data set with a size of .125 and an uncorrected two-way cluster-robust standard error of .112 is not significant.

We further check robustness of our previous results by using Probit-adapted OLS (POLS) proposed by van Praag and Ferrer-i-Carbonell (2008). Using POLS we acknowledge that our dependent variable is ordinal.[41] Columns 4, 7, and 10 of Table 5 show the results. Note that POLS coefficients need to be interpreted in units of standard deviation of the dependent variable and cannot be readily compared to

[39] Note that region fixed effects are used in all models, either with nested data as part of the individual fixed effects where region is constant over time, or by adding region dummies when we use non-nested data.

[40] The two-way variance estimator, as proposed by Cameron et al. (2011) and Thompson (2011), is implemented in the Stata command -xtivreg2- (Schaffer, 2010).

[41] POLS requires that $u_i = \beta' x_i + \varepsilon_i$ is approximately normally distributed. Other estimation methods based on the fixed-effects ordered logit model do not make this assumption. However, some of these methods dichotomize the dependent variable, which has the disadvantage of losing information. The BUC estimator proposed by Baetschmann et al. (2011) uses all information. We tried to use the BUC estimator, but in some cases the within-variation in our dependent variable was apparently not sufficient. In these cases the estimation did not converge so that we would not be able to report BUC results consistently.

the size of OLS coefficients. We observe that signs of the coefficients do not change in any case. The only difference with respect to the significance level occurs for the growth rate model in the UK sample where the wild cluster bootstrap result suggests that the coefficient is significant at the 5 percent level while our earlier result with OLS had not suggested any significance.

We base our inference mainly on results from the wild bootstrap-*t* procedure, but we are still interested in the behavior of the estimated standard errors. From column 2 onwards we compare conventional (i.i.d.) standard errors, robust standard errors (Huber, 1967; White, 1980), one-way cluster robust standard errors (Liang and Zeger, 1986), and two-way cluster-robust standard errors (Cameron et al., 2011; Thompson, 2011). Our first observation is that robust standard errors are never smaller than conventional standard errors, which otherwise could have been a worrying sign (Angrist and Pischke, 2009, chapter 8.1). The second observation is that conventional and robust standard errors are similar in size, while larger differences could have been a sign for misspecification problems (King and Roberts, 2012). The third observation is that most of the cluster-robust standard errors are larger than the conventional standard errors. This could be a sign that cluster correlation exists indeed or it could be a sign for misspecification in general according to King and Roberts (2012). However, we would not know how to better account for cluster correlation by re-specifying our models. The fourth observation is that some of our clustered standard errors fall below the robust standard errors. We can think of two reasons: intracluster correlation is negative in these cases, and/or downward bias occurs for our clustered standard errors with few clusters. We are not able to finally determine the reasons for the smaller clustered standard errors, because we neither have a reliable measure for the intracluster correlation nor a measure for the downward bias at hand.[42]

In order to avoid that our results are affected by minor irregularities and outliers in the raw data, we also determined if the results of the benchmark models are robust to the exclusion of Berlin in the Western and Eastern German samples, to the exclusion of Hamburg and Bremen in the Western German sample, and to the exclusion of London in the UK sample.[43] Results are shown in Table B.2 of Appendix B. According to *p*-values obtained from the wild bootstrap-*t* procedure, none of the coefficients of GDP per capita is significant in these robustness checks where we exclude the potentially problematic regions.

---

[42] Sribney (2009) explains how negative intracluster correlation can lead to clustered standard errors that are smaller than conventional ones. Given the downward bias of estimates of intracluster correlation coefficients if the number of clusters is small (Feng et al., 2001), we cannot think of a precise measure of intracluster correlation in our data, and we indeed get an estimate of zero for the intracluster correlation coefficient in all cases, which seems odd to us. Rogers (1993) explains that the downward bias of clustered standard errors if the number of clusters is small stems from “mathematical constraints on the residuals” (p. 22).

[43] The calculation of GDP per capita is not entirely satisfactory for East Berlin. The SOEP differentiates between East and West Berlin, but GDP per capita is not available for East Berlin from the German Federal Statistical Office. Hence, we assign GDP per capita for Berlin as a whole to individuals from East and West Berlin. The German city-states Hamburg and Bremen as well London in the UK are outliers in terms of GDP per capita. Here, regional GDP per capita is apparently overestimated because of the large amount of commuters working in these regions.

7. Brief discussion of the results

Our theoretical framework shows three plausible effects of a shock of GDP per capita on subjective well-being (upper branches of Fig. 1). One effect works through individual income. We did not find evidence for adaptation to individual income, and we could not find satisfactory constructs to measure relative individual income effects. What we did find in our data is a robust positive individual income effect (somewhat mitigated in the UK by a surprisingly insignificant result for the model with stationary variables). Controlling for the individual income effect, our main analysis and robustness checks have not produced robust evidence for a significant relationship between GDP per capita and life satisfaction.

For the benchmark models, we conclude from the robustness checks that restricting the data may lead to an overstatement of significance for the coefficient of the growth rate of regional GDP per capita in Eastern Germany. We also get insignificant results for the growth rate in Eastern Germany if we exclude unusual observations (see Appendix C). Concerning the growth rate coefficient in the UK sample, we find evidence that the significance level is understated if the ordinal character of the life satisfaction variable is not taken into account.

We conclude from the finding of an insignificant relation between GDP per capita and life satisfaction that the following assumed, but unmeasured effects cancel each other out: the positive effect of, e.g., political stability, the negative effect of, e.g., pollution, and the relative income effect (see Fig. 1). This interpretation does not hold for the growth rate model in the UK. Future studies might find a way to separately measure the magnitude of these effects, and dig deeper into the mechanics of GDP and subjective well-being.

Even though our results cannot be interpreted causally, we briefly discuss how such coefficients could provide a preliminary indicator for the relative importance of life events. For example, our Western German results would suggest that doubling equivalized household income is ceteris paribus linked with an average rise in life satisfaction of .178 on the 10-point scale (= $.257 \cdot \ln[.257]/\log_2[.257]$).[44] On the other hand, life satisfaction of Western Germans would be reduced on average by .624 points if an individual becomes unemployed. And life satisfaction of Western German individuals would be reduced on average by .528 points if an elderly person in the household requires help. If our significantly positive OLS coefficient for economic *growth* in Eastern Germany would appear in an analysis that allows causal interpretation, it would mean that a doubling of GDP per capita is associated with an average rise in Eastern German life satisfaction of .245 points on the 11-point scale. This result could be put into relation with the fact that it took 36 years until the 1970 value of real GDP per capita had doubled in Germany, while it took 32 years in the UK (World Bank, 2013).

[44] For the specification in column 3 of Tables 4a–c, we show results for all control variables in Table B.3 of Appendix B. Whenever the natural log is applied to a variable, note that the coefficient should be interpreted as the change in the dependent variable if the regressor increases by a factor of approx. 2.7 (Euler's number), rather than by the factor 2 as is often incorrectly stated.

8. Conclusion

The aim of this study was to seek empirical evidence regarding the dynamic relationship between GDP per capita and subjective well-being in Germany and the UK. We do not find evidence for a robustly significant dynamic relationship between regional GDP per capita and life satisfaction in our samples. While some results, especially for the UK, indicate a weakly significant relationship, they do not withstand our robustness checks. On the other hand, we confirm earlier evidence of a significantly positive dynamic relation between individual income and life satisfaction.

What would be the implication of our findings for public policy? Our results support the view that economic growth should be seen as a by-product, not as an end. However, the view of growth as an end still remains a basic tenet of policy making. For example, economic growth is enshrined in German federal law since 1967 (in the "Act to Promote Economic Stability and Growth"), making it a legal obligation of the federal government to pursue steady economic growth. Our results also confirm the finding of Diener et al. (2013) that GDP per capita is not a reliable indicator for average individual income. We therefore strongly suggest that politicians and economists referring to GDP per capita should not use fuzzy synonyms like "income", but should rather refer to more precise terms like "national income".

We propose that future tests of the Easterlin hypothesis should acknowledge some important methodological issues: 1) The use of micro data can make all the difference. Macro data would have shown us very different results for a test of the Easterlin hypothesis. Eventually, micro data and individual fixed effects allow the true analysis of Easterlin's happiness-income paradox, namely the dynamic relationship between national income and subjective well-being, a relationship that should not be confounded by the variation between individuals. 2) We observe that models without trended variables can lead to divergent results. Thus, we propose that unit-root problems should be discussed and avoided. 3) Clustering is a likely issue when analyzing GDP and subjective well-being. Not only should standard errors be adjusted for clustering, one also needs to be aware of the possible downward bias if the number of clusters is small. Ignoring this possibility would overstate the significance of some of our results. We used wild cluster bootstrap as the most promising method to present alternative statistics for inference.

Recent studies focus on a test of the Easterlin hypothesis on a global level and find evidence in some data sets for a significantly positive relationship between economic growth and life satisfaction (e.g., Diener et al., 2013; Sacks et al., 2011). Our study tests the Easterlin hypothesis on the country level. We cannot generally reject the Easterlin hypothesis for Germany and the UK. Together with the evidence presented in Stevenson and Wolfers (2008), we now count three countries for which Easterlin's happiness-income hypothesis cannot be rejected: the United States, Germany, and the United Kingdom.

We conclude that the evidence thus far shows that economic growth may improve the human lot – or it may not.

Acknowledgements

The data used in this publication were made available by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW Berlin) in Berlin, and the Institute for Social & Economic Research at the University of Essex. This article has benefitted from valuable comments from Stefan Bergheim, Jörg Breitung, Ada Ferrer-i-Carbonell, Bruce Headey, Andreas Knabe, Martin Kroh, Christian Müller, and Bert Van Landeghem. We thank participants of the 2011 SOEP Life Satisfaction Workshop at Bremen University, seminar participants at the University of Münster, and participants of the 2011 Market & Happiness Conference in Milan for their helpful input. The following people kindly answered our questions on the treatment of cluster correlation and/or shared their code with us: Tobias Böhm, Mike Campolieti, Adrian Chadi, Joilson Dias, Christina Gathmann, Kevin Lang, Leandro Magnusson, Dan McCaffrey, Doug Miller, Jakob Roland Munch, and Steve Pischke. All remaining errors are ours.

Appendix A

Description of control variables

GDP per capita: Real GDP per capita at the regional level (NUTS 1, Bundesländer), calculated at price levels of 1995 in Euro and obtained from the German Federal Statistical Office. Population weighted averages were calculated for the states Rhineland-Palatinate and Saarland for 1984–1999 since the SOEP did not differentiate between the two states before 2000 due to privacy-related restrictions on the data.

GVA per capita: Real GVA per capita at the regional level (NUTS 1), calculated at price levels of 2005 in GBP and obtained from the UK Office for National Statistics.

Household income: Annual net equivalized real household income calculated at price levels of 1995 in Euros for Germany, and calculated at price levels of 1987 in GBP for the UK. Household income is equivalized according to the modified-OECD scale (De Vos and Zaidi, 1997). Equivalization means here to break household income down to the individual level while attaching lower weight for additional family members due to effects of economies of scale and lower consumption of children. The SOEP variable (*i11102*) refers to household income of the previous year. The BHPS variable (*hhnyrde2*) refers to household income of the 12 months interval up to September 1 of the year of the respective wave.

Age squared: The square of the respondent's age in years.

Marital status: The variables *pgfamstd* (SOEP) and *mlstat* (BHPS) are recoded into the categories *"married"*, *"separated/divorced"*, *"single"*, *"widowed"*. In the SOEP samples, married couples living separately appear in the category *"separated/divorced"*.

Number of children in household: Number of household members aged 0–15 years in the SOEP, generated from the variables *h11103*, *h11104*, *h11105*, *h11106*, *h111107*, and *h11108*. Number of own children in household in the BHPS (*nchild*).

Health satisfaction: Subjective health satisfaction, ranging from 0 *"totally unsatisfied"* to 10 *"totally satisfied"* in the SOEP (*p0080*), and from 1 *"not satisfied at all"* to 7 *"completely satisfied"* in the BHPS (*lfsat1*).

Employment status: The variables *pglfs* (SOEP) and *jbstat* (BHPS) are recoded into the categories *"working"*, *"non-working"*, and *"unemployed"*.

House owner: Dummy variable indicating whether a person owns a home, generated from the variables *hgowner* (SOEP) and *tenure* (BHPS). The variable serves as a proxy for personal wealth.

Person requiring help in household: Dummy variable with the following wording in the SOEP (*h2750*): *"Does someone in your household need care or assistance on a constant basis due to age, sickness or medical treatment?"*. The wording in the BHPS (*aidhh*) is: *"Is there anyone living with you who is sick, disabled or elderly whom you look after or give special help to?"*.

Self-administered interview: The dummy variable indicates whether the interview was executed self-administered in contrast to a face-to-face or telephone interview. The dummy is generated from the SOEP variable *hghmode*. Chadi (2012) and Conti and Pudney (Conti and Pudney, 2011) find that the interview mode can have a significant influence on the answering behavior for satisfaction questions, which is confirmed by our SOEP results. In the BHPS, all respondents answer the life satisfaction question in a self-completion questionnaire.

Appendix B

Please see tables B.1–B.3.

Appendix C

Regression diagnostics

We performed a series of diagnostic tests on our two benchmark models (columns 3 and 6 in Tables 4a–c) in order to check compliance with assumptions of OLS regression and hypothesis testing. Concerning unusual observations, we identified "multivariate outliers" with added-variable plots. In order to exclude that unusual observations influence our results, we re-estimated the benchmark models without the 50 most unusually low and the 50 most unusually high values of the respective GDP per capita variable given all other independent variables. We found that dropping the unusual observations only had a sizeable impact on the coefficient for the GDP per capita growth rate in the Eastern German sample, where the coefficient shrunk from .254 to .105, while the cluster-robust standard error slightly increased from .090 to .105. The normality assumption should not be a problem in large samples such as ours. Residual-versus-fitted plots do not show signs of strong heteroskedasticity, and the cluster-robust standard errors used for inference are also robust to heteroskedasticity. Concerning multicollinearity, we look at variance inflation factors (VIFs). VIFs above 10.0 are conventionally considered to be problematic. The largest VIF we find is 5.19, indicating that multicollinearity is not a major issue in our models. Augmented component-plus-residual plots confirm the approximate linear relationship between life satisfaction and each of the independent variables in the benchmark models. These plots and theoretical considerations corroborate the model specifications we have chosen.

Appendix D

Regional price indices

Due to regional differences, price index data should be used at the smallest geographical level possible. We do not use data at the regional (Bundesland) level because the German Federal Statistical Office (DESTATIS) provides only data from 1995 onwards and does not cover three states (Bremen, Hamburg, Schleswig-Holstein). Thus, we have to take data at the level of Western and Eastern Germany until 1994. For the three states without data points, we would have to take the Western German price index as a proxy. We see a consistency problem here and avoid this. DESTATIS provides price index data starting in 1962 and splits the price indices into Western and Eastern Germany for the years 1991–1999. From 2000 onwards, DESTATIS does not provide separate price indices for Western and Eastern Germany. Using data at the regional level, we can calculate averages for Western and Eastern Germany for subsequent years (only roughly, since three states in Western Germany are missing). These approximated regional averages from 2000–2010 show a maximum difference of .9 percentage points between Western and Eastern Germany. We consider these minor differences from 2000–2010 negligible and use the price index for Germany as a whole from 2000 onwards. The base year for the separate price indices for Western and Eastern Germany is 1995. In the original DESTATIS data we use from 2000 onwards, base year for the national price index is 2005. Thus, we adapt the national price index to the base year 1995 in order to combine both series.

Figures

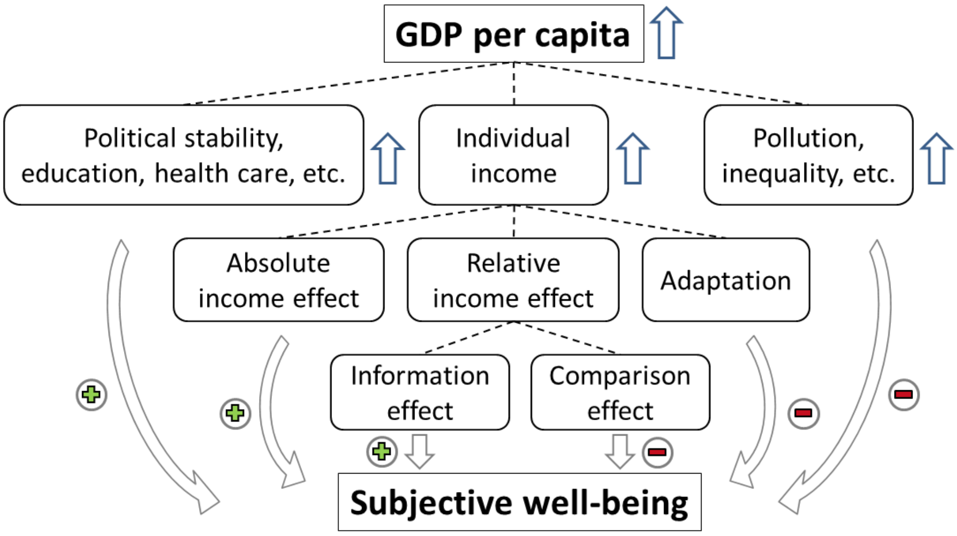


Fig. 1. Positive and negative channel effects of GDP per capita on subjective well-being. Upward arrows indicate a rise in each category. Source: Own depiction.

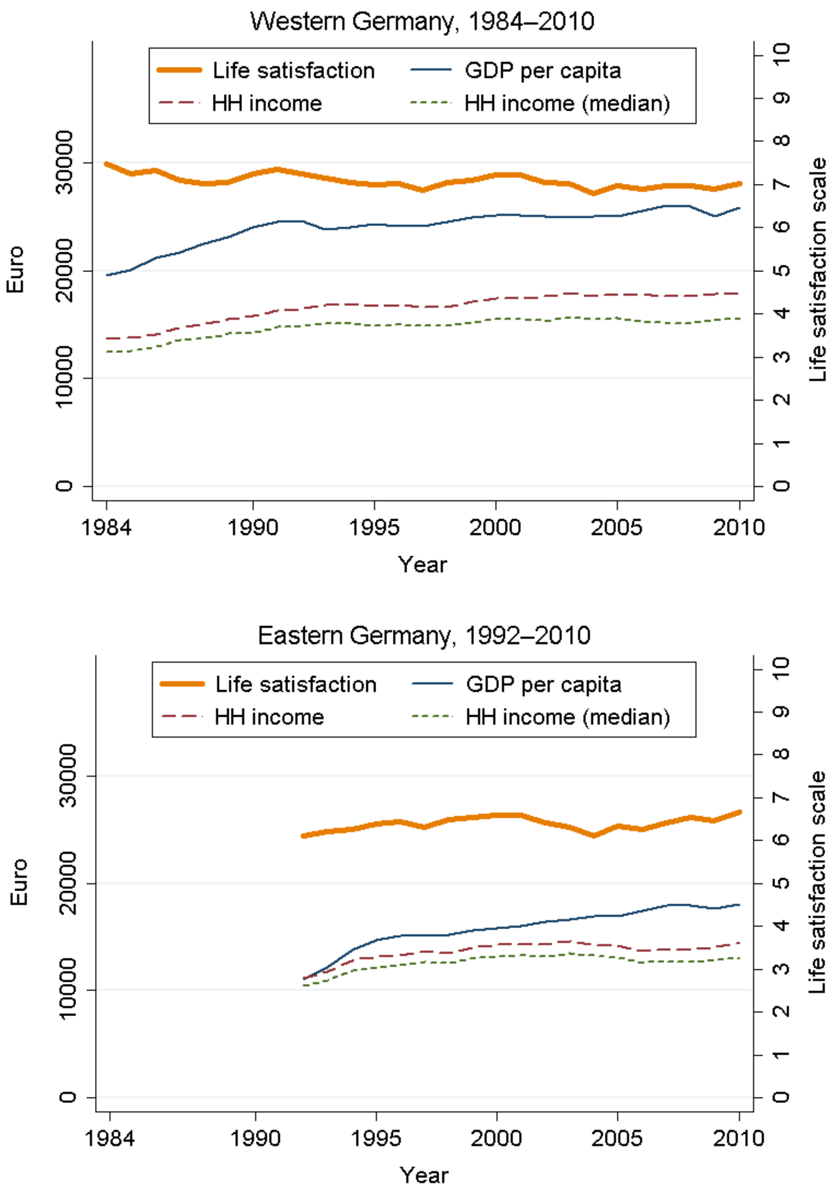
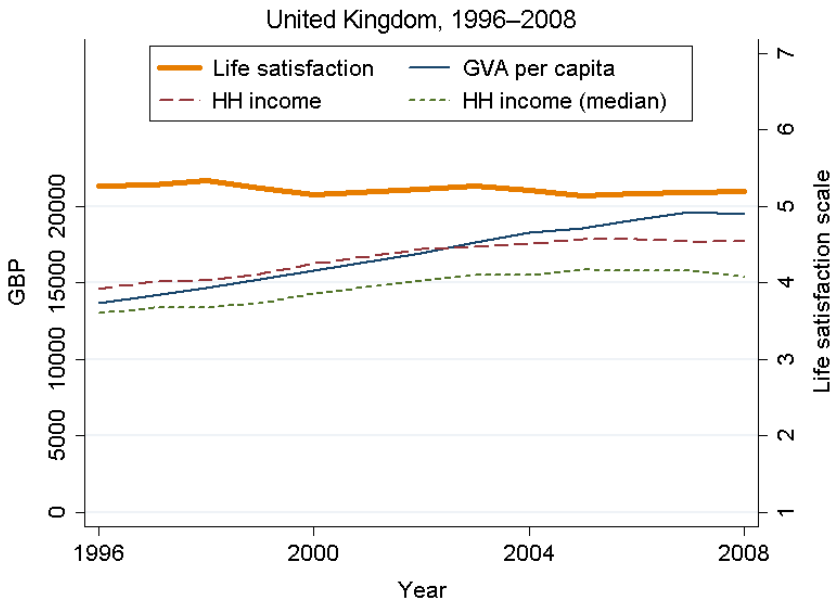


Fig. 2. GDP per capita, household income, and life satisfaction in Western Germany (1984–2010) and Eastern Germany (1992–2010). Western Germany: GDP per capita including East Berlin, household income and life satisfaction excluding East Berlin. Eastern Germany: GDP per capita, household income, and life satisfaction including East Berlin. Household income and life satisfaction are weighted cross-sectional averages. GDP per capita and household income are in real terms (price levels of 1995). Household income is annual, net and equivalized. Data source: German Federal Statistical Office and SOEP (2011).



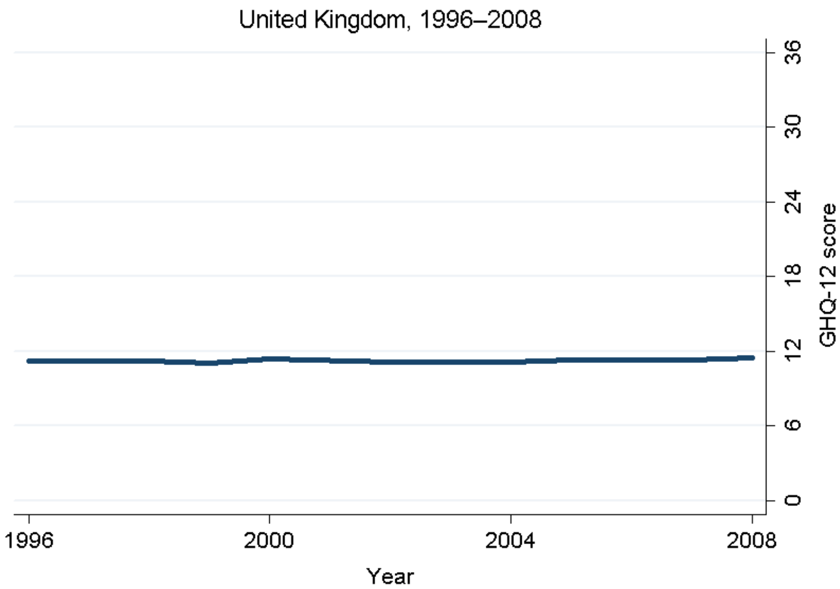


Fig. 3. GVA per capita, household income, life satisfaction, and the GHQ-12 score in the United Kingdom (1996–2008). Data on life satisfaction are missing for 2001. BHPS data on Northern Ireland are only available from 2002. Household income, life satisfaction, and the GHQ-12 score are weighted cross-sectional averages. GVA per capita and household income are in real terms (price levels of 2005 and 1987, respectively). Household income is annual, net and equivalized. Data source: UK Office for National Statistics and BHPS (2012).

Tables

Table 1: Studies with regressions of subjective well-being on GDP per capita

Study	Data source	Time period (years/waves)	Count-ries	Obser-vations	Well-being variable (scale points)	ln(Real GDP per capita)	Controlling for individual income	Country and/or year/wave dummies	Other control variables	Method
Static relation: average subjective well-being (macro level)										
Deaton (2008)	Gallup World Poll	2006 (1/1)	123	123	Cantril's ladder (11)	0.838*** (0.051)	no	no	no	*not stated*
Sacks et al. (2010)	Gallup World Poll	2006 (1/1)	131	131	Cantril's ladder (11)	0.342*** (0.019)	no	no	no	OLS
	Pew Global Attitudes Survey	2002 (1/1)	44	44	Cantril's ladder (11)	0.204*** (0.037)	no	no	no	OLS
Static relation: individual subjective well-being (micro level)										
Diener et al. (2010)	Gallup World Poll	2006 (1/1)	132	136,839	Cantril's ladder (11)	1.01*** (0.028)	yes	no	5	*not stated*
Sacks et al. (2010)	Gallup World Poll	2006 (1/1)	131	291,383	Cantril's ladder (11)	0.378*** (0.019)	no	no	4	OLS
	Pew Global Attitudes Survey	2002 (1/1)	44	37,974	Cantril's ladder (11)	0.204*** (0.037)	no	no	4	OLS
Dynamic relation: average subjective well-being (macro level)										
Hagerty (2000)	World Database of Happiness	1972–1994 (23/?)	8	61	Life satisfaction (4)	0.061*** (0.013) [a)]	no	yes	1	OLS
Easterlin (2005b)	General Social Survey	1972–2002 (31/24)	1	24	Happiness (3)	$-0.11e^{-5}$ ($0.13e^{-5}$) [a)]	no	no	no	OLS
Easterlin (2005a)	World Database of Happiness	1958–1987 (30/?)	1	29	Life satisfaction (4)	0.0692 (0.0633)	no	no	no	OLS
Inglehart et al. (2008)	World Values Survey	1981–2007 (27/5)	41	41	Life satisfaction (10)	0.045** (0.019) [a)]	no	no	3	OLS
					Happiness (4)	0.005 (0.005) [a)]	no	no	3	OLS
Sacks et al. (2011)	World Values Survey	1981–2008 (28/)	63	195	Life satisfaction (10)	0.54 (0.10) [b)]	no	yes	no	OLS
					Happiness (4)	0.16 (0.12) [b)]	no	yes	no	OLS
	Eurobarometer	1973–2009 (37/35)	33	994	Life satisfaction (4)	0.17 (0.05) [b)]	no	yes	no	OLS
	ISSP	1991–2008 (18/5)	39	144	Happiness (4)	0.55 (0.20) [b)]	no	yes	no	OLS
	Gallup World Poll	2005–2011 (7/7)	141	591	Cantril's ladder (11)	0.36 (0.21) [b)]	no	yes	no	OLS
	Pew Global Attitudes Survey	2002–2010 (9/3)	39	66	Cantril's ladder (11)	0.56 (0.38) [b)]	no	yes	no	OLS
	Latinobarometro	2001–2007 (7/7)	17	73	Life satisfaction (4)	0.27 (0.61) [b)]	no	yes	no	OLS
	"Panel of Panels"	1973–2011 (39/?)	159	2124	Standardized	0.33 (0.09) [b)]	no	yes	no	OLS
Dynamic relation: individual subjective well-being (micro level)										
Di Tella et al. (2003)	Eurobarometer	1975–1992 (18/18)	12	271,224	Life satisfaction (4)	1.094** (0.335) [c)]	yes	yes	17	Ordered probit
						1.220 (0.763) [c)]	yes	yes	19	Ordered probit
Di Tella & MacCulloch (2008)	Eurobarometer	1975–1997 (23/22)	12	344,294	Life satisfaction (4)	0.539** (0.235) [d)]	yes [g)]	yes	27	Ordered probit
Stevenson & Wolfers (2008)	Eurobarometer	1973–2007 (35/33)	31	850,153	Life satisfaction (4)	0.737*** (0.181)	no	no	no	Ordered probit
						0.192*** (0.066)	no	yes	no	Ordered probit
						0.208** (0.099)	no	yes	no	Ordered probit
Di Tella & MacCulloch (2010)	Eurobarometer	1975–2002 (28/27)	16	605,020	Life satisfaction (4)	0.65** (n.a.) [e) f)]	yes	yes	7	Ordered probit
						1.28** (0.40) [e)]	yes	yes	12	Ordered probit
Sacks et al. (2010)	World Values Survey	1980–2004 (25/4)	79	234,093	Life satisfaction (10)	0.364*** (0.034)	no	yes	4	Ordered probit

Notes: Standard errors are given in parentheses. Asterisks denote statistical significance lower than or equal to the * 10 percent, ** 5 percent, and *** 1 percent level. a) GDP per capita not used as log. b) Significance level not given in study. c) GDP per capita not used as log and scaled by a factor of 10,000. d) GDP per capita scaled by a factor of 1,000. e) GDP per capita scaled by a factor of 2,000. f) Standard error not correct in publication (personal communication with authors). g) Personal income position (logarithm of individual income relative to mean income).

Table 2: Summary statistics for key variables

Variable	Sample	Observations	Mean	Median	Std.dev.	Min.	Max.
Life satisfaction (0–10)	Western Germany (1984–2010)	311,307	7.08	7	1.85	0	10
	Eastern Germany (1992–2010)	81,127	6.39	7	1.83	0	10
Life satisfaction (1–7)	United Kingdom (1996–2008)	78,687	5.22	5	1.27	1	7
Regional GDP per capita (*t*-1)	Western Germany (1984–2010)	281	24,941	23,361	5,420	16,108	40,924
	Eastern Germany (1992–2010)	114	16,391	16,283	3,217	8,775	24,407
Regional GVA per capita (*t*-1)	United Kingdom (1996–2008)	139	15,443	14,761	3,527	10,648	29,585
Equivalized net household income	Western Germany (1984–2010)	311,307	16,648	14,740	11,610	4,148	2,448,534
	Eastern Germany (1992–2010)	81,127	13,630	12,558	6,455	4,157	324,689
	United Kingdom (1996–2008)	78,687	16,586	14,620	10,245	1,942	315,086

Notes: Data on life satisfaction and household income are weighted. GDP per capita and household income in Germany are in price levels of 1995 (Euro). GVA per capita and household income in the UK are in price levels of 2005 and 1987, respectively (GBP). UK data are without 2001 because life satisfaction is missing in the BHPS. GDP data are from the German Federal Statistical Office. GVA data are from the UK Office for National Statistics. Other data are from SOEP (2011) and BHPS (2012).

Table 3: Regressions of life satisfaction on GDP per capita without individual fixed effects

	Western Germany (1984–2010)		Eastern Germany (1992–2010)		United Kingdom (1996–2008)	
Dependent variable: life satisfaction	Macro data (OLS)	Micro data (Ordered probit)	Macro data (OLS)	Micro data (Ordered probit)	Macro data (OLS)	Micro data (Ordered probit)
ln(GDP per capita, t-1)	-1.129***	-0.181**	0.521***	0.358***	-0.288***	-0.166***
	(0.265)	[0.091]	(0.120)	[0.092]	(0.081)	[0.046]
		p = 0.090		p = 0.038		p = 0.002
ln(GDP per capita, t-1) with region dummies		-0.517***		0.422***		-0.214***
		[0.130]		[0.072]		[0.062]
		p = 0.042		p = 0.004		p = 0.056
ln(GDP per capita, t-1) with wave dummies		0.004		0.134*		-0.159**
		[0.074]		[0.079]		[0.070]
		p = 0.958		p = 0.448		p = 0.060
ln(GDP per capita, t-1) with region and wave dummies		0.081		0.015		0.249
		[0.210]		[0.057]		[0.330]
		p = 0.752		p = 0.754		p = 0.476
ln(GDP per capita, t) with region and wave dummies		0.084		-0.008		0.426
		[0.215]		[0.065]		[0.274]
		p = 0.732		p = 0.952		p = 0.220
Further controls	No	No	No	No	No	No
Adj. R^2	0.328	—	0.250	—	0.398	—
Observations	27	318,346	19	81,956	12	125,095

Notes: Life satisfaction is on a 0–10 scale for Germany, and on a 1–7 scale for the UK. Asterisks denote statistical significance lower than or equal to the * 10 percent, ** 5 percent, and *** 1 percent level. Robust standard errors are in parentheses. The standard errors in brackets are adjusted for regional clustering. The reported p-values stem from the pairs cluster bootstrap-t procedure (999 replications). Cutoffs from ordered probit regression are not reported. Macro regressions are with weighted yearly averages of life satisfaction and with GDP per capita on the national level. Micro regressions are with GDP per capita on the regional level. Instead of GDP per capita, GVA per capita is used for the UK data. GDP and GVA per capita are used in real terms. See notes of Table 2 for further details and data sources.

Table 4a: OLS regressions of life satisfaction on regional GDP per capita with individual fixed effects, Western Germany, 1984–2010

Dependent variable: life satisfaction	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Regional GDP per capita, t-1)	0.262	0.141	0.125		0.177		
	(0.258)	(0.211)	(0.207)		(0.232)		
	p = 0.451	p = 0.551	p = 0.577		p = 0.501		
ln(Regional GDP per capita, t)				0.067			
				(0.217)			
				p = 0.789			
ln(Household income)			0.257***	0.257***	0.257***		
			(0.012)	(0.012)	(0.012)		
ln(Avg. regional household income)					-0.185		
					(0.177)		
Regional GDP per capita growth (t-2, t-1)						-0.305	
						(0.282)	
						p = 0.643	
Regional GDP per capita growth (t-1, t)							-0.258*
							(0.143)
							p = 0.091
Household income growth (t-1, t)						0.059***	0.059***
						(0.006)	(0.006)
Wave dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further controls	No	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.03	0.14	0.14	0.14	0.14	0.14	0.14
Individuals	33,426	33,354	33,354	33,354	33,354	28,588	28,588
Observations	318,346	316,896	316,896	316,896	316,896	278,190	278,190

Notes: Asterisks denote statistical significance lower than or equal to the * 10 percent, ** 5 percent, and *** 1 percent level. The standard errors in parentheses are adjusted for regional clustering. The reported p-values stem from the wild cluster bootstrap-t procedure (999 replications, Rademacher weights, null hypothesis imposed). GDP per capita and household income are used in real terms. Household income is net and equivalized. Data are nested in region, and region is a time-invariant variable controlled for in the individual fixed-effects regressions. Average regional household income is weighted. Further control variables are age squared, marital status, number of children in household, health satisfaction, employment status, house ownership, person requiring help in household, and self-administered interview (see Appendix A). Data in columns 6 and 7 are from 1985–2010 because household income growth cannot be calculated for 1984. See Table 2 for further details and data sources.

Table 4b: OLS regressions of life satisfaction on regional GDP per capita with individual fixed effects, Eastern Germany, 1992–2010

Dependent variable: life satisfaction	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Regional GDP per capita, t-1)	0.153**	0.093**	0.082**		0.126**		
	(0.063)	(0.041)	(0.041)		(0.057)		
	p = 0.266	p = 0.230	p = 0.276		p = 0.104		
ln(Regional GDP per capita, t)				0.013			
				(0.079)			
				p = 0.946			
ln(Household income)			0.290***	0.290***	0.290***		
			(0.015)	(0.015)	(0.016)		
ln(Avg. regional household income)					-0.165		
					(0.199)		
Regional GDP per capita growth (t-2, t-1)						0.254***	
						(0.090)	
						p = 0.046	
Regional GDP per capita growth (t-1, t)							-0.706*
							(0.365)
							p = 0.014
Household income growth (t-1, t)						0.137***	0.136***
						(0.015)	(0.015)
Wave dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further controls	No	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.01	0.10	0.10	0.10	0.10	0.10	0.10
Individuals	8,766	8,754	8,754	8,754	8,754	7,720	7,720
Observations	81,956	81,631	81,631	81,631	81,631	71,852	71,852

Notes: Asterisks denote statistical significance lower than or equal to the * 10 percent, ** 5 percent, and *** 1 percent level. The standard errors in parentheses are adjusted for regional clustering. The reported p-values stem from the wild cluster bootstrap-t procedure (999 replications, Rademacher weights, null hypothesis imposed). Data in columns 6 and 7 are from 1993–2010 because household income growth cannot be calculated for 1992. See notes of Table 4a for further details and notes of Table 2 for data sources.

Table 4c: OLS regressions of life satisfaction on regional GVA per capita with individual fixed effects, United Kingdom, 1996–2008

Dependent variable: life satisfaction	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(Regional GVA per capita, t-1)	0.398**	0.230	0.228		0.257		
	(0.180)	(0.158)	(0.159)		(0.157)		
	p = 0.224	p = 0.300	p = 0.298		p = 0.268		
ln(Regional GVA per capita, t)				0.200			
				(0.177)			
				p = 0.416			
ln(Household income)			0.039***	0.039***	0.048***		
			(0.009)	(0.009)	(0.006)		
ln(Avg. regional household income)					-0.161**		
					(0.066)		
Regional GVA per capita growth (t-2, t-1)						0.682*	
						(0.365)	
						p = 0.114	
Regional GVA per capita growth (t-1, t)							0.278
							(0.552)
							p = 0.684
Household income growth (t-1, t)						0.001	0.001
						(0.004)	(0.004)
Wave dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Further controls	No	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.00	0.11	0.11	0.11	0.11	0.11	0.11
Individuals	19,095	18,869	18,869	18,869	16,492	16,897	16,897
Observations	125,095	122,012	122,012	122,012	110,242	107,342	107,342

Notes: Asterisks denote statistical significance lower than or equal to the * 10 percent, ** 5 percent, and *** 1 percent level. The standard errors in parentheses are adjusted for regional clustering. The reported p-values stem from the wild cluster bootstrap-t procedure (999 replications, Rademacher weights, null hypothesis imposed). GVA per capita and household income are used in real terms. Household income is net and equivalized. Further control variables are age squared, marital status, number of children in household, health satisfaction, employment status, house ownership, and person requiring help in household (see Appendix A). See Table 2 for further details and data sources.

Table 5: Robustness checks for benchmark models

	Western and Eastern Germany, 1996–2010	Western Germany, 1984–2010			Eastern Germany, 1992–2010			United Kingdom, 1996–2008		
	OLS	OLS		POLS	OLS		POLS	OLS		POLS
	Without interstate movers (nested)	Without interstate movers (nested)	With interstate movers (non-nested)	Without interstate movers (nested)	Without interstate movers (nested)	With interstate movers (non-nested)	Without interstate movers (nested)	Without interstate movers (nested)	With interstate movers (non-nested)	Without interstate movers (nested)
Dependent variable: life satisfaction	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Specification of column 3 in Tables 4a–c										
ln(Regional GDP per capita, t-1)	0.344	0.125	0.087	0.061	0.082	0.092	0.061	0.228	0.273	0.178
	(0.107)	(0.103)	(0.100)	(0.056)	(0.111)	(0.107)	(0.061)	(0.158)	(0.155)	(0.117)
	[0.114]	[0.118]	[0.115]	[0.063]	[0.127]	[0.122]	[0.069]	[0.161]	[0.157]	[0.119]
	{0.299}	{0.207}	<0.183>	{0.118}	{0.041}	<0.059>	{0.025}	{0.159}	<0.152>	{0.125}
	p = 0.308	p = 0.577		p = 0.657	p = 0.276		p = 0.326	p = 0.298		p = 0.328
Region dummies	—	—	Yes	—	—	Yes	—	—	Yes	—
Clusters	16	11	11	11	6	6	6	12	12	12
Observations	272,244	316,896	324,066	316,896	81,631	83,563	81,631	122,012	124,896	122,012
Specification of column 6 in Tables 4a–c										
Regional GDP per capita growth (t-2, t-1)	-0.114	-0.305	-0.312	-0.172	0.254	0.125	0.133	0.682	0.719	0.570
	(0.227)	(0.217)	(0.214)	(0.117)	(0.309)	(0.303)	(0.169)	(0.439)	(0.432)	(0.326)
	[0.229]	[0.235]	[0.231]	[0.123]	[0.334]	[0.327]	[0.181]	[0.449]	[0.441]	[0.330]
	{0.352}	{0.282}	<0.272>	{0.162}	{0.090}	<0.112>	{0.063}	{0.365}	<0.361>	{0.240}
	p = 0.768	p = 0.643		p = 0.643	p = 0.046		p = 0.046	p = 0.114		p = 0.048
Region dummies	—	—	Yes	—	—	Yes	—	—	Yes	—
Clusters	16	11	11	11	6	6	6	12	12	12
Observations	244,458	278,190	284,388	278,190	71,852	73,410	71,852	107,342	109,855	107,342

Notes: Conventional standard errors are reported in parentheses. Heteroskedasticity-robust standard errors using formulas in White (1980) are reported in brackets. The standard errors in braces are adjusted for regional clustering using formulas in Liang and Zeger (1986). The standard errors in angle brackets are adjusted for two-way clustering (individual over time and region) using formulas in Cameron et al. (2011) and Thompson (2011). The reported p-values stem from the wild cluster bootstrap-t procedure (999 replications, Rademacher weights, null hypothesis imposed). All regressions are estimated with individual fixed effects. All models are with further control variables. Life satisfaction is on a 0–10 scale for Germany, and on a 1–7 scale for the UK. See notes of Tables 4a–c for further details.

Table B.1a: Summary statistics, Western Germany, 1984–2010

Variable	Observations	Mean	Std.dev.	Min.	Max.
Life satisfaction	N = 318,346	7.156	1.798	0	10
Between	n = 33,426		1.351	0	10
Within	$\bar{T}$ = 9.5		1.297	-1.997	14.809
Regional GDP per capita (t-1)	N = 281	24,941	5,420	16,108	40,924
Between	n = 11		5,046	20,177	37,202
Within	$\bar{T}$ = 25.5		2,265	16,495	29,118
ln(Regional GDP per capita, t-1)	N = 281	10.103	0.200	9.687	10.619
Between	n = 11		0.184	9.909	10.519
Within	$\bar{T}$ = 25.5		0.089	9.806	10.257
Regional GDP per capita (t)	N = 281	25,198	5,419	16,656	40,924
Between	n = 11		5,138	20,384	37,617
Within	$\bar{T}$ = 25.5		2,054	17,759	29,024
ln(Regional GDP per capita, t)	N = 281	10.114	0.198	9.721	10.619
Between	n = 11		0.185	9.920	10.531
Within	$\bar{T}$ = 25.5		0.080	9.838	10.267
Regional GDP per capita growth (t-2, t-1)	N = 281	0.010	0.025	-0.196	0.069
Between	n = 11		0.004	0.003	0.016
Within	$\bar{T}$ = 25.5		0.025	-0.189	0.069
Regional GDP per capita growth (t-1, t)	N = 281	0.011	0.026	-0.196	0.069
Between	n = 11		0.004	0.003	0.016
Within	$\bar{T}$ = 25.5		0.025	-0.188	0.069
Household income	N = 318,346	17,675	15,876	4,148	2,448,534
Between	n = 33,426		15,123	4,202	1,323,781
Within	$\bar{T}$ = 9.5		9,152	-1,210,793	1,324,869
ln(Household income)	N = 318,346	9.649	0.483	8.33	14.711
Between	n = 33,426		0.454	8.343	13.484
Within	$\bar{T}$ = 9.5		0.247	6.984	12.273
Household income growth (t-1, t)	N = 284,191	0.046	0.334	-0.981	23.655
Between	n = 33,391		0.198	-0.876	7.870
Within	$\bar{T}$ = 8.5		0.315	-7.629	20.915

Table B.1b: Summary statistics, Eastern Germany, 1992–2010

Variable	Observations	Mean	Std.dev.	Min.	Max.
Life satisfaction	N = 81,956	6.461	1.774	0	10
Between	n = 8,766		1.370	0	10
Within	$\bar{T}$ = 9.3		1.228	-1.094	13.184
Regional GDP per capita (t-1)	N = 114	16,391	3,217	8,775	24,407
Between	n = 6		2,674	15,012	21,816
Within	$\bar{T}$ = 19.0		2,082	10,153	19,295
ln(Regional GDP per capita, t-1)	N = 114	9.685	0.204	9.080	10.103
Between	n = 6		0.151	9.602	9.989
Within	$\bar{T}$ = 19.0		0.150	9.163	9.876
Regional GDP per capita (t)	N = 114	16,733	2,837	10,611	23,581
Between	n = 6		2,448	15,488	21,696
Within	$\bar{T}$ = 19.0		1,735	11,722	19,193
ln(Regional GDP per capita, t)	N = 114	9.711	0.171	9.270	10.068
Between	n = 6		0.135	9.639	9.984
Within	$\bar{T}$ = 19.0		0.118	9.341	9.867
Regional GDP per capita growth (t-2, t-1)	N = 108	0.028	0.045	-0.035	0.209
Between	n = 6		0.017	-0.006	0.040
Within	$\bar{T}$ = 18.0		0.043	-0.041	0.197
Regional GDP per capita growth (t-1, t)	N = 114	0.028	0.044	-0.035	0.209
Between	n = 6		0.016	-0.005	0.040
Within	$\bar{T}$ = 19.0		0.042	-0.041	0.197
Household income	N = 81,956	14,603	7,698	4,152	324,689
Between	n = 8,766		7,153	4,179	180,648
Within	$\bar{T}$ = 9.3		4,378	-53,216	277,195
ln(Household income)	N = 81,956	9.500	0.405	8.331	12.691
Between	n = 8,766		0.377	8.338	12.094
Within	$\bar{T}$ = 9.3		0.215	7.817	12.090
Household income growth (t-1, t)	N = 73,162	0.040	0.265	-0.933	6.215
Between	n = 8,758		0.145	-0.810	3.827
Within	$\bar{T}$ = 8.4		0.253	-2.628	5.810

Table B.1c: Summary statistics, United Kingdom, 1996–2008

Variable	Observations	Mean	Std.dev.	Min.	Max.
Life satisfaction	N = 125,095	5.225	1.291	1	7
Between	n = 19,095		1.041	1	7
Within	$\bar{T}$ = 6.6		0.823	-0.275	10.368
Regional GVA per capita (t-1)	N = 139	15,443	3,527	10,648	29,585
Between	n = 12		2,962	12,442	23,189
Within	$\bar{T}$ = 11.6		2,028	9,578	21,839
ln(Regional GVA per capita, t-1)	N = 139	9.623	0.206	9.273	10.295
Between	n = 12		0.170	9.424	10.037
Within	$\bar{T}$ = 11.6		0.123	9.346	9.881
Regional GVA per capita (t)	N = 139	15,858	3,637	10,856	29,706
Between	n = 12		3,150	12,683	24,166
Within	$\bar{T}$ = 11.6		1,953	9,823	21,399
ln(Regional GVA per capita, t)	N = 139	9.649	0.206	9.292	10.299
Between	n = 12		0.175	9.443	10.079
Within	$\bar{T}$ = 11.6		0.115	9.375	9.869
Regional GVA per capita growth (t-2, t-1)	N = 139	0.031	0.012	0.002	0.060
Between	n = 12		0.005	0.025	0.044
Within	$\bar{T}$ = 11.6		0.011	-0.005	0.055
Regional GVA per capita growth (t-1, t)	N = 139	0.027	0.017	-0.023	0.060
Between	n = 12		0.006	0.020	0.043
Within	$\bar{T}$ = 11.6		0.016	-0.023	0.055
Household income	N = 125,095	16,270	10,332	1,940	529,787
Between	n = 19,095		8,387	2,536	176,245
Within	$\bar{T}$ = 6.6		6,336	-148,451	448,308
ln(Household income)	N = 125,095	9.557	0.527	7.571	13.180
Between	n = 19,095		0.462	7.825	11.883
Within	$\bar{T}$ = 6.6		0.286	7.092	12.762
Household income growth (t-1, t)	N = 111,722	0.095	0.666	-0.990	86.918
Between	n = 18,988		0.477	-0.966	42.969
Within	$\bar{T}$ = 5.9		0.589	-43.854	44.044

Table B.2: Fixed-effects OLS regressions excluding regions with statistical irregularities

	Western Germany without East Berlin (1984–2010)	Eastern Germany without East Berlin (1992–2010)	Western Germany without Hamburg and Bremen (1984–2010)	United Kingdom without London (1996–2008)
Dependent variable: life satisfaction	(1)	(2)	(3)	(4)
Specification of column 3 in Tables 4a–c				
ln(Regional GDP per capita, t-1)	0.034	0.004	0.101	0.186
	(0.346)	(0.518)	(0.217)	(0.259)
	$p = 0.954$	$p = 0.979$	$p = 0.724$	$p = 0.687$
Observations	307,782	76,319	308,356	115,468
Specification of column 6 in Tables 4a–c				
Regional GDP per capita growth (t-2, t-1)	0.103	0.300	-0.256	0.602
	(0.370)	(0.249)	(0.305)	(0.440)
	$p = 0.862$	$p = 0.313$	$p = 0.706$	$p = 0.235$
Observations	270,159	67,246	270,806	101,429

Notes: The standard errors in parentheses are adjusted for regional clustering. The reported p-values stem from the wild cluster bootstrap-t procedure (999 replications, Rademacher weights, null hypothesis imposed). Life satisfaction is on a 0–10 scale for Germany, and on a 1–7 scale for the UK. See notes of Tables 4a–c for further details.

Table B.3: Fixed-effects OLS regressions showing all control variables

Dependent variable: life satisfaction	Western Germany	Eastern Germany	United Kingdom
ln(Regional GDP per capita, t-1)	0.125	0.082**	0.228
	(0.207)	(0.041)	(0.159)
ln(Household income)	0.257***	0.290***	0.039***
	(0.012)	(0.015)	(0.009)
Age squared/1000	-0.015	-0.072	-0.001
	(0.028)	(0.060)	(0.024)
Marital status (Base category: Married)			
Separated/divorced	-0.241***	-0.074*	-0.095***
	(0.023)	(0.043)	(0.023)
Single	-0.138***	-0.025	-0.052**
	(0.034)	(0.033)	(0.024)
Widowed	-0.433***	-0.145***	-0.320***
	(0.045)	(0.037)	(0.034)
Number of children in household	0.010**	0.072***	-0.004
	(0.005)	(0.015)	(0.009)
Health satisfaction	0.266***	0.229***	0.259***
	(0.004)	(0.004)	(0.003)
Employment status (Base category: Working)			
Non-working	-0.043***	-0.102***	-0.006
	(0.014)	(0.019)	(0.018)
Unemployed	-0.624***	-0.628***	-0.248***
	(0.024)	(0.018)	(0.025)
House owner	0.038*	0.034	0.016
	(0.020)	(0.025)	(0.011)
Person requiring help in household	-0.528***	-0.509***	-0.089***
	(0.040)	(0.094)	(0.017)
Self-administered interview	-0.186***	-0.056**	
	(0.010)	(0.022)	
R^2	0.14	0.10	0.11
Individuals	33,354	8,754	18,869
Observations	316,896	81,631	122,012

Notes: Life satisfaction is on a 0–10 scale for Germany, and on a 1–7 scale for the UK. Asterisks denote statistical significance lower than or equal to the * 10 percent, ** 5 percent, and *** 1 percent level. The standard errors in parentheses are adjusted for regional clustering. Instead of GDP per capita, GVA per capita is used for the UK data. GDP per capita, GVA per capita and household income are used in real terms. See notes of Table 2 for further details.

References

Angrist, J.D., Pischke, J.-S., 2009, Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton and Oxford.

Antonakis, J., Bendahan, S., Jacquart, P., Lalive, R., 2010. On Making Causal Claims: A Review and Recommendations. Leadership Quarterly Yearly Review, 21(6), 1086-1120.

Arrow, K.J., Dasgupta, P.S., 2009. Conspicuous Consumption, Inconspicuous Leisure. The Economic Journal, 119(541), F497-F516.

Baetschmann, G., Staub, K.E., Winkelmann, R., 2011, Consistent Estimation of the Fixed Effects Ordered Logit Model. IZA Discussion Paper No. 5443.

Bell, R.M., McCaffrey, D.F., 2002. Bias Reduction in Standard Errors for Linear Regression with Multi-stage Samples. Survey Methodology, 28(2), 169-181.

BHPS, 2012, Britisch Household Panel Survey. Institute for Social & Economic Research, University of Essex.

Brickman, P., Campbell, D.T., 1971, Hedonic Relativism and Planning the Good Society. In: M.H. Appley (Ed.). Adaptation-level Theory: A Symposium. Academic Press, New York, pp. 287-302.

Cameron, C.A., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. The Review of Economics and Statistics, 90(3), 414-427.

Cameron, C.A., Gelbach, J.B., Miller, D.L., 2011. Robust lnference With Multiway Clustering. Journal of Business & Economics Statistics, 29(2), 238-249.

Cameron, C.A., Miller, D.L., 2011, Robust Inference with Clustered Data. In: A. Ullah, D.E.A. Giles (Eds.). Handbook of Empirical Economics and Finance. Taylor & Francis, Boca Raton, pp. 1-28.

Cameron, C.A., Trivedi, P.K., 2010, Microeconometrics Using Stata. Stata Press, College Station.

Chadi, A., 2012. I Would Really Love to Participate in Your Survey! Bias Problems in the Measurement of Well-being. Economics Bulletin, 32(4), 3111-3119.

Clark, A.E., Diener, E., Georgellis, Y., Lucas, R.E., 2008a. Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis. The Economic Journal, 118(529), F222-F243.

Clark, A.E., Etilé, F., Postel-Vinay, F., Senik, C., Van der Straeten, K., 2005. Heterogeneity in Reported Well-Being: Evidence from Twelve European Countries. The Economic Journal, 115(502), C118-C132.

Clark, A.E., Frijters, P., Shields, M.A., 2008b. Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles. Journal of Economic Literature, 46(1), 95-144.

Clark, A.E., Senik, C., 2010a. Who Compares to Whom? The Anatomy of Income Comparisons in Europe. The Economic Journal, 120(544), 573-594.

Clark, A.E., Senik, C., 2010b, Will GDP Growth Increase Subjective Happiness in Developing Countries? 8th AFD-EUDN Conference. Measure for Measure: How Well Do We Measure Development? AFD Publications, Paris, pp. 97-176.

Conti, G., Pudney, S., 2011. Survey Design and the Analysis of Satisfaction. The Review of Economics and Statistics, 93(3), 1087-1093.

Davidson, R., Flachaire, E., 2008. The Wild Bootstrap, Tamed at Last. Journal of Econometrics, 146(1), 162-169.

De Vos, K., Zaidi, M.A., 1997. Equivalence Scale Sensitivity of Poverty Statistics for the Member States of the European Community. Review of Income and Wealth, 43(3), 319-333.

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.

Di Tella, R., Haisken-DeNew, J.P., MacCulloch, R., 2010. Happiness Adaptation to Income and to Status in an Individual Panel. Journal of Economic Behavior & Organization, 76(3), 834-852.

Di Tella, R., MacCulloch, R., 2008. Gross National Happiness as an Answer to the Easterlin Paradox? Journal of Development Economics, 86(1), 22-42.

Di Tella, R., MacCulloch, R., 2010, Happiness Adaptation To Income Beyond "Basic Needs". In: E. Diener, J.F. Helliwell, D. Kahneman (Eds.). International Differences in Well-Being. Oxford University Press, Oxford, New York, pp. 217-246.

Di Tella, R., MacCulloch, R.J., Oswald, A.J., 2003. The Macroeconomics of Happiness. The Review of Economics and Statistics, 85(4), 809-827.

Diener, E., Lucas, R.E., 1999, Personality and Subjective Well-Being. In: D. Kahneman, E. Diener, N. Schwarz (Eds.). Well-Being: The Foundations of Hedonic Psychology. Russel Sage Foundation, New York, pp. 213-229.

Diener, E., Ng, W., Harter, J., Arora, R., 2010. Wealth and Happiness Across the World: Material Prosperity Predicts Life Evaluation, Whereas Psychosocial Prosperity Predicts Positive Feeling. Journal of Personality and Social Psychology, 99(1), 52-61.

Diener, E., Tay, L., Oishi, S., 2013. Rising income and the subjective well-being of nations. Journal of Personality and Social Psychology, 104(2), 267-276.

Dolan, P., Peasgood, T., White, M., 2008. Do We Really Know What Makes Us Happy? A Review of the Economic Literature on the Factors Associated with Subjective Well-Being. Journal of Economic Psychology, 29(1), 94-122.

Donald, S.G., Lang, K., 2007. Inference With Difference-in-differences and Other Panel Data. The Review of Economics and Statistics, 89(2), 221-233.

Easterlin, R.A., 1974, Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In: P.A. David, M.W. Reder (Eds.). Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz. Academic Press, New York, pp. 88-125.

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., 2005a. Diminishing Marginal Utility of Income? Caveat Emptor. Social Indicators Research, 70(3), 243-255.

Easterlin, R.A., 2005b. 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., Angelescu, L., 2009, Happiness and Growth the World Over: Time Series Evidence on the Happiness-Income Paradox. IZA Discussion Paper No. 4060.

Easterlin, R.A., Angelescu McVey, L., Switek, M., Sawangfa, O., Smith Zweig, J., 2010. The Happiness-Income Paradox Revisited. Proceedings of the National Academy of Sciences of the United States of America, 107(52), 22463-22468.

Easterlin, R.A., Sawangfa, O., 2010, Happiness and Economic Growth: Does the Cross Section Predict Time Trends? Evidence from Developing Countries. In: E. Diener, J.F. Helliwell, D. Kahneman (Eds.). International Differences in Well-Being. Oxford University Press, Oxford, New York, pp. 166-216.

Feng, Z., Diehr, P., Peterson, A., McLerran, D., 2001. Selected Statistical Issues in Group Randomized Trials. Annual Review of Public Health, 22, 167-187.

Ferrer-i-Carbonell, A., Frijters, P., 2004. How Important Is Methodology for the Estimates of the Determinants of Happiness? The Economic Journal, 114(497), 641-659.

Fleurbaey, M., 2009. Beyond GDP: The Quest for a Measure of Social Welfare. Journal of Economic Literature, 47(4), 1029-1075.

Frederick, S., Loewenstein, G.F., 1999, Hedonic Adaptation. In: D. Kahneman, E. Diener, N. Schwarz (Eds.). Well-Being: The Foundations of Hedonic Psychology. Russel Sage Foundation, New York, pp. 302-329.

Friedman, B.M., 2005, The Moral Consequences of Economic Growth. Alfred A. Knopf, New York.

Frijters, P., Haisken-DeNew, J.P., Shields, M.A., 2004. Money Does Matter! Evidence from Increasing Real Income and Life Satisfaction in East Germany Following Reunification. The American Economic Review, 94(3), 730-740.

Graham, C.C., 2010. Adaptation Amidst Prosperity and Adversity: Insights from Happiness Studies from Around the World. The World Bank Research Observer, 26(1), 105-137.

Granger, C.W.J., Newbold, P., 1974. Spurious Regressions in Econometrics. Journal of Econometrics, 2(2), 111-120.

Greene, W.H., 2008, Econometric Analysis, vol. 6. Pearson Prentice Hall, Upper Saddle River.

Gujarati, D.N., Porter, D.N., 2009, Basic Econometrics, vol. 5. McGraw Hill, New York.

Guriev, S., Zhuravskaya, E., 2009. (Un)Happiness in Transition. Journal of Economic Perspectives, 23(2), 143-168.
Hagerty, M.R., 2000. Social Comparisons of Income in One's Community: Evidence from National Surveys of Income and Happiness. Journal of Personality and Social Psychology, 78(4), 764-771.
Hagerty, M.R., Veenhoven, R., 2003. Wealth and Happiness Revisited: Growing National Income Does Go with Greater Happiness. Social Indicators Research, 64(1), 1-27.
Hirshman, A.O., Rothschild, M., 1973. The Changing Tolerance for Income Inequality in the Course of Economic Development. The Quarterly Journal of Economics, 87(4), 544-566.
Huber, P.J., 1967, The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions. In: L.M. Le Cam, J. Neyman (Eds.). Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley, pp. 221-233.
Ibragimov, R., Müller, U.K., 2010. t-Statistic Based Correlation and Heterogeneity Robust Inference. Journal of Business & Economic Statistics, 28(4), 453-468.
Imbens, G.W., Kolesar, M., 2012, Robust Standard Errors in Small Samples: Some Practical Advice. NBER Working Paper 18478.
Inglehart, R., Foa, R., Peterson, C., Welzel, C., 2008. Development, Freedom, and Rising Happiness: A Global Perspective (1981–2007). Perspectives on Psychological Science, 3(4), 264-285.
Kahneman, D., Deaton, A., 2010. High Income Improves Evaluation of Life but not Emotional Well-Being. Proceedings of the National Academy of Sciences of the United States of America, 107(38), 16489-16493.
King, G., Roberts, M., 2012, How Robust Standard Errors Expose Methodological Problems They Do Not Fix. Mimeo, Harvard University.
Layard, R., Mayraz, G., Nickell, S., 2008. The Marginal Utility of Income. Journal of Public Economics, 92(8-9), 1846-1857.
Layard, R., Mayraz, G., Nickell, S., 2010, Does Relative Income Matter? Are the Critics Right? In: E. Diener, J.F. Helliwell, D. Kahneman (Eds.). International Differences in Well-Being. Oxford University Press, Oxford and New York, pp. 139-165.
Liang, K.-Y., Zeger, S.L., 1986. Longitudinal Data Analysis Using Generalized Linear Models. Biometrika, 73(1), 13-22.
Moulton, B.T., 1990. An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. The Review of Economics and Statistics, 72(2), 334-338.
Oswald, A.J., Powdthavee, N., 2008. Does Happiness Adapt? A Longitudinal Study of Disability with Implications for Economists and Judges. Journal of Public Economics, 92(5), 1061-1077.
Pfaff, T., 2013a, Clustered Standard Errors and Few Clusters: Some Practical Advice for Adjusting Inference. Mimeo.
Pfaff, T., 2013b, Income Comparisons, Income Adaptation, and Life Satisfaction: How Robust Are Estimates from Survey Data? SOEPpaper No. 555 on Multidisciplinary Panel Data Research. DIW Berlin.
Pollak, R.A., 1970. Habit Formation and Dynamic Demand Functions. Journal of Political Economy, 78(4), 745-763.
Putnam, R.D., 2000, Bowling Alone: The Collapse and Revival of American Community. Simon & Schuster, New York.
Rogers, W., 1993. Regression Standard Errors in Clustered Samples. Stata Technical Bulletin, 3(13), 19-23.
Sacks, D.W., Stevenson, B., Wolfers, J., 2010, Subjective Well-Being, Income, Economic Development and Growth. NBER Working Paper 16441.
Sacks, D.W., Stevenson, B., Wolfers, J., 2011, Growth in Income and Subjective Well-Being Over Time, Mimeo, New York.
Sacks, D.W., Stevenson, B., Wolfers, J., 2012. The new stylized facts about income and subjective well-being. Emotion, 12(6), 1181-1187.
Schaffer, M.E., 2010, xtivreg2: Stata Module to Perform Extended IV/2SLS, GMM and AC/HAC, LIML and k-class Regression for Panel Data Models. http://ideas.repec.org/c/boc/bocode/s456501.html.

Senik, C., 2004. When Information Dominates Comparison: Learning from Russian Subjective Panel Data. Journal of Public Economics, 88(9-10), 2099-2123.
Senik, C., 2008. Ambition and Jealousy: Income Interactions in the 'Old' Europe versus the 'New' Europe and the United States. Economica, 75(299), 495-513.
SOEP, 2011, Socio-Economic Panel Study, Data for years 1984–2010, Version 27, beta version of SOEPlong. DIW Berlin, SOEP.
Sribney, W., 2009, Comparison of Standard Errors for Robust, Cluster, and Standard Estimators. Stata FAQ Online, StataCorp. http://www.stata.com/support/faqs/statistics/standard-errors-and-vce-cluster-option/.
Stevenson, B., Wolfers, J., 2008. Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox. Brookings Papers on Economic Activity, Spring 2008, 1-87.
Stiglitz, J.E., Sen, A., Fitoussi, J.-P., 2010, Mismeasuring Our Lives: Why GDP Doesn't Add Up. The New Press, New York and London.
Thompson, S.B., 2011. Simple Formulas for Standard Errors that Cluster by Both Firm and Time. Journal of Financial Economics, 99(1), 1-10.
van den Bergh, J., 2009. The GDP Paradox. Journal of Economic Psychology, 30(2), 117-135.
van Praag, B.M.S., 1971. The Welfare Function of Income in Belgium: An Empirical Investigation. European Economic Review, 2(3), 337-369.
van Praag, B.M.S., Ferrer-i-Carbonell, A., 2008, Happiness Quantified: A Satisfaction Calculus Approach, vol. 2. Oxford University Press, Oxford and New York.
Vendrik, M.C.M., Woltjer, G.B., 2007. Happiness and loss aversion: Is utility concave or convex in relative income? Journal of Public Economics, 91(7-8), 1423-1448.
Vieweg, B.W., Hedlund, J.L., 1983. The General Health Questionnaire (GHQ): A comprehensive review. Journal of Operational Psychiatry, 14(2), 74-81.
Wagner, G.G., Frick, J.R., Schupp, J., 2007. The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. Schmollers Jahrbuch, 127, 139-169.
White, H., 1980. A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica, 48(4), 817-838.
Wooldridge, J.M., 2003. Cluster-Sample Methods in Applied Econometrics. The American Economic Review, 93(2), 133-138.
World Bank, 2013, World Development Indicators. Indicator Code: NY.GDP.PCAP.KD. http://data.worldbank.org/data-catalog/world-development-indicators.

Bisher erschienen:

Diskussionspapiere des Centrums für Interdisziplinäre Wirtschaftsforschung

DP-CIW 1/2011: Die Gemeinschaft der Lehrenden und Lernenden: Festvortrag zur Promotionsfeier der Wirtschaftswissenschaftlichen Fakultät am 24. November 2010 in der Aula des Schlosses
Alexander Dilger
Januar 2011

DP-CIW 2/2011: Plädoyer für eine problemorientierte, lerntheoretisch und fachlich fundierte ökonomische Bildung
Gerd-Jan Krol, Dirk Loerwald und Christian Müller
Februar 2011

DP-CIW 3/2011: Gefangen im Dilemma? Ein strategischer Ansatz der Wahl- und Revolutionsteilnahme
Marie Möller
April 2011

DP-CIW 4/2011: Overconfidence and Team-Performance: An Analysis of NBA-Players' Self-Perception
Hannah Geyer, Hanke Wickhorst
April 2011

DP-CIW 5/2011: Kompetenzziele für das allgemein bildende Fach „Wirtschaft/ Ökonomie“ in der Sekundarstufe I
AGOEB – Arbeitsgruppe Ökonomische Bildung
Mai 2011

DP-CIW 6/2011: Coping with Unpleasant Surprises in a Complex World: Is Rational Choice Possible in a World with Positive Information Costs?
Roger D. Congleton
Juni 2011

DP-CIW 7/2011: Warum der Baseler Ausschuss für Bankenaufsicht mit seinem antizyklischlen Kapitalpuffer falsch liegt
Björn Ludwig
Juli 2011

DP-CIW 8/2011: Bestimmungsgründe für die Beschäftigung und Rekrutierung von Älteren sowie für das Angebot an altersspezifischen Personalmaßnahmen
Christian Lehmann
August 2011

DP-CIW 9/2011: Das „Bruttonationalglück“ als Leitlinie der Politik in Bhutan – eine ordnungspolitische Analyse
Tobias Pfaff
September 2011

DP-CIW 10/2011: Economic Voting and Economic Revolutionizing? The Economics of Incumbency Changes in European Democracies and Revolutionary Events in the Arab World
Marie Möller
Oktober 2011

DP-CIW 11/2011: Geschlechtsspezifische Verdienstunterschiede und Diskriminierung am Arbeitsmarkt – Eine Untersuchung unter Berücksichtigung von Voll- und Teilzeitarbeit
Nele Franz
November 2011

DP-CIW 1/2012: Toward a More General Approach to Political Stability in Comparative Political Systems
Thomas Apolte
Januar 2012

DP-CIW 2/2012: An Empirical Study of the Limits and Perspectives of Institutional Transfers
Marie Möller
Februar 2012

DP-CIW 3/2012: Wie (un-) fair sind Ökonomen? Neue empirische Evidenz zur Marktbewertung und Rationalität
René Ruske, Johannes Suttner
September 2012

DP-CIW 1/2013: Zur Ethik von Rankings im Hochschulwesen – Eine Betrachtung aus ökonomischer Perspektiven
Harry Müller
Februar 2013

DP-CIW 2/2013: Which Qualifications Does a Minister of the German Federal Government Need to Be Reoccupied?
Katrin Scharfenkamp
März 2013

DP-CIW 3/2013: Unkonventionelle Geldpolitik – Warum die Europäische Zentralbank ihre Unabhängigkeit nicht verloren hat
Carsten Schwäbe
Mai 2013

Herausgeber:
Westfälische Wilhelms-Universität Münster
CIW – Centrum für Interdisziplinäre Wirtschaftsforschung
Scharnhorststr. 100
D-48151 Münster

Tel: +49-251/83-25329
Fax: +49-251/83-28429

www.wiwi.uni-muenster.de/ciw