

# Panel Conditioning and Subjective Well-Being: Evidence from International Panel Data and Repeated Cross-Sections

Bert Van Landeghem  
Maastricht University, ROA  
& University of Leuven, LICOS & Department of Economics  
b.vanlandeghem@maastrichtuniversity.nl

February 2012

## **Abstract**

Using data from three European countries, this paper aims to systematically investigate whether or not subjective well-being data are subject to panel conditioning or a panel effect, i.e. whether answers on subjective well-being questions depend on whether or not one has participated in the panel before. The comparison of time trends derived from panel data and repeated cross-sections within the same country, as well as the comparison between refreshment subsamples with experienced respondents within the same panel study, suggest a substantial panel effect. The analysis takes into account potential biases caused by panel attrition and panel aging. Self-reported well-being data seem to gain importance in the world of both academics as well as policy makers, and the awareness of a potential panel effect for these data might be important to interpret results or to design evaluation studies. Since investigations of panel effects are relatively limited, the results might encourage further research in panel conditioning for other important indicators in economics.

# 1 Introduction

During the last decades, economists have paid increasing attention to ‘stated utility’ as opposed to ‘revealed preferences’. While the latter are inferred from economic agents’ actions, the former are derived from very simple survey questions which go as follows:

On a scale from 0 (very unsatisfied) to 10 (very satisfied), how satisfied are you with your life, all things together?

There is increasing evidence underpinning the internal validity of such data (Krueger and Schkade, 2008; Sgroi et al., 2010) as well as the external validity (Oswald and Wu, 2010). Many relationships between happiness<sup>1</sup> data and socioeconomic variables, whether or not to be interpreted as causal, have been examined. Happiness has shown to be strongly correlated with relational goods and social capital (Bartolini et al., 2007; Becchetti et al., 2008; Powdthavee, 2008), with major life events such as unemployment, bereavement and disability (Clark et al., 2007; Oswald and Powdthavee, 2008a,b), with social status or rank (Blanchflower et al., 2009; Luttmer, 2005; Ravallion and Lokshin, 2010; Senik, 2004, 2008a, 2009) and with expectations and aspirations (de Grip et al., forthcoming; McBride, 2010; Senik, 2008b).

Subjective well-being data have become so important in economics that even in mainstream economic journals, methodological issues of such data are repeatedly being discussed. Bertrand and Mullainathan (2001) and Conti and Pudney (2011) argue that subtle changes in question wording and questionnaire design can influence results, not only the average scores in the raw data but also the correlations with socioeconomic variables. Oswald (2008) argues that one’s stated well-being might be a concave function of true well-being. This would imply that the log-linear relationship, which is often found between income and reported well-being, does not necessarily reflect the functional relationship between income and true well-being. Ferrer-i-Carbonell and Frijters (2004) find that treating happiness data as cardinal rather than as ordinal does not have a substantial impact on regression results, but taking into account fixed effects *does* have a dramatic impact on the results. Others have been concerned about anchoring effects, i.e. that people do not report the absolute level of happiness but rather their happiness relative to a reference point, and this reference point might be different across people, countries or time. Beegle *et al.* (2012) using data from the Tajikistan household survey,

---

<sup>1</sup>Happiness, life satisfaction, subjective well-being or self-reported well-being are often used interchangeably in the literature.

try to address this issue using vignettes. Subjective questions were asked in two parts of the survey. During the Subjective Poverty and Security module, respondents were asked a question which is known as the ladder question:

Imagine a 6-step ladder where on the bottom, the first step, stand the poorest people, and the highest step, the sixth, stand the rich. On which step are you today?

In another section, people were given four vignettes, each describing a hypothetical household. Respondents were required to put the four households as well as themselves on the ladder. The average rank respondents gave to themselves was slightly higher in the second section (post-vignette) than in the first section (pre-vignette), and 25% of them placed themselves on a different rung. Although it seems that individuals have different scalings, they do not seem to affect very much the estimated relationship with objective economic indicators. It even seems that pre-vignette measures are correlated more strongly with economic indicators than post-vignette measures. Bonsang and van Soest (2011) use vignettes to address differences in anchoring across countries, and find for older individuals in 10 European countries that cross-country differences in subjective well-being are more in line with cross-country differences in income after correcting scores with vignettes. In a study comparing the Netherlands and the United States, Kapteyn et al. (2011) use vignettes to correct for anchoring effects both within and across countries. They find that correcting subjective well-being scores with vignettes leads to almost identical distributions of subjective well-being in both countries, and leads to a substantially higher correlation between subjective well-being and income, particularly in the United States.

Another methodological issue one might be concerned about is a phenomenon called ‘panel conditioning’ or a ‘panel effect’, which implies that answers to questions depend on having participated in the panel before. Das et al. (2011) and Toepoel et al. (2009) give a cross-disciplinary overview of the literature on panel conditioning, and by comparing a fresh sample and a more experienced sample, they find that the panel effect is especially strong for knowledge questions but not for questions asking about attitudes. The authors note, however, that research on panel conditioning is still rather limited, and that often no attempt is made to distinguish panel conditioning from attrition. In the context of life satisfaction measures, Kassenboehmer and Haisken-Denew (2008), who try to explain the often-found U-shaped pattern of subjective well-being over the life cycle, find for German panel data that people report lower well-being scores the longer they are

in the panel. However, in a model which includes individual fixed effects, age and time effects, it might be difficult to interpret a variable measuring the length of staying in the panel. Indeed, multicollinearity problems force researchers (often unconsciously) to make arbitrary identifying assumptions, which however can have dramatic impacts on the results. Such identification issues have led to quite some confusion in the recent subjective well-being literature regarding the course of subjective well-being over the life cycle. These identification issues, discussed in more depth in Van Landeghem (2012), explain why studies using the same data sometimes report different results (see e.g. Clark and Oswald, 2006 versus Frijters and Beatton, 2011 and Gerstorf et al., 2010).

This paper aims to investigate the existence of panel conditioning in several ways, first by comparing trends in panel data with trends in repeated cross-sections within the same country, and second by comparing reported well-being of refreshment samples with reported well-being of more experienced respondents within the same panel dataset. Issues of panel ageing and panel attrition, which might cloud results, will be addressed.

This paper also somehow relates to a growing literature studying trends of subjective well-being over time. Blanchflower and Oswald (2004), find that happiness runs flat over time in Britain, that it is increasing for blacks in America but decreasing for white American women. In a recent paper, Stevenson and Wolfers (2009), using repeated cross-sections, document how in the United States, over the last 3 decades, women's happiness has decreased both in absolute terms as well as relative to men and that nowadays men are even happier than women while it was the reverse 3 decades ago. They find an upward-sloping trend in life satisfaction for both European men and women, though the increase was smaller for women than for men.

This study documents recent trends of well-being in Europe for three countries: Germany, the United Kingdom, and Switzerland. The paper will analyze the trend of well-being over time for men and women separately. The paper tries to be explicit about the identification assumptions and the interpretation of the estimates, and argues for a very simple regression specification. As the main theme of the paper is to cleanly identify the existence of a panel effect, this paper does not aim to enter the debate on how macroeconomic indicators are correlated with subjective well-being. The latter topic is discussed in papers such as Di Tella et al. (2001, 2003) and Wolfers (2003). In particular, the paper does not enter the recently revived debate on the bivariate (absence of) correlation between economic growth and happiness, a topic discussed in works such as Clark and Senik (2010), Deaton (2008), Di Tella and MacCulloch (2008),

Easterlin (1974, 1995, 2001), Easterlin et al. (2010), Sacks et al. (2010) and Stevenson and Wolfers (2008).

The remainder of the paper is as follows. Section two motivates the core of the empirical strategy. Section three describes the data that will be used throughout the analysis. Section four presents and discusses the econometric results, and Section five concludes.

## 2 The Empirical Strategy

This section offers some background for the regression specifications, and aims to illustrate why the most simple regression specification has been opted for when estimating time trends.

Subjective well-being can vary across time periods, but happiness may also differ across birth cohorts, and change with age (or evolve over the life cycle). The introduction of this paper already mentions several studies that have documented time trends or age trends in subjective well-being data. One can also imagine that baseline happiness can vary across cohorts or years-of-birth, as the one cohort has experienced different circumstances than the other cohort. Pezzini (2005) finds that women who had experienced less restrictive abortion laws during their child-bearing age are, on average, happier than cohorts who had still been subject to more stringent rules. Oreopoulos (2007) finds, among other things, that the cohorts who had been subject to an increase in minimum school-leaving age were, on average, happier than cohorts who had not been subject yet to the new compulsory schooling law.

Since time, year-of-birth, and age are perfectly collinear, they cannot be included simultaneously in a regression framework without any functional assumptions.

Let's assume that there are  $L$  years-of-birth (indexed with dummies  $c_1$  to  $c_L$ , and  $K$  calendar years for which data are available (indexed with dummies  $t_1$  to  $t_K$ ). If  $n_l$  denotes the birth year of the cohort  $c_l$ , and  $p_k$  the calendar year associated with period  $t_k$ , then an individual's age  $age_j$  can be computed as follows:

$$age_j = \left( \sum_{k=1}^K (p_k * t_k) \right) - \left( \sum_{l=1}^L (n_l * c_l) \right) \quad (1)$$

When including the three vectors of age dummies  $a_j$ , time dummies  $t_k$ , and birth cohort dummies  $c_l$  in a regression framework as independent variables together with a constant, there will not be just 3 distinct linear dependencies across the vectors (as is always the case with exhaustive and mutually exclusive sets of dummy variables), but there will be a fourth dependency, namely each age dummy can be written as:

$$a_j = \frac{1}{n_j} * \left( \sum_{k=1}^K (p_k * t_k) \right) - \left( \sum_{l=1}^L (n_l * c_l) \right) \quad (2)$$

Leaving out a dummy from one of each set, as is the standard practice to overcome trivial dependencies with dummy variables, will thus not be enough. One might opt for leaving out a second dummy from one of the sets. However, this implies that one needs to assume that two time effects, age effects, or cohort effects are equal. In case this normalization assumption is incorrect, the estimated curves will be a rotated version of the actual curves. Not including a constant, and leaving out one dummy from each set leads to the same problem. Indeed, leaving out a constant and taking individuals of age category  $a_j$  and time period  $t_k$  as baseline implies that the baseline cohort is  $c_{j-k+K}$ . The sum of the three effects associated with the omitted dummies need to be assumed to be equal to the constant, and the individuals in the baseline category thus need to be, on average, in their steady-state. Differencing the data in an attempt to get rid of the constant and the cohort effects leaves us with similar issues. Two sets of differenced dummies require us to leave out a differenced time or age dummy, meaning that we again have to assume that 2 consecutive age effects or time effects are equal to each other.<sup>2</sup>

One may proceed by regressing the dependent variable only on a set of time dummies and a constant, which yields the following regression specification:

$$WB_k = \alpha + \sum_{k=2}^K \beta_k t_k \quad (3)$$

---

<sup>2</sup>A small algebraic example showing how the curves rotate when making the wrong assumptions in the case of a first-difference regression is provided in Van Landeghem (2012), Appendix 1. A salient graphical example to underpin the nontriviality of the concern is given in the results section of this paper.

Assuming that the age distribution in all time periods is the same<sup>3</sup>, the estimate  $\beta_k$  then equals

$$E[WB_k - \sum_{\tilde{k}=1, \tilde{k} \neq k}^K WB_{\tilde{k}}] = T_k + \sum l = 1^L (W_{lk} C_l) \quad (4)$$

where  $T_k$  and  $C_l$  denote the actual time and cohort effect respectively,  $W_{lk}$  denotes a weight for the cohort effect  $C_l$  in period  $k$ , and  $\sum_{l=1}^L w_{lk} = 1$  for any  $k = 1 \dots K$ . Indeed, these weights will differ across periods: cohorts can age in and out of the panel as there are a finite number of age groups in the data. The oldest cohort will therefore only be observed in the first time period, while the youngest cohort will only be observed in the last period.

Hence, in the specification shown in Equation 3, the changes in the coefficients on time dummies reflect changes in well-being due to changes in time-varying factors, plus changes in population composition (older cohorts age out, while younger cohorts age in).

One should note that controlling for cohort effects would not lead us to estimates which have a more natural interpretation, on the contrary even. Indeed, when conditioning on the birth cohort, the  $\beta_k$ -coefficient would then equal:

$$E[WB_k - \sum_{\tilde{k}=1}^K E[WB_{\tilde{k}} | c_1, c_2, \dots, c_L]] \quad (5)$$

In words, by introducing cohort dummies, one is comparing individuals from the same birth cohort across time periods. Comparing individuals from the same birth cohort across time periods implies comparing individuals with different ages. Even though there is still some disagreement about the true age–wellbeing relationship in the literature<sup>4</sup>, researchers seem to agree that the pattern is smooth rather than transitory fluctuations around the mean. In case there exists a trend of well-being in age rather than transitory fluctuations around a mean, it will be picked up by the  $\beta_k$ -coefficients. The interpretation of the  $\beta_k$ -coefficients is then far less straightforward.

---

<sup>3</sup>In reality, this assumption will not hold completely. When estimating regressions, results from panel data will be used with weights constructed to make sure that all ages get the same weight in each calendar year. Indeed, since a panel study interviews individuals year after year and since refreshment is limited, panels will generally age faster than the population which the panel aims to represent.

<sup>4</sup>Often, researchers find that well-being declines over the life cycle until midlife, rises until old age and then declines again (Blanchflower and Oswald, 2008; Fischer, 2009; Gwozdz and Sousa-Poza, 2010). Others put forward other shapes as more plausible (Easterlin, 2006; Frijters and Beaton, 2011).

In order to fix ideas, let's consider an example with two time periods and two age groups. Both age groups are present in each survey year, which means that there are three cohorts. Moreover, let's assume that we are in a world where happiness does not change across time or across birth cohorts, but where happiness changes linearly with ageing by a factor  $\lambda$ . In case the age distribution is the same in both time periods, estimating Equation 3 should yield a  $\beta_2$  of 0: well-being does not change over time. In case one estimates Equation 5, one needs to match cohorts across the two time periods. Only information from cohort two will be used in the estimation since it is the only cohort that is present both in time period one and time period two. Since cohort two is one year older in time period two than in time period one, the expected difference in well-being of cohort two between period two and one will be  $\lambda$ , and not 0. Hence, the coefficient  $\beta_2$  will pick up the linear age trend, and will not reflect the changes in average happiness in a society.

## 3 Data

### 3.1 Panel Data

Among the three panel datasets analyzed, the German Socio-Economic Panel (GSOEP) is probably the dataset which allows the most extensive analysis and which is most being used in happiness research. The German Socio-Economic Panel is provided by DIW Berlin, and is repeated with yearly intervals, running from 1984 for West Germany and 1990 for East Germany (see Wagner et al., 2007).

The observations can be categorized in eight<sup>5</sup> different samples:

- Sample A (started in 1984) represents the West-German population, while sample C (started in 1990) represents the East German population.
- Samples E, F and H (started in 1998, 2000 and 2006 respectively) are refreshment samples in order to reduce the effects of panel ageing and attrition.
- Samples B and D (started in 1984 and 1994 respectively) are immigrant subsamples, and sample G (started in 2002) comprises high-income households. The latter

---

<sup>5</sup>In 2009, a ninth sample has been started and aims to gather information that is not available in the other samples in order to allow the investigation of new research questions.

three samples are excluded from the analysis in order to avoid that kinks in well-being trends are caused by the oversampling of one of these categories.

One then ends up with a sample of in total around 31000 individuals, 21000 in West Germany and 10000 in East Germany.

The following question on subjective well-being, asked at the end of a face-to-face questionnaire, has been in the survey from 1984 onwards and has since then reappeared in every survey round.

On a scale from 0 (completely dissatisfied) to 10 (completely satisfied): How satisfied are you with your life, all things considered?

Across all rounds, the average life satisfaction score in West Germany equals 7.15 with standard error of 1.82. In east Germany, life satisfaction is, on average, considerably lower than in West Germany, with an average score of 6.41 and standard error of 1.80.

The British Household Panel Survey (BHPS) is made available through the ESRC Data Archive and originally collected by the ESRC Research Centre on Micro-social Change at the University of Essex. The BHPS started in 1991. In 1999, two additional samples were included in order to oversample households in Wales and Scotland. In 2001, a Northern Ireland sample was started so that the survey covered the UK area rather than just Great Britain. From 1997 to 2001, additional households were interviewed for the purpose of the European Community Household Panel, and those households who agreed that their information was passed on to the University of Essex constitute another subsample in the BHPS, the ECHP subsample. The latter four samples are, however, not included in the analysis again to avoid artificial results, which leaves us with data for around 15300 individuals and 77000 person-year observations.

Questions asking about people's well-being are asked at the beginning of a self-completion questionnaire. Unfortunately, however, not all rounds contain all the information that should ideally be used throughout the analysis. A general life satisfaction question is only included from the year 1996 onwards, and has been skipped in 2001, which means that we have subjective well-being data for the period 1996-2000 and 2002-2008. The 'satisfaction with life' data for waves 6-10 and 12-18 are based on the following survey question:

Please tick the number which you feel best describes how dissatisfied or satisfied you are with your life overall? 1 (not satisfied at all) to 7 (completely satisfied)

Across these 12 waves and all surveyed individuals, respondents report an average life satisfaction score of 5.22 on a scale of 7, with a standard deviation of 1.29. In 1.48% of the responses, an individual indicates being completely dissatisfied, while in 14.49% of cases, respondents are completely satisfied.

All 18 rounds, however, include questions from a General Health Questionnaire (GHQ), which are designed to identify depression. Due to a lack of subjective well-being data, the GHQ data are often used instead (e.g. Metcalfe et al., 2011). More in particular, these questions go as follows:

Here are some questions regarding the way you have been feeling over the last few weeks. For each question please tick the box next to the answer that best describes the way you have felt.

Have you recently...

1. been able to concentrate on whatever you're doing? ++
2. lost much sleep over worry? -
3. felt that you were playing a useful part in things? ++
4. felt capable of making decisions about things? ++
5. felt constantly under strain ? -
6. felt you couldn't overcome your difficulties ? -
7. been able to enjoy your normal day-to-day activities ? ++
8. been able to face up to problems ? ++
9. been feeling unhappy or depressed ? -
10. been losing confidence in yourself ? -
11. been thinking of yourself as a worthless person ? -
12. been feeling reasonably happy, all things considered ? ++

For the items followed by '+' the response scale is as follows:

1. More so than usual

2. About the same as usual
3. Less so than usual
4. Much less than usual

And for the items followed by ‘-’:

1. Not at all
2. Not more than usual
3. Rather more than usual
4. Much more than usual

In order to construct an aggregate measure, a dummy is created for each item taking 1 when the score on the item equals 3 or 4. By summing up the dummies within a person-year observation, one obtains a 0-12 depression scale. Subsequently, the score is reversed, so that 12 = no depressive symptoms, and 0 = severely depressed. This measure (labelled as GHQ score) will be used as an alternative dependent variable during the regression analysis instead of the life satisfaction measure. Across all rounds and individuals, the average GHQ score equals 9.30, with a standard deviation of 3.20. 2.26% of respondents have the lowest GHQ score of 0, while 31.85% of respondents have the highest GHQ score of 12. Hence, there are considerably more individuals with the highest GHQ score than with the highest life satisfaction score, which illustrates the upper truncation of the former measure. The GHQ score is appropriate to identify different degrees of depression, but it cannot help us to distinguish the satisfied individuals from the extremely satisfied individuals.

The Swiss Household Panel (SHP) is a panel repeated with yearly intervals which started in 1999 and which is run by the FORS, the Swiss Centre of Expertise in the Social Sciences. Currently, there are 12 rounds available (up to 2010), with more than 95000 person-year observations and 18300 individuals. In 2004, a refreshment sample was started.

Respondents are interviewed by telephone, and a general life satisfaction question is asked from the year 2000 onwards and goes as follows:

In general, how satisfied are you with your life if 0 means ‘not at all satisfied’ and 10 means ‘completely satisfied’?

The average score across all person-year observations equals 8.01 (with a standard error of 1.48). 16.5% of individuals consider themselves as completely satisfied.

In all three panels, one aims, by definition, to re-interview the same individuals in successive rounds. new individuals can enter, however, for several reasons. As discussed above, new samples can be added to the survey, e.g. to oversample some minorities or to refresh the panel. Apart from adding new samples, the datasets will contain every year a small number of new respondents. First, there are members from interviewed households who reach the eligible age to enter the questionnaire. Second, if a new member eligible for the survey moves into the household, one will attempt to interview this new member as well. Third, if a household member leaves the household, one aims to follow the respondent and one will also aim to interview the other members eligible for the survey in his new household. Generally, new respondents not stemming from a refreshment sample account for around 5% of the respondents, and they ensure that each age group is represented in each survey round.

## 3.2 Repeated Cross-Sections

In order to compare results from panel data with results from repeated cross-sections, the paper draws upon data from the World Values Survey and the Eurobarometer Survey.

The World Values Survey (WVS) is a face-to-face survey conducted worldwide and aims to facilitate the investigation of sociocultural and political changes across regions and over time. The World Values Survey now counts five waves spanning the periods 1981-1984, 1989-1993, 1994-1999, 1999-2004, and 2005-2008 respectively. Currently, a sixth wave is carried out. The World Values Survey does not follow the same people over time. Nevertheless, many countries have been included in the survey in multiple rounds, which offers us repeated cross-sectional data for these countries. In particular, the 3 countries studied in this paper have been surveyed at least twice, which means that a comparison between panel data and repeated cross-sectional data is possible. There are data for Germany for the years 1997 and 2006, for Britain for the years 1998 and 2006, and for Switzerland for the years 1989, 1996, and 2007. For each country, around 1100 to 2000 individuals are surveyed in one round. The life satisfaction variable in the World Values Survey is measured with the following question:

Taking all things together, would you say you are:

- Very happy
- Quite happy
- Not very happy
- Not at all happy

During the analysis, scores 1 and 4, and 2 and 3 are swapped so that 1 denotes not at all happy and 4 denotes very happy, which makes the results easier to compare with results from the panel data. For the different waves and the three countries together, respondents report an average life satisfaction score of 3.11 (standard deviation 0.67). 31.2% of individuals report the highest happiness score. This is a significantly larger percentage than in the other data sets, but this might have to do with the more restricted scale which allows respondents less differentiation. Only 1.6% of individuals report the lowest happiness score.

A second repeated cross-sectional data set is the Mannheim Eurobarometer trend file, which runs from 1970 to 2000. A variable is defined as a trend variable if it can be found in a Eurobarometer survey (conducted with face-to-face interviews) in at least five years. Compared to the World values Survey, this dataset has the advantage to contain cross-sections repeated with yearly intervals. A first subjective well-being question goes as follows:

On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the life you lead?

As with the subjective well-being measure in the WVS, scores 1 and 4 are swapped, as well as scores 2 and 3, in order to have a subjective well-being variable which goes from low to high well-being. The question has been asked at a yearly basis from 1973 to 1998, except for the years 1974 and 1996.<sup>6</sup> Averaging all available data for West and East Germany and Britain gives us a life satisfaction score of 3.13. 2% of individuals report the lowest well-being score, while 29% report the highest well-being score.

---

<sup>6</sup>The Eurobarometer trend file also contains a happiness question. This question has not been asked any more after 1986, which means that these data are not of any use for comparison with the panel data sets.

## 4 Results

### 4.1 Results from Panel Data Compared with Results from Repeated Cross-Sections

This subsection examines the estimates from the very parsimonious and straightforward regression as of the form displayed in Equation (3), which means a pooled OLS regression containing only a constant and time dummies. The curve describing the pattern of well-being over time is thus driven by time-varying variables, as well as by changes in the population's cohort composition. Indeed, younger cohorts age into the data, while older cohorts age out of the data. The latter regression is run for the 3 concerned panel datasets on the whole sample and for men and women separately. Subsequently, these trends (women and men aggregated) are compared with available repeated cross-sectional data.

#### 4.1.1 Trends in Germany

Figure 1 shows us trends of life satisfaction from the GSOEP, for the entire West German sample, as well as for men and women separately. During the first five years of the panel, we see an overall decline in well-being from 7.4 to 7.0. In the next three years, following reunification, we see an upsurge in well-being, after which there is an almost steady decline until 1997, where well-being reaches a dip of 7. During the next four years (which was a period of economic recovery), well-being increased again up to a score of 7.3 in 2001. From 2002 to 2010, the pattern is a bit more irregular but overall downward sloping (with a dip in 2004). The GSOEP data seem to suggest, for West Germany, that well-being over time is overall very slightly downward sloping. It seems that, over these 27 years, well-being of men and women have known a very similar pattern. Moreover, women and men report, on average, the same level of life satisfaction both at the beginning and at the end of the period analyzed.

Figure 2 shows us trends of life satisfaction for East Germany from 1990 onwards. Life satisfaction in East Germany is considerably lower than in West Germany. This gap narrows near the end of the sample but is still very pronounced, even 20 years after reunification. The huge drop in life satisfaction between 1990 and 1991 followed by a recovery is remarkable, and is consistent with results found for East Germany and other

transition countries by Sanfey and Teksoz (2007), Easterlin (2009) and Easterlin and Plagnol (2009). From 1991 onwards, the overall pattern for East Germany is increasing. This discrepancy with West German data might be interpreted as evidence of a reduction in relative deprivation (in many ways) compared to neighbouring countries. Moreover, while male and female life satisfaction are on average the same at the beginning of the time span, women are considerably happier at the end of the time span compared to men.

Figure 3 illustrates the claim in Section 2 that including age and time dummies and differencing away the cohort effect still requires us to make a nontrivial normalization assumption. The two curves in Figure 3 are based on regressions with different normalizations. The regression underlying the first curve assumes that the age effects of age 18 and age 19 are equal to each other. For the second curve is assumed that the age effect of ages 57 and 58 are equal to each other. It is clear that the 2 different normalizations lead to completely different results. According to the first specification, there is a steady and huge drop in life satisfaction over time. Life satisfaction in 1985 is almost 4 points higher than life satisfaction in 2010. According to the second specification, the drop in life satisfaction is much smaller, around 0.9 points over a 26-year period.

If we now turn to German data from the World Values Survey, comparison of means between outcomes for 1997 and 2006 reveals no downward trend. If anything, average well-being is even slightly higher in the latter year than in the former (2.973 versus 2.966). Figure 4 shows us trends in standardized life satisfaction scores from GSOEP data and Eurobarometer data over the period 1984-1998. As subjective well-being data is missing in the Eurobarometer Survey for 1996, a score for this year is imputed by averaging scores of 1995 and 1997. Generally, well-being in the Eurobarometer survey is upward sloping and higher in 1995 than in 1984. The sudden dip at the end of the period is very remarkable. Stevenson and Wolfers (2009), have been puzzling over these West German results as well. When looking at each country separately in the Eurobarometer survey, they find West Germany being the exception where life satisfaction was lower at the end of the period than at the beginning of the period. One should note, however, that the context in which the trend question is being asked can vary across different survey rounds. For example, the 1995 data are drawn from a questionnaire conducted during the period April-May which placed the life satisfaction second in row (after asking about one's nationality. In 1997, the life satisfaction question was the 36th question in a questionnaire conducted during the period March-April. As earlier research suggests,

placement of a question in a questionnaire might influence reported well-being scores (Conti and Pudney, 2011). Moreover, conducting a survey earlier in the year might influence results for reported well-being e.g. due to a lower amount of sunshine (Guven, forthcoming).

Figure 5 shows us the results from the same data sources but now for East Germany for the period 1990-1998. The comparison between both datasets suggests that the very significant drop in life satisfaction between 1990 and 1991 shown by GSOEP data is most likely not entirely due to a panel effect.

#### 4.1.2 Trends in Britain

Figure 6 shows us trends in GHQ scores for the whole sample and men and women separately from 1991 to 2008 onwards. A GHQ score of 0 means ‘seriously depressed’ and 12 means ‘no depressive symptoms’. It is remarkable that GHQ scores for women are considerably lower than GHQ scores for men.<sup>7</sup> This reminds us of the fact that a depression index is conceptually different from a life satisfaction score. The GHQ scores show an overall slightly downward sloping pattern, with peaks in 1999 and a dip in the first years of the millennium. In 2008, there is a huge jump in GHQ scores, both for men and for women, of more than 1 point. The positioning of the GHQ-questions does not seem to have changed in the questionnaire and the increase is not caused by new individuals entering the panel.

Figure 7 shows us trends in life satisfaction from 1996 onwards. As no data are available for 2001, the value for 2001 has been imputed by averaging the mean of scores in 2000 and 2002. Women and men report more or less the same life satisfaction scores. It is also interesting to note that life satisfaction peaks in 1998, which is one year earlier than the GHQ scores. There is also a peak in life satisfaction in 2008.

According to the World Values Survey, the British report, on average, a well-being score of 3.21 in 1998, which increases to 3.42 in 2006. This increase is statistically significant at any significance level. Figure 8 shows us standardized GHQ scores from the BHPS and standardized life satisfaction scores from the Eurobarometer Survey for the period 1991-1998. The standardized GHQ scores from the BHPS (aggregated for

---

<sup>7</sup>This finding is not just limited to Britain, as do suggest unreported results from the Panel Study of Belgian Households.

men and women) show an overall downward sloping trend over the examined period, while there is no such observation for the life satisfaction scores in the Eurobarometer survey, which do not show a clear trend.

### 4.1.3 Trends in Switzerland

Life satisfaction time trends derived from the Swiss Household Panel are shown in Figure 9 for the whole population, and men and women separately, for the years 2000 to 2010.

The life satisfaction measure in the SHP has a 0-10 range, as does the life satisfaction measure in the GSOEP. This means that we can quite easily compare the averages of the scores. It is then interesting to see that Swiss respondents rate their satisfaction with life on average much higher than Germans. In West Germany, average life satisfaction over time ranges from 6.9 to 7.4, while in Switzerland the average scores range from 7.9 to 8.2. Of course, even though the questions are identical, the design of the survey is not the same. For example, while the GSOEP gathers data through face-to-face interviews, the interviews for the SHP are conducted by telephone. Anchoring effects might play a role as well.

In line with the other datasets, the Swiss data show some cyclicity in well-being over time, and on average, there seems to be a slight decrease over time. The time span is rather short, so difficult to observe convergence or divergence in male and female happiness. Female happiness is slightly higher than male happiness in the beginning of the period, but is equal to male happiness at the end of the time span.

As for Swiss WVS data, average well-being scores equal 3.29 in 1989, 3.34 in 1996 and 3.35 in 2007. The increase between 1989 and 1996 is statistically significant at the 7% significance level, the increase over the period 1996 to 2007 is not economically significant, nor statistically significantly different from 0. The increase between 1989 and 2007, however, is statistically significant at the 4% significance level, and seems to be of a substantial magnitude as well (0.05) given the limited scale of the variable and the relatively short time span. Unfortunately, Eurobarometer data for Switzerland are not available.

## 4.2 A Comparison between New Respondents and Experienced Respondents within a Panel

The previous subsection offered at least some mild first evidence for a panel effect: answers on subjective well-being questions seem to be influenced by whether or not one participated in the panel before. Comparing results from panel data with the existing repeated cross-sections is not ideal, however, to identify a panel effect. First, the design of the questionnaires used to construct the panel datasets and the repeated cross-sectional datasets are different, as well as the question wording and the answering scales for the well-being indicators. Moreover, comparisons are made difficult by the fact that repeated cross-sectional data is not always available with yearly intervals, and because subtle changes in survey design in both panel data and repeated cross-sections might be the cause of artificial results (Conti and Pudney, 2011).

An alternative way to investigate the existence of a panel effect is therefore to compare the score of a refreshment sample with the score of more experienced respondents for that same year within the same panel dataset. As is clear from Section 3, this robustness check can be performed in the German Socio-Economic Panel and in the Swiss Household Panel. In the years in which the refreshment samples are started, there is a very substantial number of new respondents, which facilitates the test. When comparing a more experienced sample with a refreshment sample within the same panel study, one does not need to worry about differences in survey design. The differences in average scores for the new respondents in the refreshment samples on the one hand, and the experienced respondents in the corresponding calendar year are represented in Table 1 for West Germany, East Germany, and Switzerland. Again, weights are used in order to make sure that each age group has the same weight in both samples.

The German Socio-Economic Panel contains 3 refreshment samples in the years 1998, 2000 and 2006, while the SHP introduces a refreshment sample in 2004. For West Germany, the three refreshment samples offer us 1550, 8661, and 2044 new respondents, respectively while this is 332, 2097, and 542 for East Germany. In the Swiss Household Panel, the 2004 panel refreshment brings us data on new respondents for 5371 individuals.

In all cases, scores in the calendar year in which a refreshment sample is started are higher for the refreshment sample than for the more experienced sample, even after

using weights to correct for differences in age distributions. The results are different from 0 at any significance level, and the magnitude is substantial, varying between 0.16 and 0.68 on a 0-10 scale.

One might be worried that the samples are nevertheless substantially different. A first check to test whether results are artificial is to compute the difference between real household income in the refreshment subsample on the one hand, and in the experienced sample on the other hand, again using weights to correct for differences in age distributions in both samples. The results suggest that differences in average income between the refreshment samples and experienced samples is rather small (see Table 2). Real household income in the refreshment samples for Germany are on average 0 to 17 per cent lower than in the more experience sample. As for the Swiss Household Panel, real household income is around 6% lower in the refreshment sample than in the experienced sample. These results thus seem to reassure that the differences between well-being in refreshment and experienced samples shown in Table 1 indeed reflect a panel effect, since it is well-known that life satisfaction scores are correlated positively with income. The observation that refreshment samples have, on average, a higher income than the more experienced sample is in line with the observation in the literature that people with lower income are likely to exit the panel. of course, reported household income might suffer from panel conditioning as well.

A second robustness check, which aims to investigate the possibility that the results in Table 1 are driven by a selection (or attrition) effect, instead of a panel effect, makes use of the fact that all refreshment samples are introduced well before the last calendar year for which data are available. This facilitates us to compare the experienced samples only with those individuals from the refreshment samples who will at least be in the panel for three more years. Hence, the refreshment and experienced samples should be more comparable. The large number of waves in the two panel datasets give us some more flexibility than have Das et al. (2011). The latter extend the framework of Keisuke et al. (2001) by recognizing that differences between an experienced sample and a refreshment sample can be due to not only panel attrition but also panel conditioning, but who have to cope with the restriction that the refreshment sample is only observed in one round.

The results of this exercise are presented in Table 3, and seem to suggest that the discrepancy is even slightly higher. It thus seems that a selection effect would bias downward a panel effect rather than the reverse. Indeed, attrition often occurs after a household has suffered from an unpleasant life event (such as bereavement of a family

member) which is believed to influence subjective well-being negatively. Moreover, low life satisfaction scores are good predictors for future attrition (Kroh, 2011; Lips, 2007).

Of course, in order for the latter robustness check to be convincing, the course of attrition over the age of a sample should be similar for the experienced and the refreshment samples. Attrition rates could be influenced by time-varying socioeconomic factors, or by factors related to the data collection (e.g. different interviewers, different management etc.). Large differences in attrition rates might then be an indication that the characteristics of the attritors across the experienced sample and the refreshment sample are not comparable.

Figures 10, 11, and 12 show the attrition rate over the life cycle for the different samples for West Germany, East Germany and Switzerland, respectively. An individual is regarded as an attritor at age  $n$  of the original sample of respondents if the individual was a respondent at age 1, and if the individual answered the subjective well-being question strictly less than  $n$  times. The evolution of the attrition rates over a sample's life cycle within a panel are rather similar, concavely increasing. Attrition in the Swiss Household Panel is much higher than in the German Socio-Economic Panel, but this goes for both the original 1999 sample as for the 2004 refreshment sample.

## 5 Conclusion

This paper has focussed on trends in subjective well-being data in panels and repeated cross-sections, with as main aim to investigate whether answers to subjective well-being questions exhibit a panel effect, or, in other words, whether the score on subjective well-being questions depends on having participated in the panel before. The paper documents very basic regressions, as they are likely to yield more useful results than more complicated specifications.

The patterns of the life satisfaction measures in the different panel datasets show that in most cases, scores are slightly lower in the last rounds of data than in the first rounds of data, even when weights are used to correct for panel aging. Patterns in life satisfaction are not very different for men than for women (likely because of the relatively short time span of the data), but the data from the British Household Panel Survey show that women are more likely to report depressive symptoms than men.

When comparing panel data results with repeated cross-sectional results from the World Values Survey and the Eurobarometer Survey, one observes some discrepancies. While panel data show for different countries (apart from East Germany) an overall downward trend, repeated cross-sections for these countries rather indicate an increase in well-being over time, and this comparison offers some first (mild) evidence for a panel effect. More convincing evidence seems to emerge from comparing refreshment samples during their first year with the more experienced sample: subjective well-being scores are substantially higher in the refreshment sample than in the more experienced sample in the corresponding calendar year. Importantly, these results do not seem to be driven by panel attrition or panel aging.

Strong evidence on the existence of a panel effect has important implications for economists and policy makers using subjective well-being data. First, the results can guide the increasing literature on trends in well-being, as they suggest that repeated cross-sections and pseudo panel data might be favoured above genuine panel data to study the latter phenomenon. Second, they might help to fine tune the design of policy evaluations. For example, oversampling a minority in a certain year of a household panel to compare its well-being scores with those of the experienced subsample might lead us to wrong conclusions, even if attrition in the older sample is negligible. Finally, the identification of a substantial panel effect in subjective well-being data might help to draw the attention to this phenomenon in other areas of economics.

## References

Bartolini, S., E. Bilancini and M. Pugno (2007) Did the Decline in Social Capital Decrease American Happiness? A Relational Explanation of the Happiness Paradox. *Working Paper*, Department of Economics, University of Siena.

Becchetti, L., A. Pelloni and F. Rossetti (2008) Relational Goods, Sociability, and Happiness. *Kyklos* 61, 343-363.

Beegle, K., H. Himelein and M. Ravallion (2012) Frame-of-Reference Bias in Subjective Welfare. *Journal of Economic Behavior and Organization* 81, 556-570.

Bertrand, M. and S. Mullainathan (2001) Do People Mean What They Say? Implications for Subjective Survey Data. *American Economic Review* 91, 67-73.

Blanchflower, D. and A. Oswald (2004) Well-Being Over Time in Britain and the USA. *Journal of Public Economics* 88, 1359-1386.

Blanchflower, D. and A. Oswald (2008) Is Well-Being U-Shaped Over the Life Cycle? *Social Science and Medicine* 66, 1733-1749.

Blanchflower, D., A. Oswald and B. Van Landeghem (2009) Imitative Obesity and Relative Utility. *Journal of the European Economic Association* 7, 528-538.

Bonsang, E. and A. van Soest (2011) Satisfaction With Job and Income Among Older Individuals Across European Countries. *Social Indicators Research* 105, 227-254.

Clark, A., E. Diener, Y. Georgellis and R. Lucas (2007) Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis. *Economic Journal* 118, 222-243.

Clark, A. and A. Oswald (2006) The Curved Relationship between Subjective Well-Being and Age. *PSE Working Papers* 2006-29.

Clark, A. and C. Senik (2010) Will GDP Growth Increase Happiness in Developing countries? *PSE Working Papers* 2010-43.

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

Das, J., A. van Soest and V. Toepoel (2011) Nonparametric Tests of Panel Conditioning and Attrition Bias in Panel Surveys. *Sociological Methods and Research* 40, 32-56.

Deaton, A. (2008) Income, Health, and Well-Being around the World: Evidence from the Gallup World Poll. *Journal of Economic Perspectives* 22, 53-72.

de Grip, a., M. Lindeboom and R. Montizaan (forthcoming) Shattered Dreams: The Effects of Changing the Pension System Late in the Game, *The Economic Journal*.

Di Tella, R. and R. MacCulloch (2008) Gross national happiness as an answer to the Easterlin Paradox? *Journal of Development Economics* 86, 22-42.

Di Tella, R., r. MacCulloch and A. Oswald (2001) Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. *American Economic Review* 91, 335-341.

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

Easterlin, R. (1974) Does Economic Growth Improve the Human Lot? In Paul A. David and Melvin W. Reder, eds., *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*, New York: Academic Press, Inc.

Easterlin, R. (1995) Will Raising the Incomes of All Increase the Happiness of All? *Journal of Economic Behavior and Organization* 27, 35-48.

Easterlin, R. (2001) Income and Happiness: Towards a Unified Theory. *Economic Journal* 111, 465-484.

Easterlin, R. (2006) Life Cycle Happiness and Its Sources. Intersections of Psychology, Economics, and Demography. *Journal of Economic Psychology* 27, 463-482.

Easterlin, R. (2009) Lost in Transition: Life Satisfaction on the Road to Capitalism. *Journal of Economic Behavior and Organization* 71, 130-145.

Easterlin, R., L. Angelescu-McVey, M. Switek, O. sawangfa, J. smith-Zweig (2010) The Happiness-Income Paradox Revisited. *Proceedings of the National Academy of Sciences* 107, 22423-22468.

Easterlin, R. and A. Plagnol (2009) Life Satisfaction and Economic Conditions in East and West Germany Pre- and Post-Unification. *Journal of Economic Behavior and Organization* 68, 433-444.

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

Fischer, J. (2009) Happiness and Age Cycles - Return to Start. . . . *MPRA Paper* No. 15249, University of Munich.

Frijters, P. and T. Beaton (2011) The mystery of the U-shaped relationship between happiness and age". *Working Paper*, Queensland University of Technology.

Gerstorff, d., N. Ram, G. Mayraz, M. Hidajat, U. Lindenberger, G. Wagner, J. Schupp (2010) Late-Life Decline in Well-Being Across Adulthood in Germany, the United Kingdom, and the United States: Something Is Seriously Wrong at the End of Life. *Psychology and Aging* 25, 477-485.

Güven, C. (Forthcoming) Reversing the Question. Does Happiness Affect Consumption and Savings Behavior? *Journal of Economic Psychology*.

Gwozdz, W. and A. Sousa-Poza (2010) Ageing, Health and Life Satisfaction of the Oldest Old: An Analysis for Germany. *Social Indicators Research* 97, 397-417.

Kapteyn, A., J. Smith and A. van Soest (2011) Are Americans Really Less Happy with Their Incomes? *Discussion Paper* 2011-059, Tilburg University, Center for Economic Research.

Kassenboehmer, S. and J. Haisken-Denew (2008) Heresy or enlightenment? The wellbeing age U-shape is really flat! *mimeo*, RWI Essen.

Keisuke, H., g. Imbens, g. Ridder, D. Rubin (2001) Combining Panel Data Sets with Attrition and Refreshment Samples. *econometrica* 69, 1645-1659.

Kroh, M. (2011) Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2010). *Data Documentation* 59, Deutsches Institut für Wirtschaftsforschung.

Krueger, A. and D. Schkade (2008) The Reliability of Subjective Well-Being Measures. *Journal of Public Economics* 92, 1833-1845.

Lips. O. (2007) Attrition in the Swiss Household Panel. *Methoden - Daten - Analysen* 1, 45-68.

Luttmer, E. (2005) Neighbours as Negatives: Relative Earnings and Well-Being. *Quarterly Journal of Economics* 120, 963-1002.

McBride, M. (2010) Money, Happiness, and Aspirations: An Experimental Study. *Journal of Economic Behavior and Organization* 74, 262-276.

Metcalf, R., N. Powdthavee and P. Dolan (2011) Destruction and Distress: Using a Quasi-Experiment to Show the Effect of the September 11 Attacks on Mental Well-Being in the United Kingdom. *Economic Journal* 121, F81-F103.

Oreopoulos, P. (2007) Do Dropouts Drop Out Too Soon? Wealth, Health and Happiness from Compulsory Schooling. *Journal of Public Economics* 91, 2213-2229.

Oswald, A. (2008) On the Curvature of the Reporting Function from Objective Reality to Subjective Feelings. *Economics Letters*, 100, 369-72.

Oswald, A. and N. Powdthavee (2008a) Death, Happiness, and the Calculation of Compensatory Damages. *Journal of Legal Studies* 37, S217-S252.

Oswald, A. and N. Powdthavee (2008b) Does Happiness Adapt? A Longitudinal Study of Disability with Implications for Economists and Judges. *Journal of Public Economics* 92, 1061-1077.

Oswald, A. and S. Wu (2010) Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the U.S.A. *Science* 327, 576-580.

Pezzini, S. (2005) The Effect of Women's Rights on Women's Welfare: Evidence from a Natural Experiment. *Economic Journal* 115, pages C208-C227.

Powdthavee, N. (2008) Putting a Price Tag on Friends, Relatives, and Neighbours: Using Surveys of Life Satisfaction to Value Social Relationships. *Journal of Socio-Economics* 37, 1459-1480.

Ravallion, M. and M. Lokshin (2010) Who Cares About Relative Deprivation? *Journal of Economic Behavior and Organization* 73, 171-185.

Sacks, D., B. Stevenson and J. Wolfers (2010) Subjective Well-Being, Income, Economic Development and Growth. *NBER Working Papers* 16441, National Bureau of Economic Research.

Sanfey, P. and U. Teksoz (2007) Does transition make you happy? *Economics of Transition* 15, 707-731.

Senik, C. (2004) When Information Dominates Comparison: Learning From Russian Subjective Panel Data. *Journal of Public Economics* 88, 2099-2123.

Senik, C. (2008a) Ambition and Jealousy: Income Interaction in the “Old” Europe versus the “New” Europe and the United States. *Economica* 75, 495-513.

Senik, C., 2008b. Is Man Doomed to Progress? *Journal of Economic Behavior and Organization* 68, 140-152.

Senik, C. (2009) Direct Evidence on Income Comparisons and Their Welfare Effects. *Journal of Economic Behavior and Organization* 72, 408-424.

SgROI, d., E. Proto, A. Oswald and A. Dobson (2010) Priming and the Reliability of Subjective Well-Being Measures. *Warwick Economic Research Papers* No. 935.

Stevenson, B. and J. Wolfers (2008) Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox. *NBER Working Papers* 14282, National Bureau of Economic Research.

Stevenson, b. and J. Wolfers (2009) The Paradox of Declining Female Happiness. *American Economic Journal: Economic Policy* 1, 190-225.

Toepoel, V., J. Das and A. van Soest (2009) Relating Question Type to Panel Conditioning: A Comparison between Trained and Fresh Respondents. *Survey Research Methods* 3, 73-80.

Van Landeghem, B. (2012) A Test for the Convexity of Human Well-Being over the Life Cycle: Evidence from a 20-Year Panel. *Journal of Economic Behavior and Organization* 81, 571-582.

Wagner, G., J. Frick and J. Schupp (2007) The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements. *Schmollers Jahrbuch* 127, 139-169.

Wolfer, J. (2003) Is Business Cycle Volatility Costly? Evidence from Surveys of Subjective Well-Being. *International Finance* 6, 1-26.

Table 1: Differences in Well-Being between Refreshment Samples and Experienced Samples

	West Germany	East Germany	Switzerland
$E[WB_{1998}^{re}] - E[WB_{1998}^{exp}]$	0.53	0.40	
$E[WB_{2000}^{re}] - E[WB_{2000}^{exp}]$	0.30	0.28	
$E[WB_{2004}^{re}] - E[WB_{2004}^{exp}]$			0.16
$E[WB_{2006}^{re}] - E[WB_{2006}^{exp}]$	0.50	0.68	

Source: German Socio-Economic Panel and Swiss Household Panel.

$WB$  = well-being, subscripts denote the survey year, superscripts  $RE$  and  $EXP$  denote ‘refreshment sample’ and ‘experienced sample’ respectively. All differences turn out to be significantly different from 0 at any significance level.

Table 2: Differences in Real Household Income between Refreshment Samples and Experienced Samples (Expressed in %)

	West Germany	East Germany	Switzerland
$E[Y_{1998}^{re}] - E[Y_{1998}^{exp}]$	-9.6***	-16.7 <sup>[***]</sup>	
$E[Y_{2000}^{re}] - E[Y_{2000}^{exp}]$	-6.8***	-6.5***	
$E[Y_{2004}^{re}] - E[Y_{2004}^{exp}]$			-6.2***
$E[Y_{2006}^{re}] - E[Y_{2006}^{exp}]$	-0.2	-0.4	

Source: German Socio-Economic Panel and Swiss Household Panel.

$Y$  = real household income, subscripts denote the survey year, superscripts  $RE$  and  $EXP$  denote 'refreshment sample' and 'experienced sample' respectively. One to three stars denote significance at the 10%, the 5% and the 1% significance level respectively.

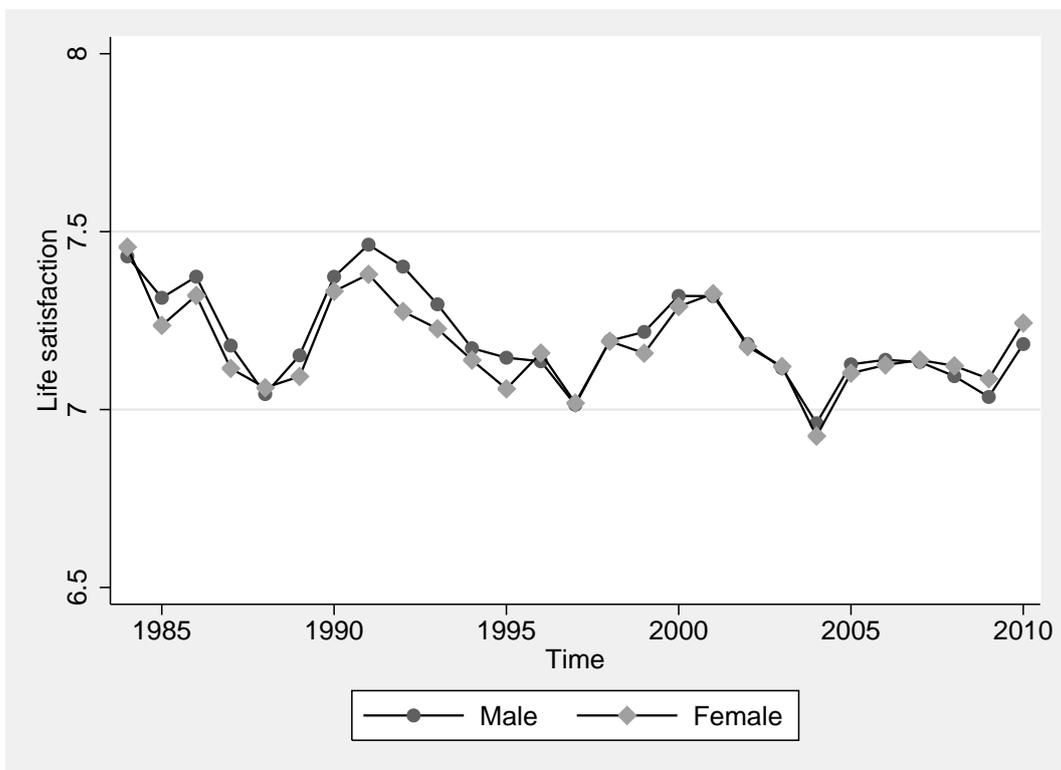
Table 3: Differences in Well-Being between Refreshment Samples and Experienced Samples: Addressing Attrition Bias

	West Germany	East Germany	Switzerland
$E[WB_{1998}^{re}] - E[WB_{1998}^{exp}]$	0.56	0.45	
$E[WB_{2000}^{re}] - E[WB_{2000}^{exp}]$	0.41	0.32	
$E[WB_{2004}^{re}] - E[WB_{2004}^{exp}]$			0.21
$E[WB_{2006}^{re}] - E[WB_{2006}^{exp}]$	0.60	0.68	

Source: German Socio-Economic Panel and Swiss Household Panel.

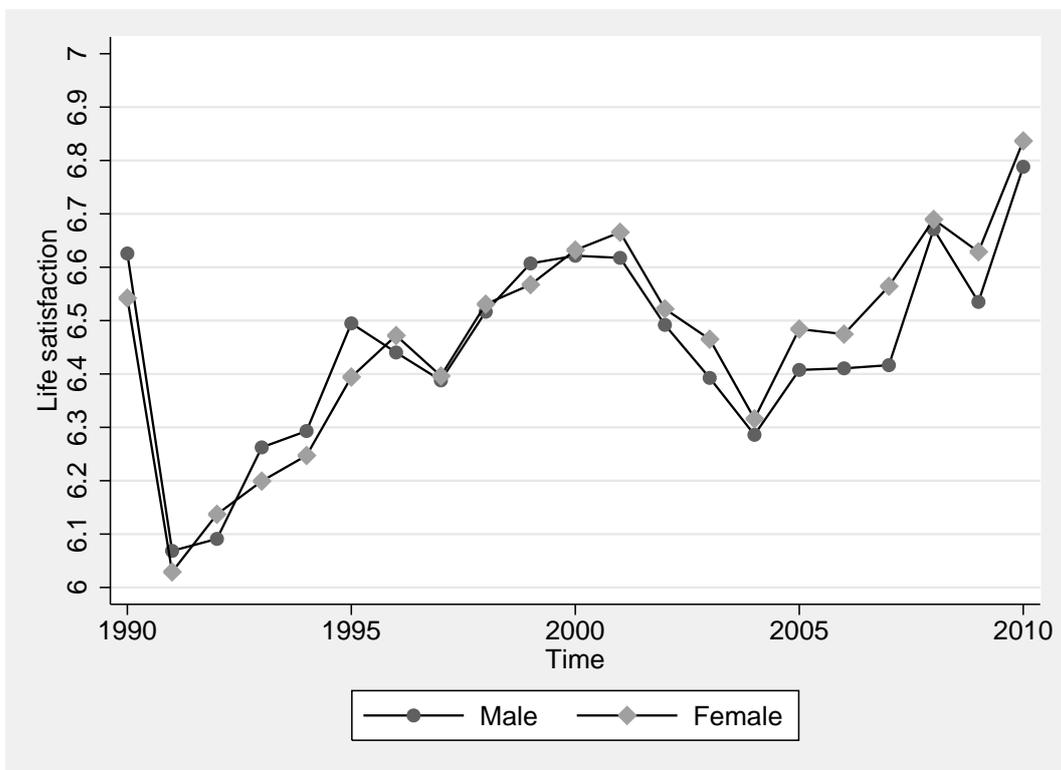
$WB$  = well-being, subscripts denote the survey year, superscripts  $RE$  and  $EXP$  denote ‘refreshment sample’ and ‘experienced sample’ respectively. All differences turn out to be significantly different from 0 at any significance level. As for the refreshment sample, only those are considered who will remain in the panel for at least 3 more years.

Figure 1: Trends of Life Satisfaction over Time: West Germany



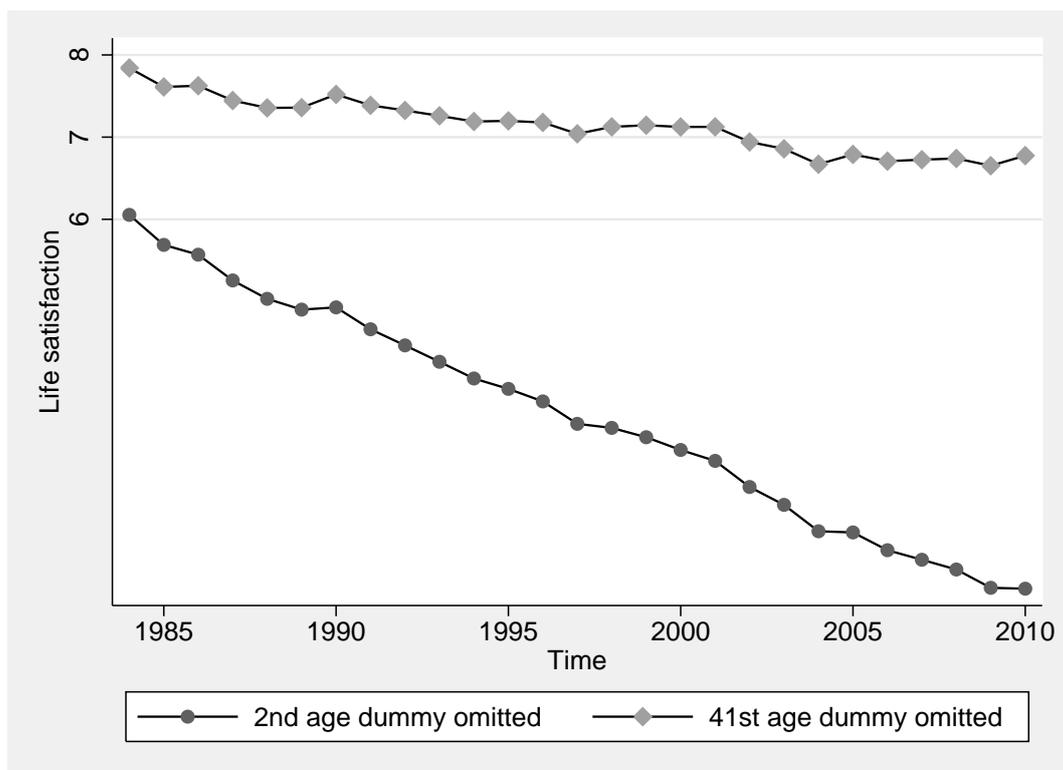
Source: German Socio-Economic Panel.

Figure 2: Trends of Life Satisfaction over Time: East Germany



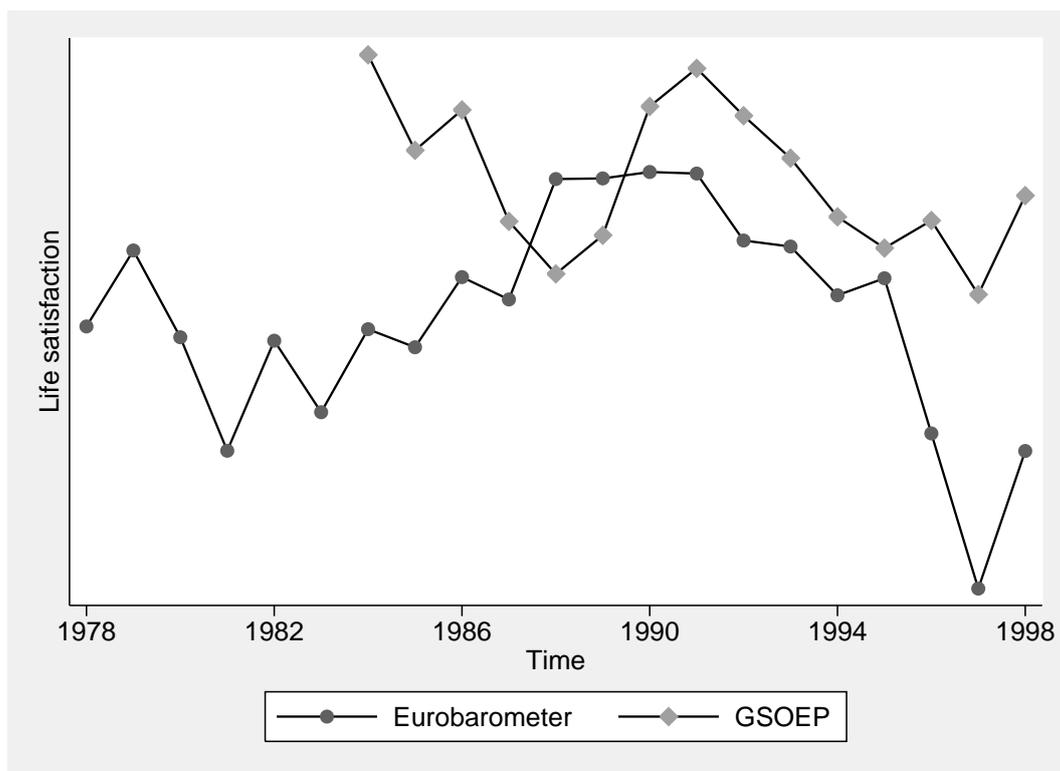
Source: German Socio-Economic Panel.

Figure 3: Trends of Life Satisfaction over Time: Illustrating Biases in Fixed Effects Using German Data



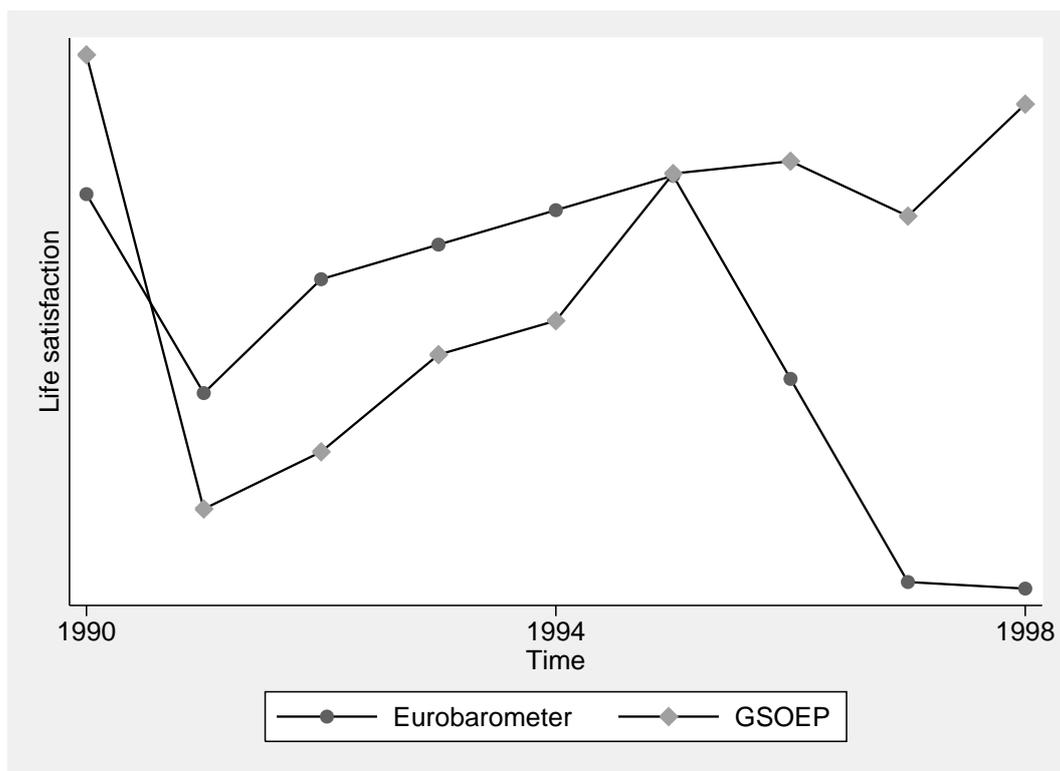
Source: German Socio-Economic Panel.

Figure 4: A Comparison between Trends in Standardized Life Satisfaction in the GSOEP and the Eurobarometer: West Germany



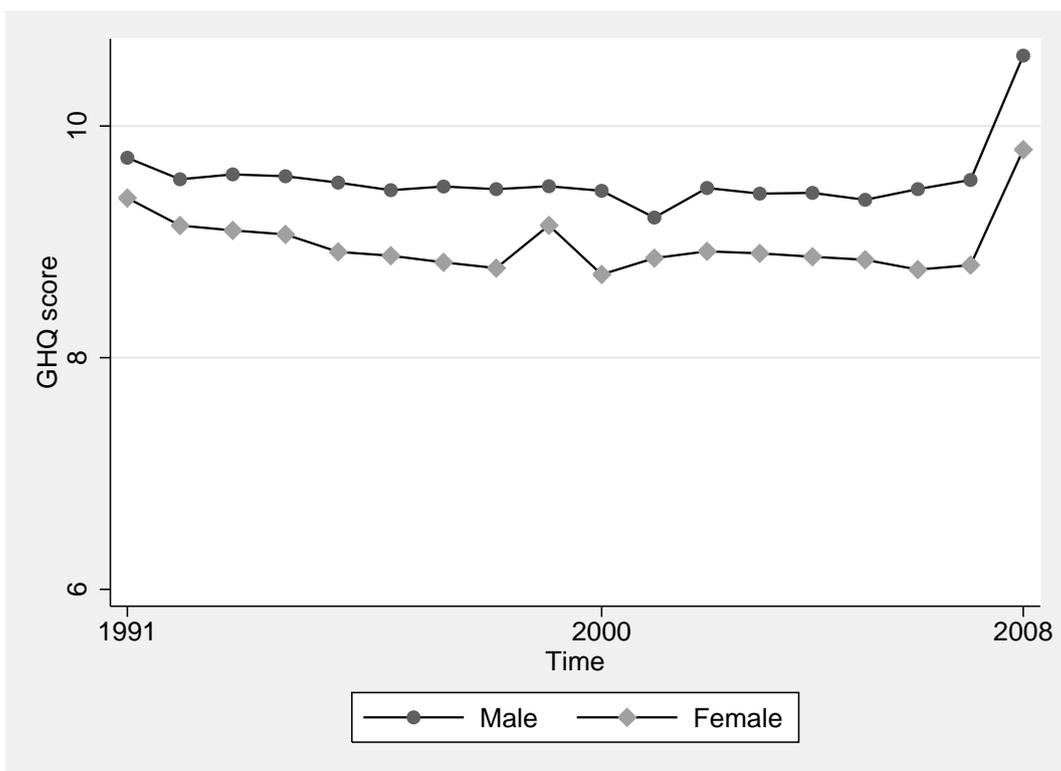
Source: German Socio-Economic Panel and Eurobarometer Survey.

Figure 5: A Comparison between Trends in Standardized Life Satisfaction in the GSOEP and the Eurobarometer: East Germany



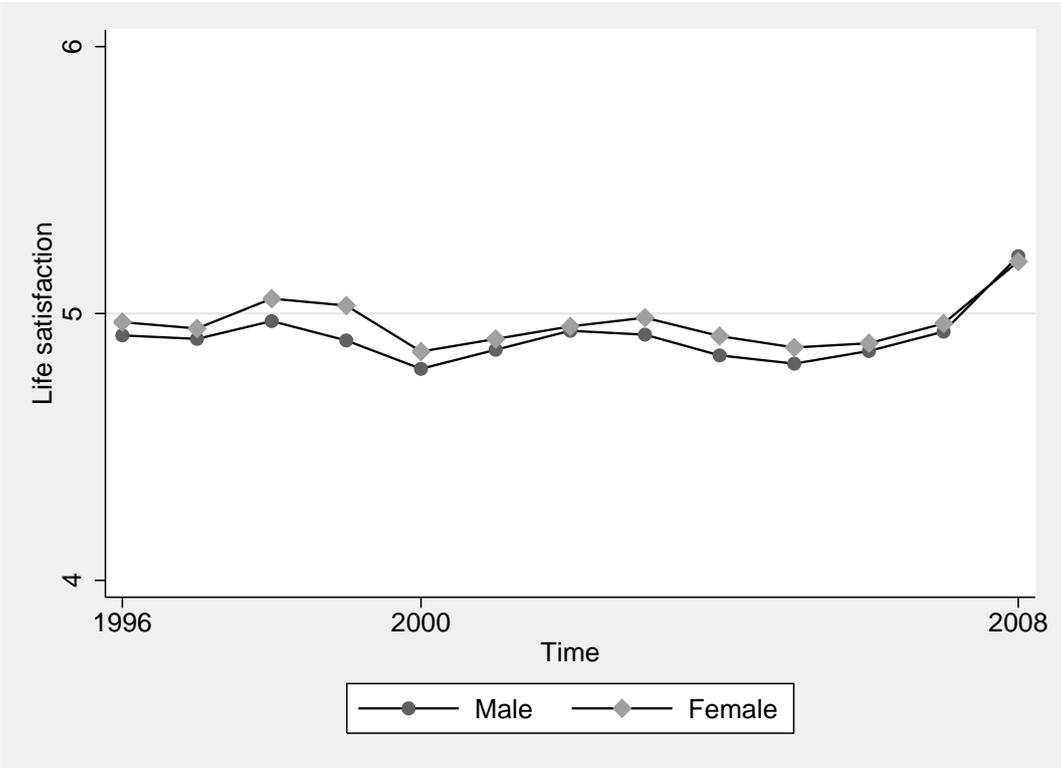
Source: German Socio-Economic Panel and Eurobarometer Survey.

Figure 6: Trends of GHQ Scores over Time: Britain



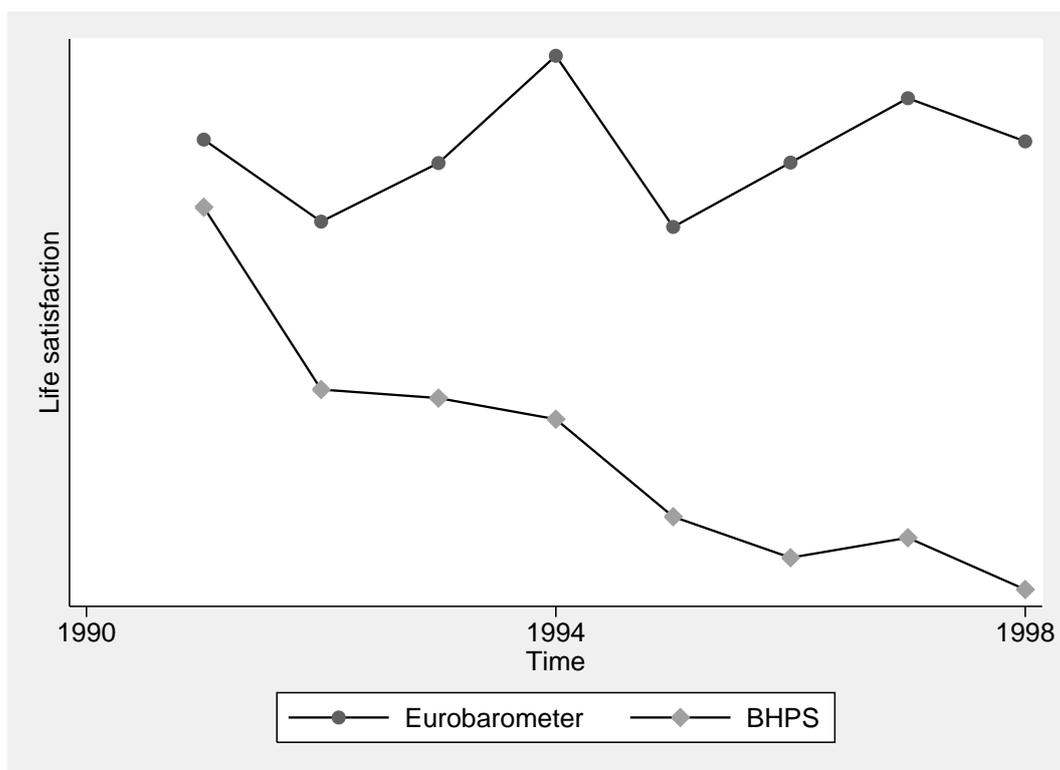
Source: British Household Panel Survey.

Figure 7: Trends of Life Satisfaction Scores over Time: Britain



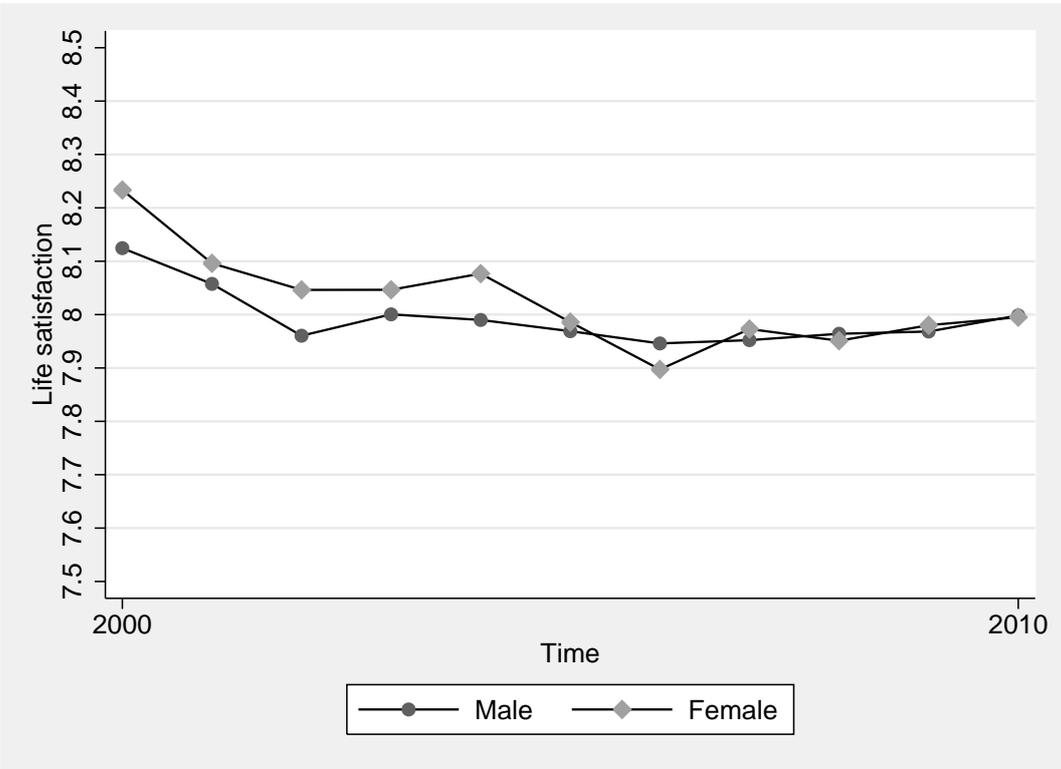
Source: British Household Panel Survey.

Figure 8: A Comparison between Trends in Standardized GHQ Scores in the BHPS and Standardized Life Satisfaction Scores in the Eurobarometer



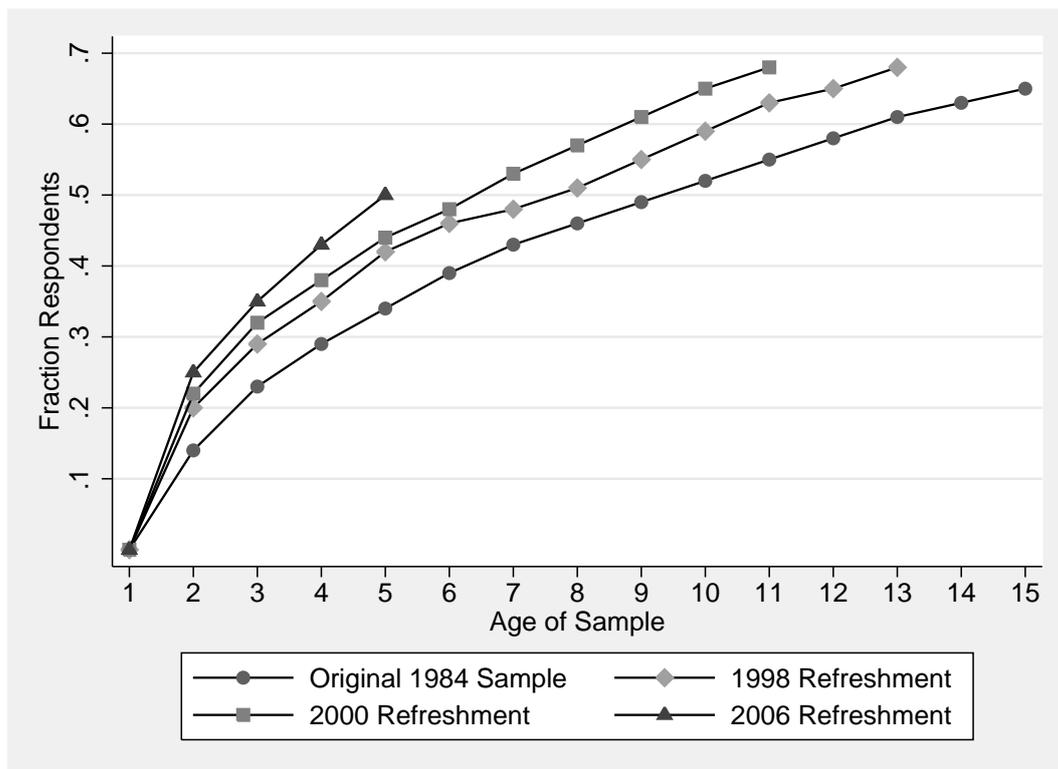
Source: British Household Panel Survey and Eurobarometer Survey.

Figure 9: Trends of Life Satisfaction Scores over Time: Switzerland



Source: Swiss Household Panel.

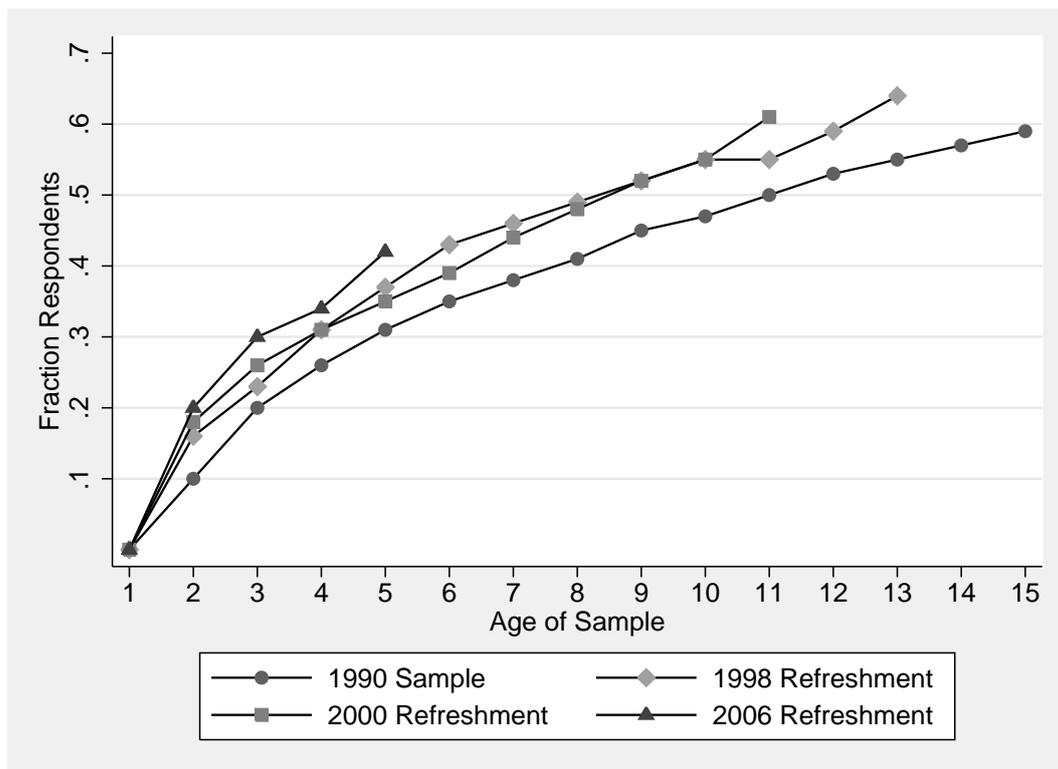
Figure 10: The Attrition Rate of Respondents to Subjective Well-Being Questions over a Sample's Life Cycle: West Germany



Source: German Socio-Economic Panel.

An individual is regarded as an attritor at age  $n$  of the panel if the individual was a respondent at age 1 and if the number of times the individual responded to a subjective well-being question is strictly less than  $n$ .

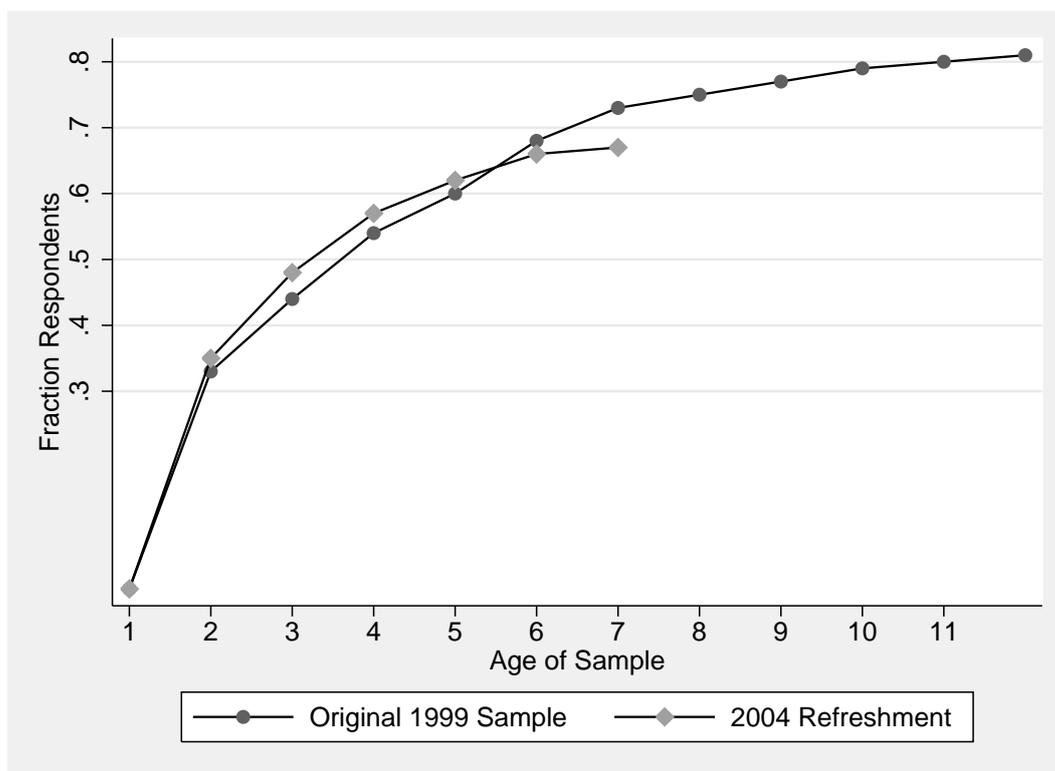
Figure 11: The Attrition Rate of Respondents to Subjective Well-Being Questions over a Sample's Life Cycle: East Germany



Source: German Socio-Economic Panel

An individual is regarded as an attritor at age  $n$  of the panel if the individual was a respondent at age 1 and if the number of times the individual responded to a subjective well-being question is strictly less than  $n$ .

Figure 12: The Attrition Rate of Respondents to Subjective Well-Being Questions over a Sample's Life Cycle: Switzerland



Source: Swiss Household Panel

An individual is regarded as an attritor at age  $n$  of the panel if the individual was a respondent at age 1 and if the number of times the individual responded to a subjective well-being question is strictly less than  $n$ .