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

Bert Van Landeghem
Maastricht University, ROA
b.vanlandeghem@maastrichtuniversity.nl

May 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 between refreshment subsamples with experienced respondents within the same panel study, as well as the comparison of time trends derived from panel data and repeated cross-sections within the same country, suggest a substantial panel effect. The analysis proposes a way to take into account panel attrition in cases where the attrition rate is substantial, and an attempt is made to estimate how the panel effect builds up over different interviews. Self-reported well-being data seem to gain increasing attention in the world of both academics as well as policy makers, and the awareness of a potential panel effect is important to interpret results (e.g. in the current debate on trends in life satisfaction over time). The results might encourage further research in panel conditioning for other important data in economics gathered through household surveys.

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[1] 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 (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., 2012; 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 self-reported life satisfaction might be a concave function of *true* life satisfaction. 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

[1] Terms such as happiness, life satisfaction, subjective well-being are often used interchangeably in the literature.

survey, 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 authors conclude that, 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 (2012) use vignettes to address differences in anchoring across countries, and find for older individuals in 11 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 a 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 generally no attempt is made to distinguish panel conditioning from panel 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, and hence explain why recent life course studies using the same data sometimes report different outcomes (see e.g. Clark and Oswald, 2006 versus Frijters and Beatton, 2012 and Gerstorf et al., 2010).

This paper aims to investigate the existence of panel conditioning in several ways. Reported life satisfaction of refreshment samples is compared with reported life satisfaction of more experienced samples within the same panel dataset. Issues of panel ageing and panel attrition, which might cloud results, will be addressed. As a robustness check, trends in panel data are compared with trends in repeated cross-sections within the same country.

The findings of this paper might be important in the 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), document how in the United States, over the last three 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. If indeed self-reported life satisfaction data are subject to panel conditioning, it would imply that trends derived from panel data are likely to be clouded, and that at least panel data and repeated cross-sections cannot be used interchangeably. Blanchflower and Oswald (2004), for example, derive trends in the United States from repeated cross-sections, while they derive trends for Britain from panel data. Stevenson and Wolfers (2009) use repeated cross-sections during their analysis, but refer to an unreported robustness check in German panel data.

It seems obvious that the existence of a panel effect could also be important information for researchers being interested in how macroeconomic indicators are correlated with subjective well-being (Di Tella et al. (2001, 2003) and Wolfers (2003). In particular, one can think of the renewed 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 documents the data that will be used, and outlines the core of the empirical strategy. Section three presents and discusses the econometric results, and Section four concludes.

2 Data and Empirical Strategy

2.1 Panel Data

The study draws upon three panel datasets, the German Socioeconomic Panel (GSOEP), the British Household Panel Survey (BHPS) and the Swiss Household Panel (SHP). 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. 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 panel. 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.

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 the most common panel dataset 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[2] different samples:

- Sample A (started in 1984) represents the West-German population, while sample C (started in 1990) represents the East German population.

[2]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.

- Samples E, F and H (started in 1998, 2000 and 2006 respectively) are refreshment samples
- 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 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 Questionaire (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 (labeled 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.

2.2 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 change 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 panel 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 survey 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.[3] 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.

2.3 empirical Strategy

As is clear from Section 2.1, two out of the three panel datasets contain refreshment samples in the strict sense, i.e., new samples drawn from the same population from which the original sample was drawn. Refreshment samples are often added in the course of a panel study to increase the sample size, allowing for a wider possibility of analyses, or at least to mitigate the decrease in sample size due to panel attrition.

Such refreshment samples can help us to identify whether or not answers on (subjective well-being) survey questions are prone to panel conditioning. Indeed, they allow us to compare the average responses between a group of first-time respondents on the one hand, and more experienced respondents on the other hand. One may of course be sceptical about whether or not the group of new respondents is comparable to the

[3]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.

group of more experienced respondents, and that the results are clouded by panel attrition. One might consider the strategy of Das Et al. (2011), who 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. Das et al. (2011) point out that, if no assumptions regarding the attrition process are to be made, one can still identify upper and lower bounds of the panel effect. They illustrate this with dichotomous responses, and calculate bounds by assuming that all attritors would have chosen either 0 or 1. For life satisfaction data, however, calculating bounds seems less useful. The range of possible ordinal responses is much larger than just 2 options, and moreover, the attrition rate in the considered data sets is relatively high (see below). In an attempt to address this problem, we can make 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 n more years. Hence, these refreshment subsample should be more comparable with the experienced samples.

Second, as a robustness check, the paper compares trends in panel data with trends in cross-sectional data within the same country. the regressions estimating trends in data are kept very parsimonious, and do not contain any other variables than calendar year dummies. In particular, including age dummies and fixed effects, which are standard controls in happiness equations, would lead to estimates that are very hard to interpret, and seemingly mild parametric restrictions to obtain identification can lead to rotations of the estimated path of well-being over time if the underlying assumptions are not entirely correct (see e.g. Van Landeghem, 2012 Appendix 1 for an algebraic example).

3 Results

3.1 A Comparison Between New Respondents and Experienced Respondents Within a Panel

This subsection documents the difference between average subjective well-being scores of a refreshment sample and the average scores of more experienced respondents for that same year within the same panel dataset. 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 sample in the corresponding calendar year, are represented in Table 1 for West Germany, East Germany, and Switzerland. 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 three refreshment samples in the years 1998, 2000 and 2006, while the Swiss Household Panel introduced 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 statistically significant at any significance level, and the magnitude is substantial, varying between 0.16 and 0.68 on a 0-10 scale. For completeness, Table 2 and 3 report results separately for men and women, respectively, and show that the results hold for both sexes. One might be worried that the refreshment samples are nevertheless substantially different from the more experienced samples. For example, Table 4 shows us probit equations for data of each survey year a refreshment sample was started, and which predict whether or not an individual belongs to a refreshment sample. The results from the regressions make clear that socioeconomic characteristics are not always completely randomly distributed across the refreshment sample and the experienced sample, either because of sampling error, attrition biases or because the explanatory variables are subject to panel conditioning themselves. The signs and size of the coefficients on the variables predicting the likelihood of belonging to a refreshment sample are, however, not very consistent across the different regressions. One common trend across the different regressions seems to be that singles are more likely to be in the refreshment sample than in the more experience sample. This is not very unlogical, since singles are more mobile and thus more prone to quit the panel. It also seems that, across the different regressions, individuals with lower household income are more likely to be in the refreshment sample. We have a closer look at the distribution of this latter variable in Table 5, which shows the difference

between real household income in the refreshment subsample on the one hand, and in the experienced sample on the other hand, using weights to correct for differences in age distributions between both samples. The results suggest that differences in average income between the refreshment samples and experienced samples is rather small (see Table 5). Real household income in the refreshment samples for Germany are on average 0 to 17% 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 individuals in the refreshment sample have, on average, a lower income than individuals in the more experienced sample, is in line with the observation in the literature that people with lower income are likely to exit the panel.

Moreover, as pointed out in Section 2.3, the refreshment samples are started well before the last round of data. In order to make experienced respondents more comparable with the fresh respondents, Table 6 compares the experienced samples only with those individuals from the refreshment samples who will at least be in the panel for three more years. The results of this exercise 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 1, 2, and 3 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 Socioeconomic Panel, but this goes for both the original 1999 sample as for the 2004 refreshment sample.

Finally, one might be interested in the actual size of the panel effect as well as in the dynamics. Is a panel effect only of importance between the first and second interview, or does it cumulate across several interview rounds? To investigate this issue, a subsample of individuals is selected from both the GSOEP and the SHP, which are, at least re-interviewed in the four consecutive years after the initial interview in which the question measuring subjective well-being was asked.[4] Life satisfaction is regressed on five dummies taking the value of one if an individual is being interviewed for the first, second ...or fifth time respectively. The baseline category consists of those who are being interviewed for more than five times. Some calendar years will have much more new respondents than average years, in particular those years in which a new sample was started. Hence, time dummies are included as controls to address the problem of nonrandom distribution of newcomers across calendar years. Since the regressions are run on a subsample of individuals who are at least in the sample for 5 consecutive survey rounds, we do not need to worry that the path of the panel effect over the different survey rounds (from the first to the fifth) is clouded by panel attrition.

Regression output is displayed in Table 7 for West Germany, East Germany and Switzerland respectively. The regression results show us that the panel effect is not entirely established between being interviewed for the first and seconde time, but that it accumulates over the different survey rounds. In West Germany, there is a substantial panel effect from the second to the fifth interview, which is also statistically significant at any conventional significance level.[5] There is an estimated panel effect of -0.13 during the second as well as third interview, a panel effect of -0.1 during the fourth interview and of -0.05 during the fifth interview. For East Germany, we see a rather similar pattern of negative panel effects reoccurring from interview to interview, although the panel effect during the fifth interview is not statistically different from zero at conventional significance levels. The pattern for Switzerland is in line with that of Germany, but no

[4]In the SHP, the first interview for an individual does not necessarily equal to the interview in which he was first asked about his life satisfaction, as a question on subjective well-being was only introduced in the second round of the panel.

[5]To be clear, a panel effect for interview n is calculated by subtracting the coefficient on the dummies for interview $n-1$ and n.

panel effect is measured any more after the third interview, and a panel effect is only statistically significant (be it at any conventional level) for the second interview.

The coefficient on the dummy for being interviewed for the fifth time can be interpreted as the negative of a residual panel effect, i.e. which will be established over the interviews after the fifth has taken place. The coefficient on this dummy for Switzerland is nearly equal to 0, while for West and East Germany, this coefficient still has a substantial magnitude of 0.25 and 0.32 respectively. The reason why the cumulation path of panel effects for Switzerland is much shorter than for Germany is rather speculative, but one should note that in the SHP, 90% of people answering the well-being question for the first time are situated in either 2000 or 2004. Moreover, one should keep in mind that , contrary to the path of panel effects from interview one to five, the estimate for a residual panel effect might be slightly clouded by panel attrition. Indeed, some respondents will not be interviewed any more after the fifth interview while others will remain in the panel for many years.

3.2 A Comparison between Time Trends in Panel Data and Repeated Cross-Sections

Comparing time trends in panel data with the repeated cross-sectional data within the same country is another, though far from ideal strategy to identify the existence of 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) and it seems difficult to disentangle a panel effect from an attrition effect. Nevertheless, it seems useful to check whether the results from this exercise are in line with the results described in the previous subsection.

Figure 4 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 5 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.

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 6 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 7 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.

Figure 8 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.[6] 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 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 9 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. 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 10 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 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.

[6]This finding is not just limited to Britain, as do suggest unreported results from the Panel Study of Belgian Households.

Life satisfaction time trends derived from the Swiss Household Panel are shown in Figure 11 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 cyclicality 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 nor 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 Conclusion

This paper has investigated whether subjective well-being data are subject to panel conditioning, i.e. whether answers on questions measuring subjective well-being depend on having participated in the panel before. The analysis has made use from panel data as well as repeated cross-sectional data for three European countries.

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

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., 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 (2012) 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 (2012) Shattered Dreams: The Effects of Changing the Pension System Late in the Game, *The Economic Journal* 122, 1-25.

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.

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

Gerstorf, 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.

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

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 fur 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.

Metcalfe, 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.

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.

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.

Wolfers, 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
$E[WB^{re}_{1998}] - E[WB^{exp}_{1998}]$	0.53	0.40	
$E[WB^{re}_{2000}] - E[WB^{exp}_{2000}]$	0.30	0.28	
$E[WB^{re}_{2004}] - E[WB^{exp}_{2004}]$			0.16
$E[WB^{re}_{2006}] - E[WB^{exp}_{2006}]$	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 Well-Being between Refreshment Samples and Experienced Samples: Male Subsample

	West Germany	East Germany	Switzerland
$E[WB^{re}_{1998}] - E[WB^{exp}_{1998}]$	0.61	0.38	
$E[WB^{re}_{2000}] - E[WB^{exp}_{2000}]$	0.35	0.31	
$E[WB^{re}_{2004}] - E[WB^{exp}_{2004}]$			0.11
$E[WB^{re}_{2006}] - E[WB^{exp}_{2006}]$	0.41	0.57	

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, except for the Swiss subsample where the difference of 0.11 is only statistically different from 0 at the 3% significance level.

Table 3: Differences in Well-Being between Refreshment Samples and Experienced Samples: Female Subsample

	West Germany	East Germany	Switzerland
$E[WB^{re}_{1998}] - E[WB^{exp}_{1998}]$	0.52	0.42	
$E[WB^{re}_{2000}] - E[WB^{exp}_{2000}]$	0.39	0.25	
$E[WB^{re}_{2004}] - E[WB^{exp}_{2004}]$			0.20
$E[WB^{re}_{2006}] - E[WB^{exp}_{2006}]$	0.46	0.77	

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 4: Exploring the Determinants of the Propensity to Belong to a Refreshment Sample: Results for West Germany, East Germany and Switzerland

	West 1998	East 1998	West 2000	East 2000	West 2006	East 2006	Switzerland 2004
log real hh. income	-0.043***	-0.027**	-0.094***	-0.011	-0.002	-0.010	-0.051***
	(0.012)	(0.011)	(0.012)	(0.019)	(0.011)	(0.011)	(0.012)
disabled	-0.056***	0.016	-0.070***	-0.005	-0.008	-0.014	0.020
	(0.015)	(0.018)	(0.017)	(0.026)	(0.012)	(0.015)	(0.013)
household size	-0.000	-0.015**	0.033***	0.006	0.009*	-0.008	-0.030***
	(0.006)	(0.006)	(0.006)	(0.009)	(0.005)	(0.006)	(0.006)
low education	-0.008	0.167***	-0.074***	0.094***	-0.033*	0.007	0.053***
	(0.017)	(0.040)	(0.018)	(0.029)	(0.019)	(0.027)	(0.020)
medium education	-0.046***	0.055***	0.017	0.234***	-0.056***	-0.050***	0.006
	(0.015)	(0.011)	(0.014)	(0.017)	(0.014)	(0.017)	(0.021)
unemployed	-0.066***	0.002	-0.018	-0.016	0.006	0.006	0.043
	(0.024)	(0.013)	(0.032)	(0.022)	(0.020)	(0.015)	(0.043)
single	0.013	0.040***	0.021	-0.062***	-0.017	0.027*	-0.047**
	(0.017)	(0.013)	(0.017)	(0.024)	(0.015)	(0.015)	(0.019)
widowed	-0.045**	-0.037***	-0.046**	-0.020	0.033	-0.004	-0.012
	(0.019)	(0.013)	(0.022)	(0.031)	(0.021)	(0.021)	(0.032)
divorced	-0.043**	-0.008	-0.047**	0.048*	-0.002	-0.012	-0.037*
	(0.020)	(0.015)	(0.022)	(0.027)	(0.016)	(0.015)	(0.022)
female	-0.019*	-0.003	-0.004	0.024*	-0.023**	0.001	-0.040***
	(0.012)	(0.009)	(0.011)	(0.014)	(0.011)	(0.011)	(0.012)
Observations	5,049	3,569	9,520	5,169	8,439	4,581	6,707

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The table shows marginal probabilities, computed at the average value of independent variables, derived from Probit Equations. Regressions are run for each survey year in which a refreshment sample was started. The dependent variable is dichotomous and takes the value one whenever an individual belongs to the refreshment sample started in that particular year.

Table 5: Differences in Real Household Income between Refreshment Samples and Experienced Samples (Expressed in %)

	West Germany	East Germany	Switzerland
$E[Y^{re}_{1998}] - E[Y^{exp}_{1998}]$	-9.6***	-16.7[* **]	
$E[Y^{re}_{2000}] - E[Y^{exp}_{2000}]$	-6.8***	-6.5***	
$E[Y^{re}_{2004}] - E[Y^{exp}_{2004}]$			-6.2***
$E[Y^{re}_{2006}] - E[Y^{exp}_{2006}]$	-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 6: Differences in Well-Being between Refreshment Samples and Experienced Samples: Addressing Attrition Bias

	West Germany	East Germany	Switzerland
$E[WB^{re}_{1998}] - E[WB^{exp}_{1998}]$	0.56	0.45	
$E[WB^{re}_{2000}] - E[WB^{exp}_{2000}]$	0.41	0.32	
$E[WB^{re}_{2004}] - E[WB^{exp}_{2004}]$			0.21
$E[WB^{re}_{2006}] - E[WB^{exp}_{2006}]$	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 individuals are considered who will remain in the panel for at least 3 more years.

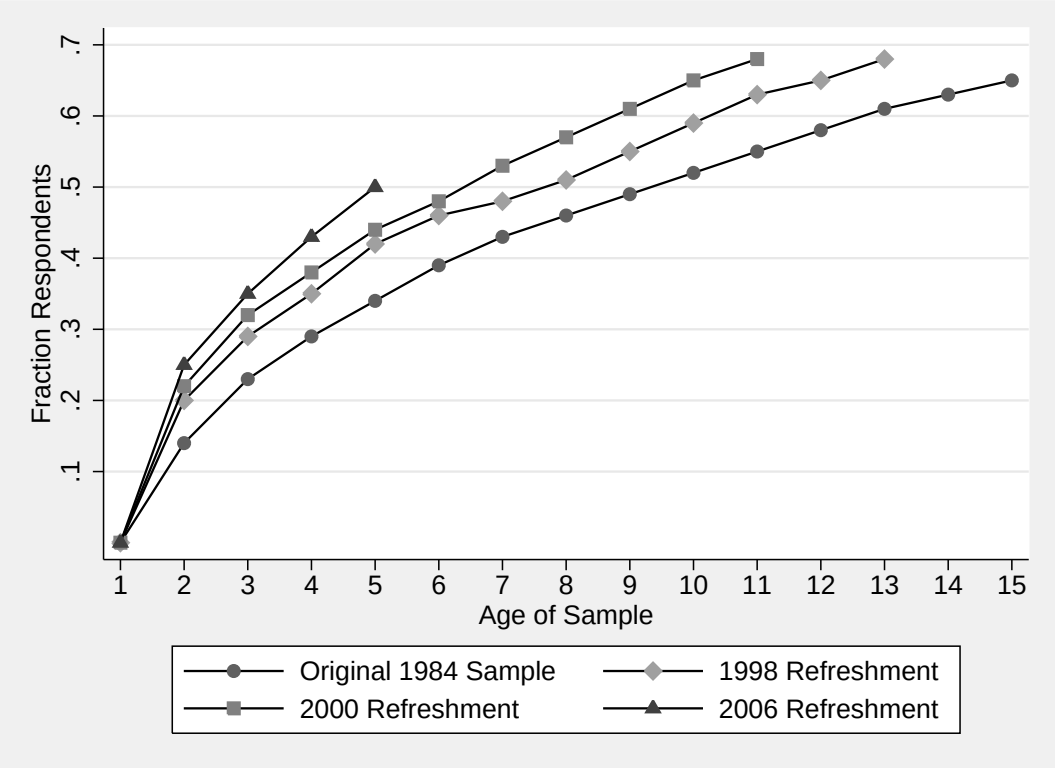
Table 7: The Path of Panel Effects over Interviews: Results for West Germany, East Germany and Switzerland

	(1)	(2)	(3)
VARIABLES	sat	sat	sat
1st Interview	0.667***	0.744***	0.234***
	(0.019)	(0.038)	(0.055)
2nd Interview	0.530***	0.546***	0.103**
	(0.018)	(0.036)	(0.043)
3rd Interview	0.399***	0.467***	0.056
	(0.017)	(0.034)	(0.039)
4th Interview	0.301***	0.357***	0.032
	(0.017)	(0.032)	(0.037)
5th Interview	0.246***	0.321***	0.042
	(0.016)	(0.032)	(0.032)
Constant	6.864***	5.506***	8.012***
	(0.033)	(0.055)	(0.065)
Observations	288,800	80,757	50,217
R-squared	0.015	0.019	0.005

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

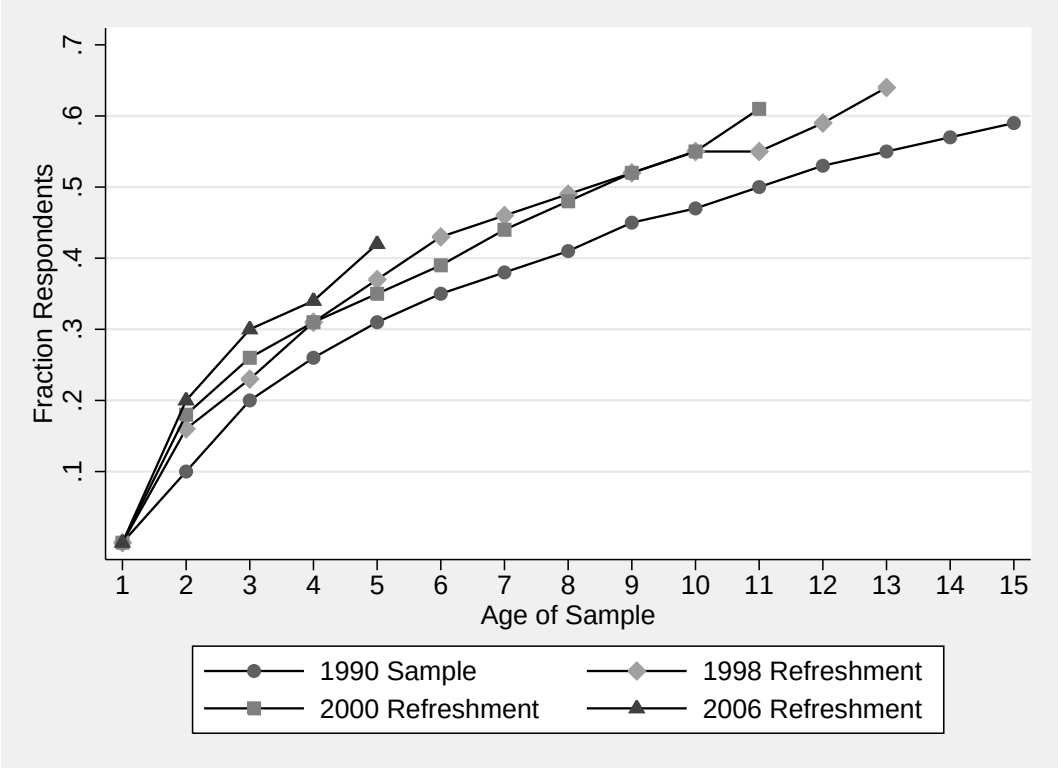
Source: German Socio-Economic Panel and Swiss Household Panel.
Regressions are run on individuals who have at least answered the subjective questions in the four survey rounds following the interview in which they were first asked the subjective well-being question.

Figure 1: The Attrition Rate over a Sample's Life Cycle: West Germany



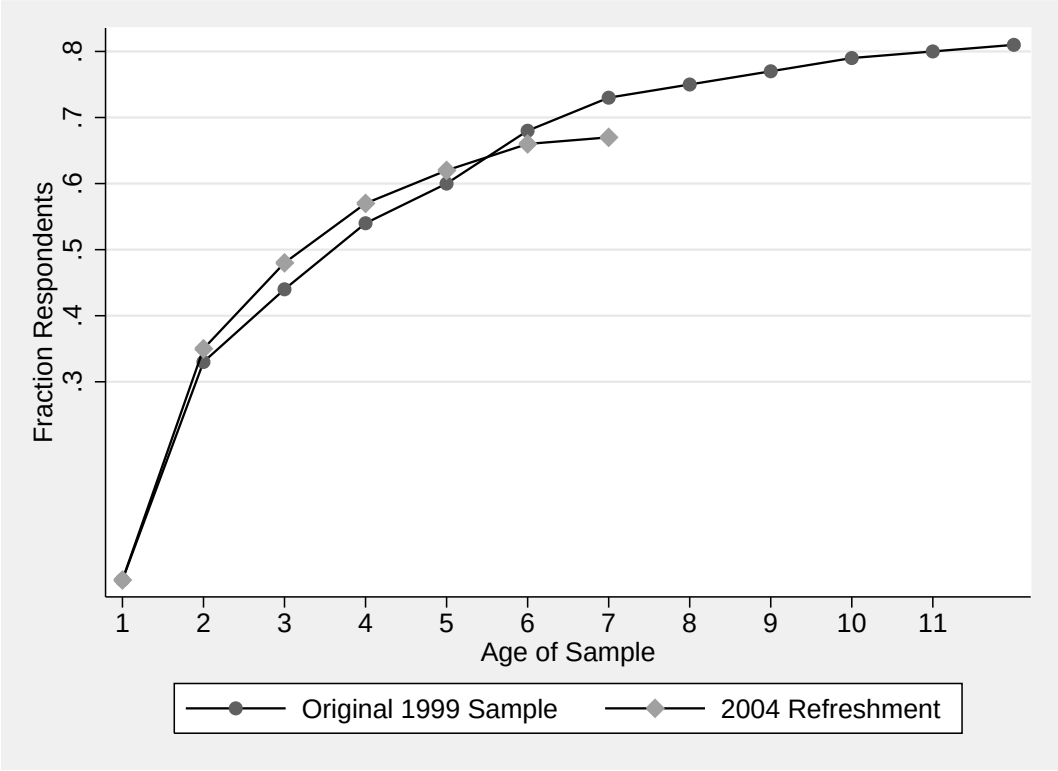
Source: German Socio-Economic Panel.
An individual is regarded as an attritor at age n of the panel if the individual was a respondent at age 1 and if the number of times the individual responded to a subjective well-being question is strictly less than n.

Figure 2: The Attrition Rate over a Sample's Life Cycle: East Germany



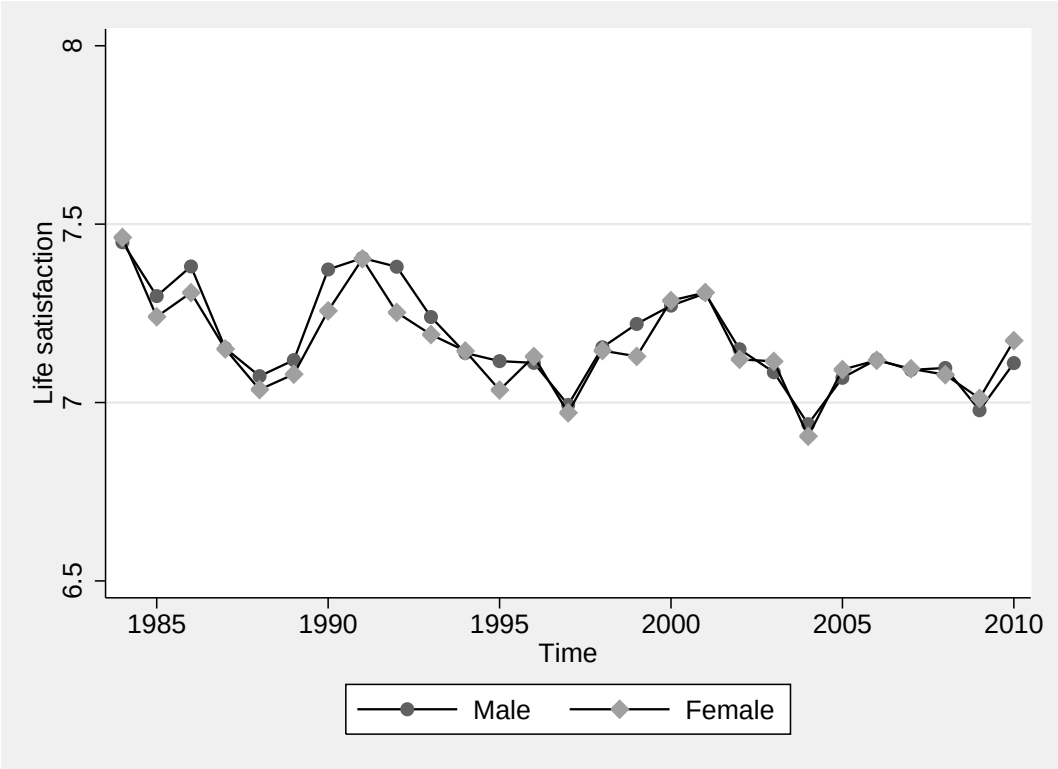
Source: German Socio-Economic Panel
An individual is regarded as an attritor at age n of the panel if the individual was a respondent at age 1 and if the number of times the individual responded to a subjective well-being question is strictly less than n.

Figure 3: The Attrition Rate over a Sample's Life Cycle: Switzerland



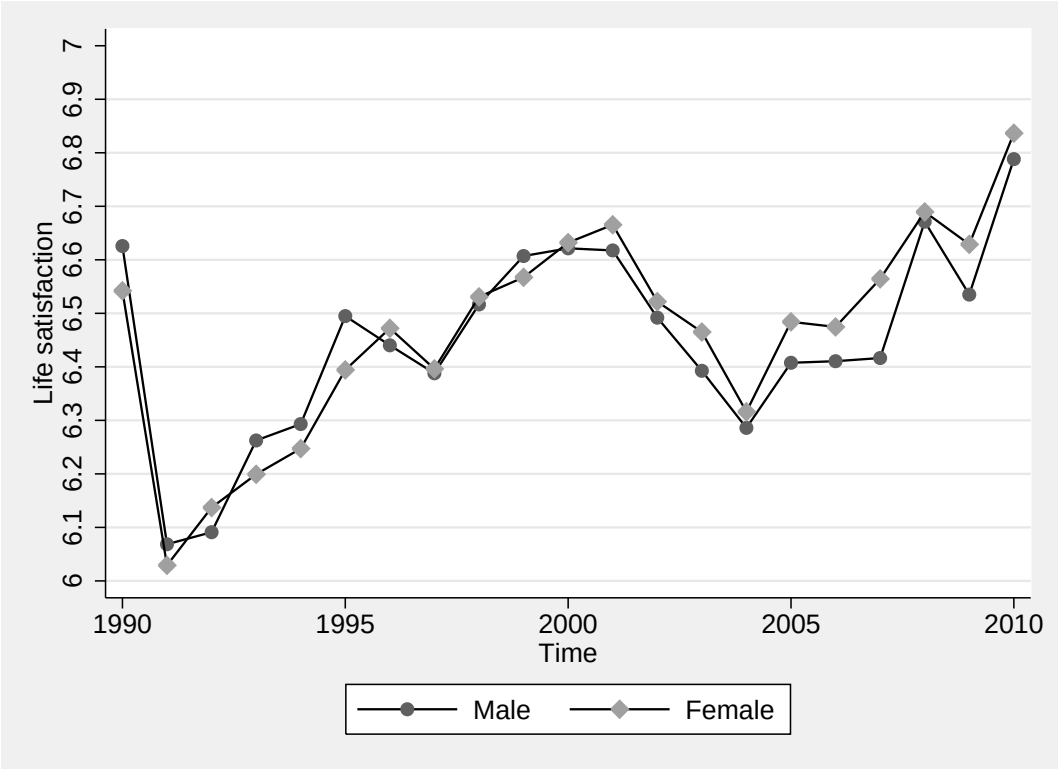
Source: Swiss Household Panel
An individual is regarded as an attritor at age n of the panel if the individual was a respondent at age 1 and if the number of times the individual responded to a subjective well-being question is strictly less than n.

Figure 4: Trends of Life Satisfaction over Time: West Germany



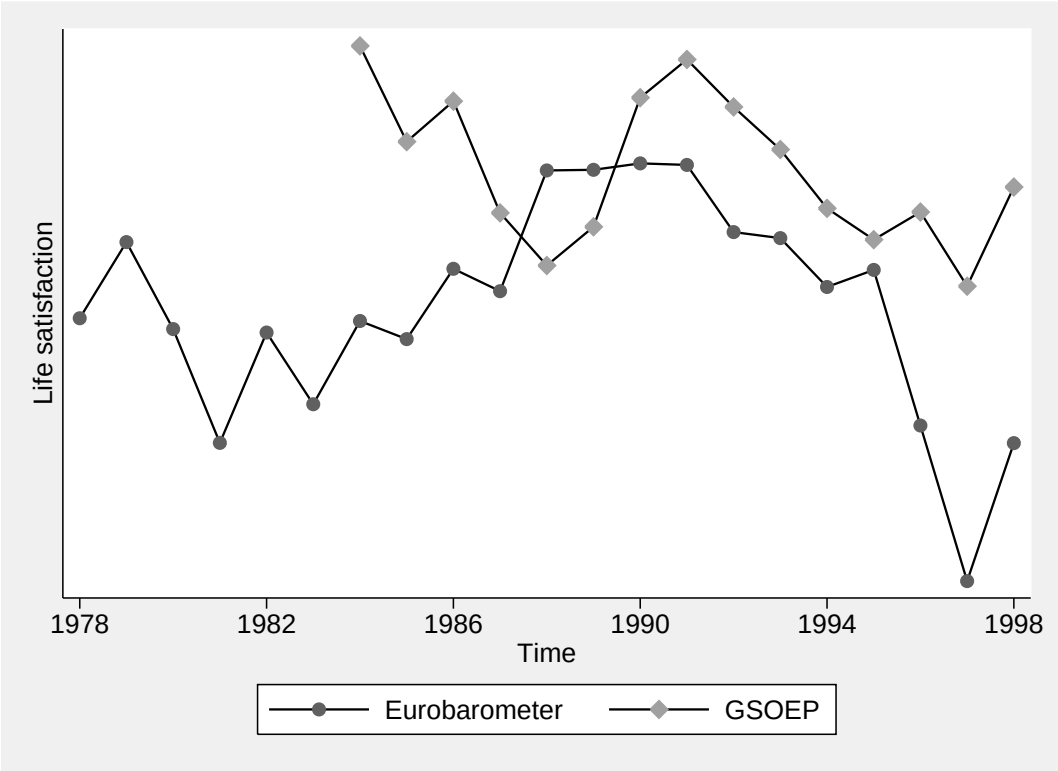
Source: German Socio-Economic Panel.

Figure 5: Trends of Life Satisfaction over Time: East Germany



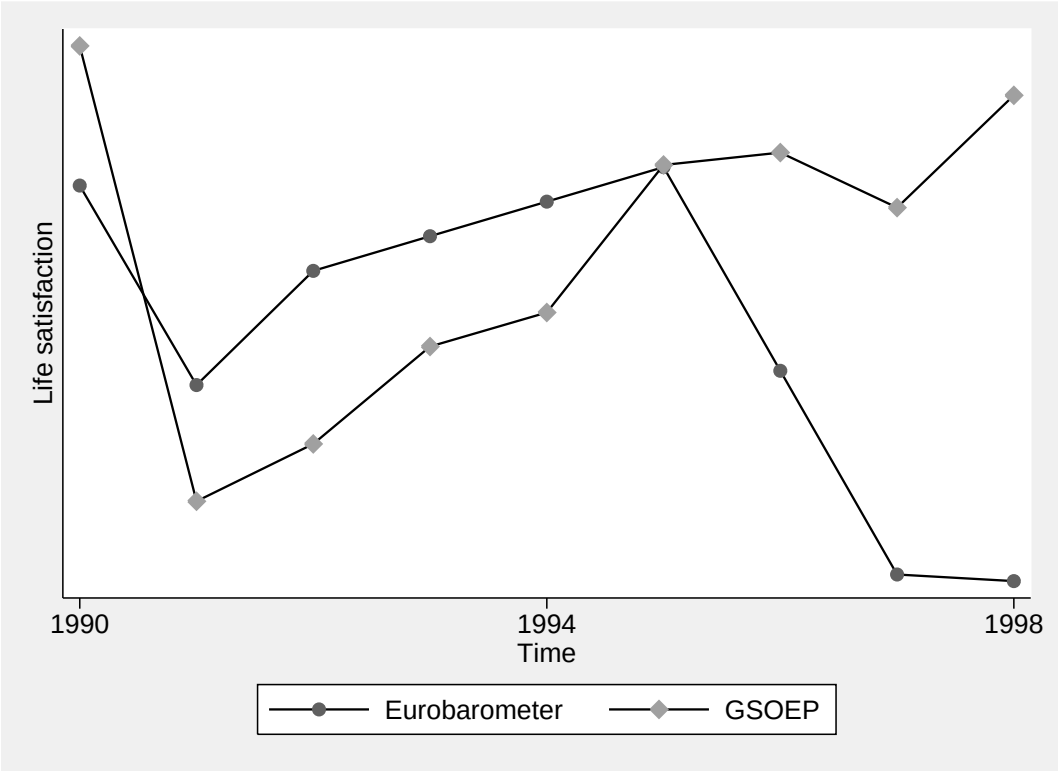
Source: German Socio-Economic Panel.

Figure 6: A Comparison between Trends in Standardized Life Satisfaction in the GSOEP and the Eurobarometer: West Germany



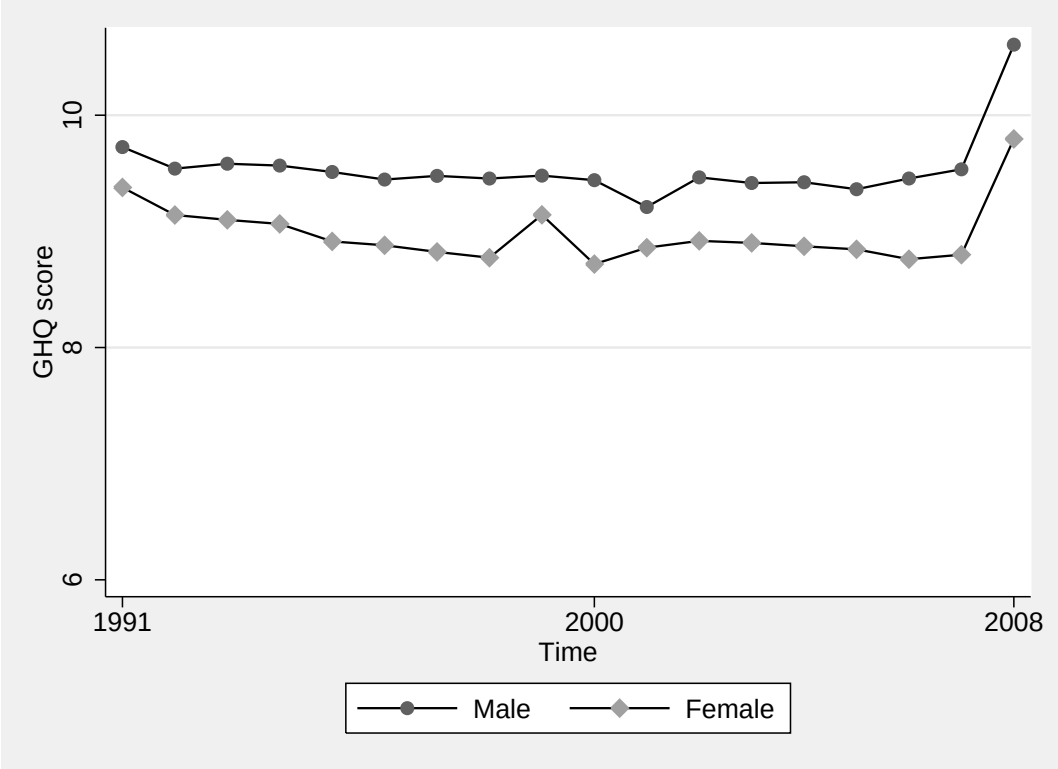
Source: German Socio-Economic Panel and Eurobarometer Survey.

Figure 7: A Comparison between Trends in Standardized Life Satisfaction in the GSOEP and the Eurobarometer: East Germany



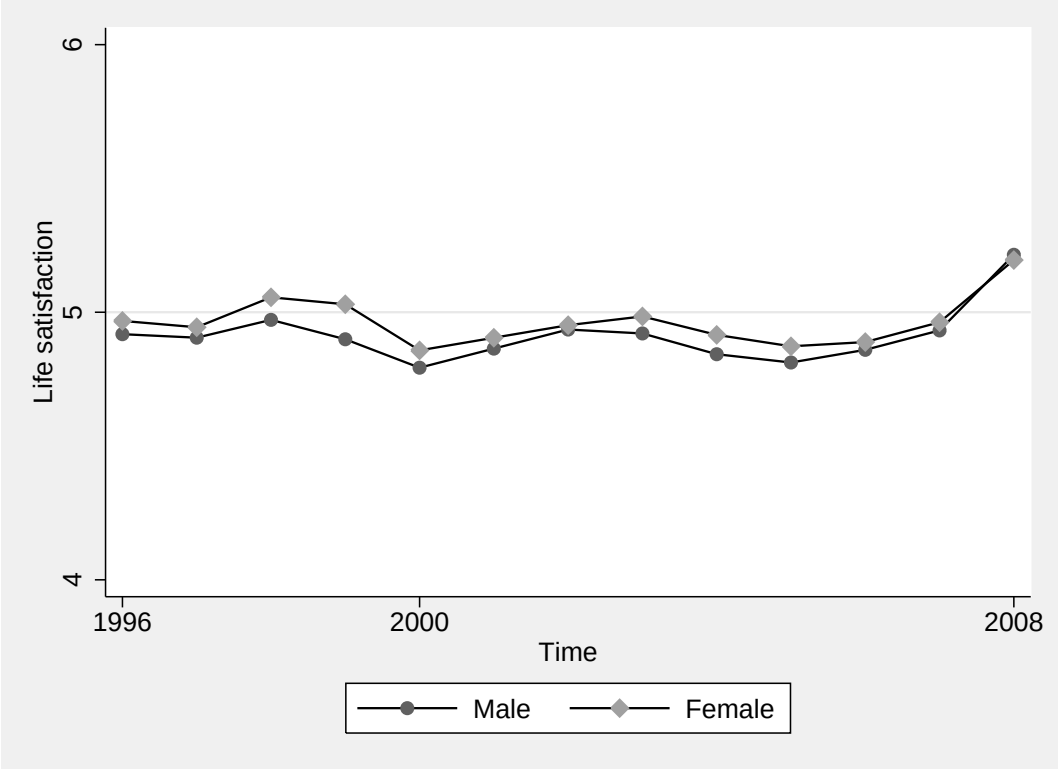
Source: German Socio-Economic Panel and Eurobarometer Survey.

Figure 8: Trends of GHQ Scores over Time: Britain



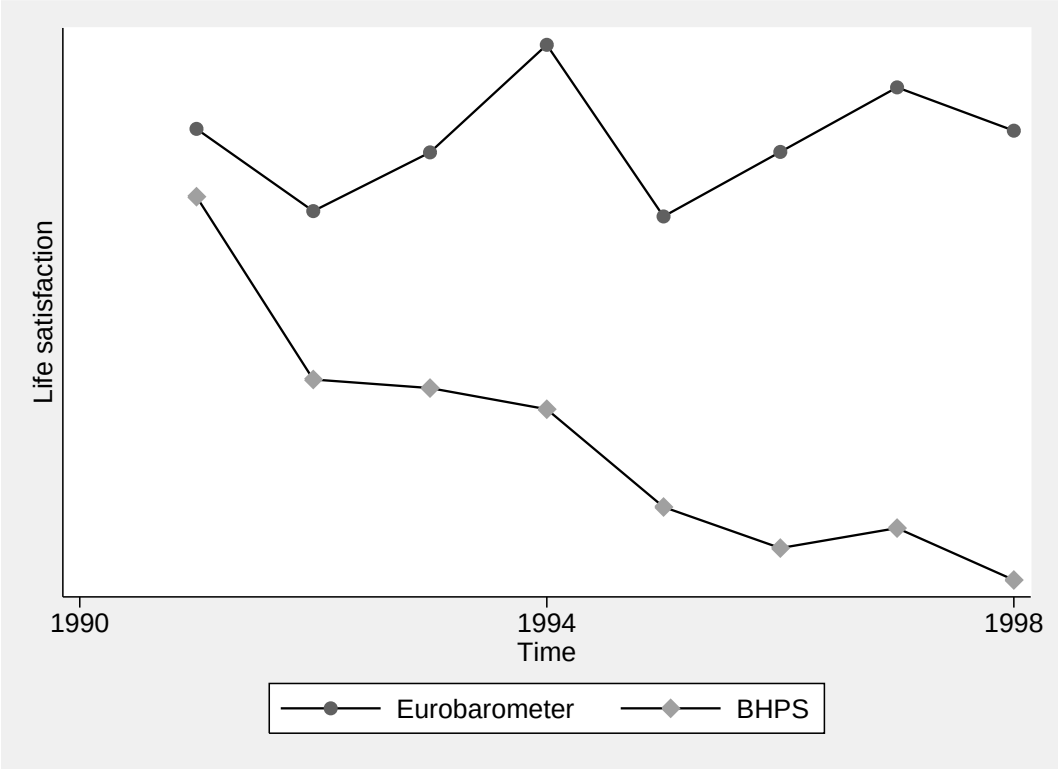
Source: British Household Panel Survey.

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



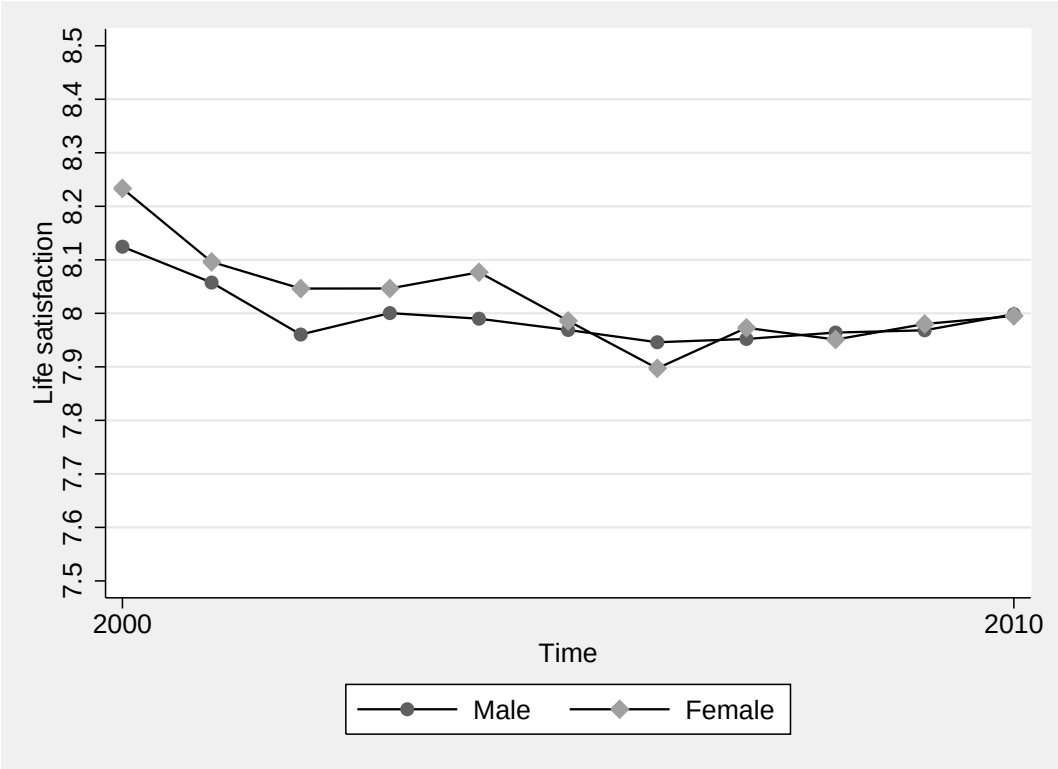
Source: British Household Panel Survey.

Figure 10: 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 11: Trends of Life Satisfaction Scores over Time: Switzerland



Source: Swiss Household Panel.