

Would you like to know what makes people happy? An overview of the data sets on subjective well-being

Nattavudh Powdthavee*

London School of Economics and University of Melbourne

15th January 2015

Abstract

This article provides a guide for young economists wishing to conduct well-being research. It describes the different data sources commonly used in the study of people's subjective well-being, as well as provides a brief discussion on the types of the well-being measures available within each data set.

1. Introduction

Economics is changing. There are now more researchers studying the determinants of people's life satisfaction, happiness and mental well-being than ever before.¹ We are inevitably drawing closer to psychology and medicine.

Having recognised the potentials for empirical research in this area, many household surveys around the world have started to incorporate a series of questions aimed specifically to elicit people's subjective well-being (SWB) into their standard set of questionnaires. This includes, but not limited to, questions that ask survey participants to self-rate their satisfaction with life as a whole, their happiness yesterday, and their usual state of mental well-being. And from studying how people respond to these subjective questions across time and space², researchers have been able to provide insights into some of the most important philosophical issues in economics, including, for example, whether or not people adapt to adverse life events

*Address: CEP, London School of Economics, Houghton Street, London, UK, WC2A 2AE. Email: n.powdthavee@lse.ac.uk. Tel: +44(0)7990 815924.

¹ According to [Kahneman and Krueger \(2006\)](#), a cross-tabulation of *EconLit* suggests that more than 100 published papers included an analysis of happiness and mental well-being data between 2001 and 2005, compared to just four papers in 1991-1995.

² For discussions on the validity of SWB responses, see, for example, [Kahneman and Krueger \(2006\)](#) and [Oswald and Wu \(2010\)](#).

such as disability and unemployment ([Lucas et al., 2004](#); [Clark et al., 2008](#); [Oswald & Powdthavee, 2008](#)); whether economic growth is good for the society in the long-run ([Easterlin, 1974; 1995](#)); and whether we can figure out a better way to value non-market goods such as health and the environment ([Van Praag & Baarsma, 2005](#); [Luechinger, 2009](#); [Powdthavee & Van Den Berg, 2011](#)).

Knowing which data sets to use may seem daunting to many young researchers coming into the field for the first time. To help ease the search process, this article seeks to provide readers with a guide to the typical data sets used in the study of human well-being. However, the list is not supposed to be exhaustive and is meant only as a guideline for PhD and young colleagues who are thinking about starting a project that uses SWB data as either the outcome variable or the input variable.

The structure of the paper is as follows. Section 2 describes the SWB measures and the kind of micro-econometric regression equations typically estimated in the literature. An overview of the SWB data sets can be found in Section 3. Section 4 discusses and concludes.

2. Measuring SWB

2.1. Affective and cognitive well-being

There are two main components of SWB ([Diener et al., 1999](#)). The first is affective well-being (AW), and the second is cognitive well-being (CW).³ By definition, AW represents moods and emotions, which can either be positive or negative and can fluctuate significantly from moment to moment. On the other hand, CW involves a cognitive evaluation process of how satisfied one is with one's life as a whole. Although both measures are highly correlated within the same data set, research in psychology has shown that the two constructs are separable and additive ([Lucas et al., 1996](#)).

AW and CW are generally measured separately and independently. A person's AW can be elicited through a variety of self-reported questionnaires about his or her current (or most recent) moods and emotions, e.g., feelings of happiness, pride, joy,

³ Latest research has also argued that there is also a third component of SWB, which is eudaimonic well being (i.e. feelings of meaningfulness and purpose) ([Huppert and So, 2013](#)).

anger, guilt, anxiety, and depression. One example of an AW question typically elicited in a household survey is

“Have you recently been feeling reasonably happy, all things considered?
1. Much less than usual, 2. Less so than usual, 3. Same as usual, 4. Much more than usual.”

To elicit CW, survey participants are commonly asked the following self-rated life satisfaction question,

“How dissatisfied or satisfied are you with your life overall? 1. Not satisfied at all, ..., 7. Completely satisfied.”

Other than a question about satisfaction with life overall, many household surveys also tend to ask their survey participants to self-rate their levels of satisfaction with different areas of life, including satisfaction with health, job, housing, marriage, and leisure time.

2.2. SWB function

The general idea is that there exists a reported well-being function

$$r = h(u(z, t, p)) + e,$$

where r is some reported number or level of either AW or CW, $u(\dots)$ is to be thought as the person’s true well-being, $h(\cdot)$ is a continuous non-differentiable function relating to reported well-being, z is a set of observable characteristics of the individual, t is time, p is a set of personality traits and other unobserved time-invariant characteristics, and e is an error term that subsumes among other factors the inability of human beings to communicate accurately their well-being level ([Blanchflower & Oswald, 2004](#)). We assume that $h(\cdot)$ rises in step with $u(\dots)$.

Responses to AW and CW questions are normally treated as ordinal and, hence, are usually estimated using ordinal regression models such as ordered logit and ordered probit. However, recent research has shown that it makes virtually no

difference whether one assumes cardinality or ordinality in the SWB data, although it is important for researchers working with SWB data to properly account for the presence of unobserved characteristics, p , from biasing the estimates of interest (Ferrer-i-Carbonell & Frijters, 2004). Generally, a good longitudinal data set (i.e., one with a reasonably large N and T) is required to overcome the problem of unobserved heterogeneity bias in cases where the explanatory variables of interest are not randomly distributed across individuals in the sample.

3. Data

3.1. Cross-national data

Perhaps the most used data set for cross-national comparisons of SWB is the World Values Surveys (WVS). The WVS is a repeated cross-national survey of individuals, which had been conducted at the global level since 1981 (Inglehart, 2000). There were 10 participating countries featured in the first wave of the WVS (1981-1984). This number grew to 52 participating members in its latest wave (Wave 6, 2010-2014), with some members joining and/or leaving in-between waves. Overall, the WVS has interviewed over 80 nations in its history. And although the number of people interviewed in each country varies greatly, the minimum of 1,000 respondents from each member is required.⁴

Data on life satisfaction are present since Wave 1 for the majority of participating nations in the WVS. The phrasing of the question in English, which has been translated to all participating countries in the WVS, is: “All things considered, how satisfied are you with your life as a whole these days?”, with possible answers ranging from 1 (“Dissatisfied”) to 10 (“Satisfied”). WVS also contains a question about the respondent’s happiness, which is more of a measure for AW than CW, and the phrasing of the happiness question is: “Taken all things together, would you say you are: very happy, quite happy, not very happy, or not at all happy?”

Another large-scale collection of cross-national data of well-being is the Gallup World Poll (GWP). Began in 2006 by the Gallup Organization, the GWP is a repeated cross-national survey of individuals from 132 countries; with the exception of Angola, Cuba, and Myanmar, where the samples are urban, the GWP samples are nationally representative of people aged 15 and older. The questionnaires in the GWP,

⁴ For details on the WVS and how to access them, see <http://www.worldvaluessurvey.org/wvs.jsp>.

like WVS, covered many aspects of subjective well-being (e.g. life satisfaction, happiness, moods, etc.), as well as several aspects of health, family, and economic status.⁵

There are also other similar repeated cross-national surveys of well-being but that are more regional specific and therefore are relatively smaller in terms of scale than the WVS and GWP. This includes the Eurobarometer, Latinobarometer, and the Afrobarometer.⁶

While two of the many advantages of using the WVS, the GWP, or the barometer surveys to study well-being lie in its incredibly large sample size and the diversity of its sampled populations, one natural concern for well-being researchers is whether the vast differences in some of the participating members' cultures could potentially bias the estimates of both AW and CW. For example, studies have shown that people from Asia systematically report lower life satisfaction than people in America. Much of these differences, argued by many researchers, can be explained by the differences in cultures (which are largely unobserved to researchers) than by the differences in socio-economic status (e.g. [Diener et al., 2003, 2009](#); [Oishi, 2006](#)), thus implying that any cross-national results have to be interpreted with care.

3.2. Cross-sectional data

One of the earliest cross-sectional data sets on SWB is the General Social Survey (GSS), which is a national-representative, repeated cross-sectional data set for the United States. Since 1972, the GSS have been asking their respondents the following AW question: "Taken together, how would you say things are these days – would you say that you are happy, pretty happy, or not too happy?" There is, however, no question on life satisfaction in the GSS.⁷ The GSS has been used in many important studies, including, for example, a paper by Richard Easterlin on the unified theory on income and SWB (Easterlin, 2001) and a paper by Andrew Oswald, Rafael Di Tella, and Robert MacCulloch on the relationship between macroeconomic indicators and SWB (2001).

⁵ For details on the GWP, see <http://www.gallupworldpoll.com/content/24046/About.aspx>.

⁶ For details on the Eurobarometer, see http://ec.europa.eu/public_opinion/index_en.htm. For the details on Latinobarometer, see <http://www.latinobarometro.org/lat.jsp>. And for the details on Afrobarometer, see <http://www.afrobarometer.org>.

⁷ For details on the GSS, see <http://www3.norc.umd.edu/gss+website>.

Another important source of information on SWB in the United States is the Behavioural Risk Factor Surveillance System (BRFSS). Established in 1984, the BRFSS is a state-based system of health surveys that collects information on health risk behaviours, preventive health practices, and health care access primarily related to chronic disease and injury. Currently data are collected monthly in all 50 US states, and more than 350,000 adults are interviewed each year, making the BRFSS the largest telephone health survey in the world. The BRFSS started collecting data on life satisfaction for the first time in 2005, which means that there is currently little work on life satisfaction using this data set.⁸

An almost equivalent survey to the BRFSS is the Health Surveys for England (HSE).⁹ The HSE is an annual, nationally representative survey and is designed to monitor the nation's health. Information is collected through a combination of face-to-face interviews, a self-completed questionnaire, and a series of medical examination conducted by the nurse. There are both AW and CW questions in the HSE, with the AW being elicited through the General Health Questionnaires (GHQ-12) and the CW being elicited through the life satisfaction questionnaire. Both BRFSS and HSE are perhaps two of the best cross-sectional data sets for researchers wishing to study the link between SWB and different measures of physical health, including biomarkers such as blood pressure and cortisol level.

3.3. Longitudinal surveys

There are perhaps three main nationally representative longitudinal data sets used in the studies of SWB. They are i) the British Household Panel Survey (BHPS), ii) the German Socio-Economic Panel (SOEP), and iii) the Household, Income and Labour Dynamics in Australia (HILDA).

The BHPS is a multi-purpose study and a nationally representative sample of British households, containing over 15,000 adult individuals in the UK (Taylor et al., 2002).¹⁰ The survey has been conducted between September and Christmas of each year from 1991 to 2008. From wave 1 onward, individuals are asked a battery of self-completed questions about their usual emotional states. These AW questions, more

⁸ One recent exception is the work by Oswald and Wu (2010). For details on the BRFSS, see <http://www.cdc.gov/brfss>.

⁹ For details on the HSE, see <http://www.hscic.gov.uk/article/3741/Health-Survey-for-England-Health-social-care-and-lifestyles>.

¹⁰ For details on the BHPS, see <https://www.iser.essex.ac.uk/bhps>.

commonly known as the General Health Questionnaire (or GHQ-12), ask respondents to indicate on a 4-point scale from 1 (no more than usual) to 4 (much more than usual) how often over the past few weeks they had lost sleep over worry, felt constantly under strain, felt they could not overcome difficulties, been feeling unhappy and depressed, been losing confidence, and been feeling like a worthless person. The same individuals are also asked to indicate on a 4-point scale from 1 (better than usual) to 4 (much less than usual) on how often over the past few weeks that they had felt that they were playing a useful part in things, felt capable of making decisions, been able to enjoy day-to-day activities, been able to concentrate, been able to face up to problems, and been feeling reasonably happy. Researchers typically use the Caseness score of GHQ, in which the number of times the person places himself or herself in the fairly stressed or highly stressed category are added up to form the total score, in their analysis. Many medical scholars have considered the scale to be a good proxy for mental stress and strain (e.g. Guthrie et al., 1998).

Individuals in the BHPS are also asked to self-rate their level of satisfaction with life as a whole on a 7-point scale from wave 6 onward.¹¹ Satisfaction with other domains of life are elicited in the BHPS, including satisfaction with income, health, housing, job, partner, social life, and leisure time.

The SOEP is a nationally representative, longitudinal survey that has closely followed approximately 13,500 West German individuals each year since 1984.¹² The survey then expanded to include residents of the former East Germany following the reunification of Germany in 1990. Since 1984, individuals in the SOEP have been asked to rate their level of satisfaction with their life as a whole on a 10-point scale. The SOEP also contains other CW questions, including satisfaction with health and job. However, unlike the BHPS, it does not contain a standard set of AW questions that would allow researchers to annually track respondents' usual mental states.

The HILDA is a longitudinal survey that has been tracking members of a nationally representative sample of Australian households since 2001, containing an initial sample of nearly 20,000 individuals (Watson and Wooden, 2012).¹³ The members of these participating households form the basis of the panel pursued in subsequent annual survey waves. Similar to the BHPS and the SOEP, interviews are

¹¹ Except for wave 10 where there is a gap in the implementation of the life satisfaction question.

¹² For details on the SOEP, see <http://www.diw.de/en/soep>.

¹³ For details on the HILDA, see <http://www.melbourneinstitute.com/hilda/>.

conducted with all adults (defined as persons aged 15 years or older) who are members of the original sample, as well as any other adults who, in later waves, are residing with an original sample member. The main CW question in the HILDA is the 10-point scale life satisfaction question, which has been asked in every wave since the survey's inauguration in 2001. The HILDA also asks its respondents about their satisfaction with different domains of life, including satisfaction with employment opportunity, health, income, job, partner, housing, leisure, local community, children, and feeling of safety, thus making both the BHPS and the HILDA two of the most comprehensive longitudinal surveys on CW in the world.

From wave 1 onward, the HILDA also asks respondents a battery of mental health (SF-36) questions. These AW questions ask respondents to indicate on a 6-point scale how, during the past four weeks, they have been feeling nervous, feeling so down in the dumps that nothing could cheer them up, feeling calm and peaceful, feeling down, and feeling happy. The responses are then recoded and added up to make a continuous single-item variable of people's usual mental health state. At present, both the BHPS and the HILDA are among two of the best panel surveys of people's usual mental states measured annually in existence.

3.4. Cohort surveys

More recent research into the determinants of individual's SWB has started looking at the long-term relationship between childhood characteristics and adult life CW and AW (Frijters et al., 2014; Layard et al., 2014). This requires a special type of data that tracks the same individuals over a long period of time throughout their life cycle. Two examples of such data sets are the National Child Development Study (NCDS) and the British Cohort Study 1970 (BCS).

The NCDS is a longitudinal study that tracks the lives of all those living in Great Britain who were born in one particular week in March 1958.¹⁴ This covers a cohort of around 17,400 children, with follow-on data having been collected in 1965 (at age 7), 1969 (age 11), 1974 (age 16), 1981 (age 23), 1991 (age 33), 1999-2000 (age 41-42), 2004 (age 46) and 2008-2009 (age 50-51). Since the 1981 sweep of the data, the cohort member has been the main respondent. At ages 33, 42, 46, and 50,

¹⁴ For details on both the NCDS and the BCS, see <http://www.cls.ioe.ac.uk>.

participants in the NCDS have been asked to rate on a 10-point scale their level of satisfaction with life as a whole.

Similar to the NCDS, the BCS is a longitudinal study that follows approximately 17,000 children living in Great Britain who were born in a single week in April 1970. Currently the data are available for nine major follow-up surveys: 1975 (at age 5), 1980 (age 10), 1986 (age 16), 1991 (age 21), 1996 (age 26), 2000 (age 30), 2004 (age 34), 2008 (age 38) and 2012 (age 42). Since the 1991 sweep of the data, the cohort member has been the main respondent. Data on life satisfaction have been collected in the BCS at ages 30, 34, and 42.

Both NCDS and BCS also collect a battery of indexes that had been designed to elicit the person's mental well-being throughout the life-course. One example of such index is the Rutter Malaise Inventory scale (Rutter et al., 1970), which is an index derived from a number of "yes" responses to: having backaches, feeling tired, feeling miserable and depressed, having headaches, worrying, having difficulty in falling asleep or staying asleep, waking unnecessarily early in the morning, worrying about health, getting annoyed by people, having twitches, becoming scared for no reason, being scared to be alone, being easily upset, being frightened of going alone, being jittery, suffering from indigestion, suffering from upset stomach, having poor appetite, being worn out by little things, experiencing racing heart, having bad pains in your eyes, being troubled by rheumatism, and having had a nervous breakdown.

4. Conclusion

All the data sources I have described should provide young researchers with plenty of opportunities to test different untested ideas and hypotheses about what makes people happy and satisfied with their life. Given that there are often direct policy implications to well-being research, coupled with the fact that many of these data sets are either made publically accessible through the Internet for free via registration or relatively affordable to purchase compared to many other data sources, there is perhaps no better time to conduct research on the economics of happiness.

References

- Blanchflower, D.G., and Oswald, A.J. 2004. Well-being over time in Britain and the USA. *Journal of Public Economics*, 88, 1359-1386.
- Clark, A.E., Diener, E., Georgellis, Y., and Lucas, R.E. 2008. Lags and leads in life satisfaction: a test of the baseline hypothesis. *Economic Journal*, 118(529), F222-F243.
- Diener, E., Suh, E.M., Lucas, R.E., and Smith, H.L. 1999. Subjective well-being: three decades of progress. *Psychological Bulletin*, 125(2), 276-302.
- Diener, E., Oishi, S., and Lucas, R. E. 2003. Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual review of psychology*, 54(1), 403-425.
- Diener, E., Oishi, S., and Lucas, R. E. 2009. 17 Subjective Well-Being: The Science of Happiness and Life Satisfaction. *Oxford handbook of positive psychology*, 187.
- Di Tella, R., MacCulloch, R. J., and Oswald, A. J. (2001). Preferences over inflation and unemployment: Evidence from surveys of happiness. *American Economic Review*, 335-341.
- Easterlin, R. A. 1974. Does economic growth improve the human lot? Some empirical evidence. *Nations and Households in Economic Growth*, 89, 89-125.
- Easterlin, R. A. 1995. Will raising the incomes of all increase the happiness of all?. *Journal of Economic Behavior & Organization*, 27(1), 35-47.
- Easterlin, R. A. 2001. Income and happiness: Towards a unified theory. *The economic journal*, 111(473), 465-484.
- Ferrer-i-Carbonell, A., and Frijters, P. 2004. How important is methodology for the determinants of happiness? *Economic Journal*, 114(497), 641-659.
- Frijters, P., Johnston, D.W., and Shields, M.A. 2014. Does childhood predict adult life satisfaction? Evidence from British Cohort Surveys, *Economic Journal*, 124(580), F688-F719.
- Guthrie, E., Black, D., Bagalkote, H., Shaw, C., Campbell, M. and Creed, F. (1998). Psychological stress and burnout in medical students: a five-year prospective longitudinal study. *Journal of the Royal Society of Medicine*, 91, 237-243.
- Huppert, F.A., and So, T.T.C. 2013. Flourishing across Europe: application of a new conceptual framework for defining well-being. *Social Indicators Research*, 110(3), 837-861.

- Inglehart, R., and colleagues. 2000. *World Values Surveys and European Values Surveys, 1981-1984, 1990-1993, and 1995-1997*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.
- Kahneman, D., and Krueger, A.B. 2006. Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1), 3-24.
- Layard, R., Clark, A.E., Cornaglia, F., Powdthavee, N., and Vernoit, J. 2014. What predicts successful life? A life-course model of well-being. *Economic Journal*, 124(580), F720-F738.
- Lucas, R.E., Clark, A.E., Georgellis, Y., and Diener, E. 2004. Unemployment alters the set point for life satisfaction. *Psychological Science*, 15(1), 8-13.
- Lucas, R.E., Diener, E., and Suh, E. 1996. Discriminant validity of well-being measures. *Journal of Personality and Social Psychology*, 71, 616-628.
- Rehdanz, K., and Maddison, D. 2008. Local environmental quality and life-satisfaction in Germany. *Ecological Economics*, 64(4), 787-797.
- Oishi, S. 2006. The concept of life satisfaction across cultures: An IRT analysis. *Journal of Research in Personality*, 40(4), 411-423.
- Oswald, A.J., and Powdthavee, N. 2008. Does happiness adapt? A longitudinal study of disability with implications for economists and judges. *Journal of Public Economics*, 92, 1061-1077.
- Oswald, A.J., and Wu, S. 2010. Objective confirmation of subjective measures of human well-being: evidence from the USA. *Science*, 327, 576-579.
- Powdthavee, N., and Van den Berg, B. 2011. Putting different price tags on the same health condition: Re-evaluating the well-being valuation approach. *Journal of Health Economics*, 30(5), 1032-1043.
- Rutter, M., Tizard, J., and Whitmore, K. 1970. *Education, Health and Behaviour*. Longman.
- Taylor, M.F., Brice, J., Buck, N., and Prentice-Lane, E. 2002. *British Household Panel Survey User Manual*. Colchester: University of Essex.
- Watson, N. and Wooden, M. 2012. The HILDA Survey: a case study in the design and development of a successful household panel study. *Longitudinal and Life Course Studies*, 3(3), 369-381.
- Van Praag, B., and Baarsma, B. E. 2005. Using Happiness Surveys to Value Intangibles: The Case of Airport Noise. *Economic Journal*, 115(500), 224-246.