



STATISTICS IN TRANSITION

new series

An International Journal of the Polish Statistical Association

CONTENTS

The Measurement of Subjective Well-Being in Survey Research

From the Editors: G. Kalton, C. Mackie and W. Okrasa.....	329
Submission information for authors	333
Mackie C., Smith C., Conceptualizing Subjective Well-Being and its Many Dimensions – Implications for Data Collection in Official Statistics and for Policy Relevance	335
Tinkler L., The Office for National Statistics Experience of Collecting and Measuring Subjective Well-Being	373
Allin P., Official Statistics on Personal Well-Being: Some Reflections on the Development and Use of Subjective Well-Being Measures in the UK	397
Fattore M., Maggino F., Arcagni A., Exploiting Ordinal Data for Subjective Well-Being Evaluation	409
Smith D. M., Using the Day Reconstruction Method to Quantify Time Spent Suffering among Older Adults with Chronic Pain.....	429
Xing Z., Qu X., An Initial Research on Output Well-Being Index Applied to Residents in Wealthy Counties from China	441
Koźmiński A. K., Noga A., Piotrowska K., Zagórski K., Operationalization and Estimation of Balanced Development Index for Poland 1999-2016	461
About the Authors	479

EDITOR IN CHIEF

Prof. W. Okrasa, *University of Cardinal Stefan Wyszyński, Warsaw, and Central Statistical Office of Poland*
w.okrasa@stat.gov.pl; Phone number 00 48 22 — 608 30 66

ASSOCIATE EDITORS

Belkindas M.,	<i>Open Data Watch, Washington D.C., USA</i>	Osaulenko O.,	<i>National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine</i>
Bochniarz Z.,	<i>University of Minnesota, USA</i>	Ostasiewicz W.,	<i>Wrocław University of Economics, Poland</i>
Ferligoj A.,	<i>University of Ljubljana, Ljubljana, Slovenia</i>	Pacakova V.,	<i>University of Economics, Bratislava, Slovak Republic</i>
Ivanov Y.,	<i>Statistical Committee of the Commonwealth of Independent States, Moscow, Russia</i>	Panek T.,	<i>Warsaw School of Economics, Poland</i>
Jajuga K.,	<i>Wrocław University of Economics, Wrocław, Poland</i>	Pukli P.,	<i>Central Statistical Office, Budapest, Hungary</i>
Kotzeva M.,	<i>Eurostat, Luxembourg</i>	Szreder M.,	<i>University of Gdańsk, Poland</i>
Kozak M.,	<i>University of Information Technology and Management in Rzeszów, Poland</i>	de Ree S. J. M.,	<i>Central Bureau of Statistics, Voorburg, Netherlands</i>
Krapavickaite D.,	<i>Institute of Mathematics and Informatics, Vilnius, Lithuania</i>	Traat I.,	<i>University of Tartu, Estonia</i>
Lapins J.,	<i>Statistics Department, Bank of Latvia, Riga, Latvia</i>	Verma V.,	<i>Siena University, Siena, Italy</i>
Lehtonen R.,	<i>University of Helsinki, Finland</i>	Voineagu V.,	<i>National Commission for Statistics, Bucharest, Romania</i>
Lemmi A.,	<i>Siena University, Siena, Italy</i>	Wesołowski J.,	<i>Central Statistical Office of Poland, and Warsaw University of Technology, Warsaw, Poland</i>
Młodak A.,	<i>Statistical Office Poznań, Poland</i>	Wunsch G.,	<i>Université Catholique de Louvain, Louvain-la-Neuve, Belgium</i>
O'Muircheartaigh C. A.,	<i>University of Chicago, Chicago, USA</i>		

FOUNDER/FORMER EDITOR

Prof. J. Kordos, *Central Statistical Office, Poland*

EDITORIAL BOARD

Prof. Witkowski, Janusz (Co-Chairman), *Central Statistical Office, Poland*
Prof. Domański, Czesław (Co-Chairman), *University of Łódź, Poland*
Sir Anthony B. Atkinson, *University of Oxford, United Kingdom*
Prof. Ghosh, Malay, *University of Florida, USA*
Prof. Kalton, Graham, *WESTAT, and University of Maryland, USA*
Prof. Krzyśko, Mirosław, *Adam Mickiewicz University in Poznań, Poland*
Prof. Wywił, Janusz L., *University of Economics in Katowice, Poland*

Editorial Office

Marek Cierpiał-Wolan, Ph.D.: Scientific Secretary
m.wolan@stat.gov.pl
Beata Witek: Secretary
b.witek@stat.gov.pl. Phone number 00 48 22 — 608 33 66
Rajmund Litkowicz: Technical Assistant

Address for correspondence

GUS, al. Niepodległości 208, 00-925 Warsaw, POLAND, Tel./fax:00 48 22 — 825 03 95

ISSN 1234-7655

The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 329–332

THE MEASUREMENT OF SUBJECTIVE WELL-BEING IN SURVEY RESEARCH

About a year ago, in recognition of the importance of measuring subjective well-being in surveys and the challenges it presents, *Statistics in Transition new series (SiT)* announced a plan to publish an issue containing a collection of papers on statistically relevant aspects of research on subjective well-being. This thematic issue of SiT is the result of that plan. It contains a set of original papers especially produced for this occasion that present some recent advances in survey-based research on subjective well-being. The main objective of this issue is to provide readers with an understanding of the conceptual and methodological, as well as practical and institutional, matters involved in collecting and analysing data in nation-wide studies of subjective well-being, in multi-national contexts. The papers collected in this issue examine the key problems in measuring subjective well-being in the survey context, with an emphasis on strategies for developing effective measures of subjective well-being, including for cross-cultural comparisons.

Some of the submissions in this volume arrived in direct response to the call for paper from the authors who were generous enough to share their results with the research community at large. Others originated from personal requests from the editors of this issue. As the Editor coordinating of this process on behalf of the Editorial Office, I feel indebted to Graham Kalton and Chris Mackie – who kindly accepted a request to serve as Guest Editors of this issue – for taking up the task of arranging for several papers from already overburdened leading experts in the field. Let me take this opportunity to express my gratitude to Chris and Graham for the invaluable contribution they made in the development of this issue.

*

This topical issue presents a collection of seven papers. Each of them makes some novel contributions, even those which are aimed at codifying a seemingly established knowledge base or an emerging research paradigm in a broadly defined topic area. In order to facilitate an overview of the papers, they are briefly summarized below.

The collection begins with a paper on *Conceptualizing Subjective Well-Being and its Many Dimensions – Implications for Data Collection in Official Statistics and for Policy Relevance*, by **Christopher Mackie** and **Conal Smith**. Noting that subjective well-being encompasses several distinct but interacting aspects of people's feelings, attitudes, and experiences, the authors discuss recent

developments in measuring these dimensions on the basis of self-reports of subjective well-being collected in survey instruments. After considering the multidimensional nature of subjective well-being, the role of national statistics offices in measuring subjective well-being and deriving official statistics is discussed. They conclude by indicating consequences that the selection of a given type of construct (a set of characteristics of well-being) may have for their use for policy purposes. The overarching conclusion of this paper is that, despite well documented limitations (methodological and conceptual), a growing consensus is emerging on how to collect and interpret data on subjective well-being that have the potential to positively contribute to better informed policies in a number of areas.

In the paper *The Office for National Statistics Experience of Collecting and Measuring Subjective Well-Being*, **Lucy Tinkler** presents a detailed description of the UK Office for National Statistics' experience with measuring subjective well-being. She recounts the commencement of the project (in 2011) within the ONS Measuring National Well-being programme that was established to monitor the quality of life and progress for populations in the UK. The paper discusses the development of the ONS subjective well-being measures, and provides information on data collection methods and data presentation considerations, along with an overview of findings and recent developments. In conclusion, the author points to the way in which user engagement has been key to the development of the ONS subjective well-being statistics, and suggests ways to engage a wide range of stakeholders, including international organisations and policy makers (as well as citizen users), in order to ensure that ONS subjective well-being analysis is relevant.

Paul Allin's paper *Official Statistics on Personal Well-Being: Some Reflections on the Development and Use of Subjective Well-Being Measures in the UK* also draws on experiences of the UK Office for National Statistics (ONS) programme to measure national well-being. The author focuses on the high-profile element of the programme in which subjective well-being measures have been collected and published (since April 2011). Although Allin touches briefly on measurement and analysis, most of the discussion covers issues related to user requirements, the international context, and political, policy, public and business uses of well-being data. He describes the main drivers of the ONS work and how these have given rise to interest both in national well-being (the "beyond GDP" agenda) and in the use of different measures (self-reported, personal well-being) in public policy. The paper concludes by noting that well-being can give a very different picture from the one constructed only with economic and largely market-based considerations in mind.

Marco Fattore, Filomena Maggino and Alberto Arcagni, in the paper *Exploiting Ordinal Data for Subjective Well-Being Evaluation*, present an original approach to measuring subjective well-being that is envisaged as an

alternative to composite indicators or counting procedures which dominate the literature to date. Using the theory of partially ordered sets, the authors demonstrate how the evaluation of multidimensional ordinal well-being can be addressed in an effective and consistent way. The proposed approach avoids the risk of inconsistencies and inefficiency in the treatment of ordinal data that may affect the quality of information provided to researchers and policy makers. First, they show that the proper evaluation space of well-being is the partially ordered set of achievement profiles, the structures of which depend on the importance assigned to various attributes. Next, they describe how evaluations can be performed by extracting information out of the evaluation space (respecting the ordinal nature of data) and producing synthetic indicators without attribute aggregation. The paper concludes with an application of their approach to subjective well-being in Italy.

The paper by **Dylan M. Smith**, *Using the Day Reconstruction Method to Quantify Time Spent Suffering among Older Adults with Chronic Pain*, explores the methodological potential of the Day Reconstruction Method (DRM) showing its suitability in the context of measuring selected (negative) aspects of experienced well-being. The DRM aims to measure time use in a manner that is more valid than traditional written summary measures, but less burdensome than real-time electronic diary methods. The lower respondent burden and administration costs of the method may create opportunities to exploit it in national survey contexts. In contrast to past studies – employing the DRM for characterizing subjective mood states during different types of activities – Smith uses the DRM to measure suffering associated with negative symptoms such as pain and fatigue in 122 older adults, most of whom suffer from chronic pain. The results indicated that the method could be administered effectively with this population, with over 98% of the sample providing interpretable responses. Time spent suffering was associated with lower income, negative mood, and lower life satisfaction and quality of life. In summarizing his findings, Smith emphasizes the adaptability of the DRM for surveying well-being, especially for capturing suffering, in addition to emotional well-being.

Zhanjun Xing and **Xiaxia Qu** present the results of *An Initial Research on Output Well-Being Index Applied to Residents in Wealthy Counties from China*. Using a sample of 855 residents from three wealthy counties in Shandong province, data on several characteristics were gathered in order to construct a well-being index. After reviewing the index's internal consistency, reliability and construct validity, it was adopted to serve as an instrument to evaluate the quality of life of Chinese citizens. When the index was applied to measuring the quality of life in the three counties, it was found that the quality of life levels of rural residents were generally higher than those of urban residents. However, the level of some indicators was not stable and the levels of subjective and objective indicators tracked somewhat consistently with one another. The well-being characteristics of the rural residents were shown to be closely related to the local

economic performance and social development of the areas in which they lived. The authors conclude that it could prove informative to use the output of well-being indexes to evaluate the degree to which the citizens' needs have been met and to which overall development has been achieved. Such measures could also be used as evaluative instruments for policy makers and planners at the local level.

The last paper, by **Andrzej K. Koźmiński, Adam Noga, Katarzyna Piotrowska, and Krzysztof Zagórski** on *Operationalization and Estimation of Balanced Development Index for Poland 1999-2016* addresses macro-level aspects of well-being and their relation to the economy, in a historical perspective. Developing from a perspective that Gross Domestic Product oversimplifies economic development and socio-economic progress, the authors propose a new conceptual and methodological approach that motivates estimation of a Balanced Development Index (BDI) assumed to capture both economic and social development in Poland. Calculations of this composite index, as well as of its four components (middle-level indexes), were made for the period from 1999 to 2013. The estimated trends in BDI suggest that socio-economic changes were less positive than those suggested by the GDP growth only, and that public expectations fluctuate to a much greater extent than does the actual situation as measured jointly by objective and subjective indicators.

Graham Kalton, Guest Editor

Christopher Mackie, Guest Editor

Włodzimierz Okrasa, Editor

STATISTICS IN TRANSITION new series, Autumn 2015
Vol. 16, No. 3, pp. 333

SUBMISSION INFORMATION FOR AUTHORS

Statistics in Transition new series (SiT) is an international journal published jointly by the Polish Statistical Association (PTS) and the Central Statistical Office of Poland, on a quarterly basis (during 1993–2006 it was issued twice and since 2006 three times a year). Also, it has extended its scope of interest beyond its originally primary focus on statistical issues pertinent to transition from centrally planned to a market-oriented economy through embracing questions related to systemic transformations of and within the national statistical systems, world-wide.

The *SiT-ns* seeks contributors that address the full range of problems involved in data production, data dissemination and utilization, providing international community of statisticians and users – including researchers, teachers, policy makers and the general public – with a platform for exchange of ideas and for sharing best practices in all areas of the development of statistics.

Accordingly, articles dealing with any topics of statistics and its advancement – as either a scientific domain (new research and data analysis methods) or as a domain of informational infrastructure of the economy, society and the state – are appropriate for *Statistics in Transition new series*.

Demonstration of the role played by statistical research and data in economic growth and social progress (both locally and globally), including better-informed decisions and greater participation of citizens, are of particular interest.

Each paper submitted by prospective authors are peer reviewed by internationally recognized experts, who are guided in their decisions about the publication by criteria of originality and overall quality, including its content and form, and of potential interest to readers (esp. professionals).

Manuscript should be submitted electronically to the Editor:
sit@stat.gov.pl.

It is assumed, that the submitted manuscript has not been published previously and that it is not under review elsewhere. It should include an abstract (of not more than 1600 characters, including spaces). Inquiries concerning the submitted manuscript, its current status etc., should be directed to the Editor by email, address above, or w.okrasa@stat.gov.pl.

For other aspects of editorial policies and procedures see the *SiT* Guidelines on its Web site: <http://stat.gov.pl/en/sit-en/guidelines-for-authors/>

The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 335–372

CONCEPTUALIZING SUBJECTIVE WELL-BEING AND ITS MANY DIMENSIONS – IMPLICATIONS FOR DATA COLLECTION IN OFFICIAL STATISTICS AND FOR POLICY RELEVANCE

Christopher Mackie¹, Conal Smith²

ABSTRACT

Subjective well-being encompasses several distinct but interacting aspects of people's feelings, attitudes, and experiences. This paper assesses the state of the art for measuring these dimensions of people's lives, which typically involves analyzing self-reports of subjective well-being collected in survey instruments; however, other potentially complementary, technology-driven tools are emerging as well. We first answer the question, "what is subjective well-being?" and unpack its multidimensionality. The role of national statistics offices in measuring subjective well-being and deriving official statistics is considered next. We conclude by discussing how different characteristics of well-being constructs shape their applicability to policy. The overarching conclusion is that—while methodological limitations are present and a number of fundamental research challenges remain—understanding of how to collect and interpret data on subjective well-being has made enormous strides in the last two decades, and policies for a wide range of domains are beginning to be usefully informed.

Key words: subjective well-being, national statistics, policy.

1. Introduction

Notions of subjective well-being (SWB) or happiness have a long tradition as central elements of the good life. However, until recently, these concepts were often deemed impossible to measure, and certainly beyond the scope of official statistics. In the past two decades, however, an increasing body of evidence has shown that SWB can be measured in surveys, that such measures are valid and reliable, and that they can inform policy making. This evidence has been reflected in the exponential growth of research in this field. As documented by Krueger and

¹ Committee on National Statistics, National Academies – Washington, DC. E-mail: cmackie@nas.edu.

² Organisation for Economic Co-operation and Development – Paris. E-mail: Conal.SMITH@oecd.org.

Schkade (2008), OECD (2013), and elsewhere, a near exponential growth in the publication of articles on happiness and related subjects has emerged since the 1990s, including in the top economics journals.³

Reflecting increasing interest in SWB from researchers, policy-makers and the public—and further adding to its legitimacy—the report of the *Commission on the Measurement of Economic Performance and Social Progress* (Stiglitz et al., 2009) recommended that national statistical agencies collect and publish measures of SWB. This was followed in 2013 by the publication of the *OECD Guidelines on the Measurement of Subjective Well-being*, aimed at encouraging the collection and publication of such measures by national statistical offices. In the same year the National Academy of Sciences published *Measuring Happiness, Suffering, and Other Dimensions of Experience*, which investigated the application of experienced well-being measures to policy in the United States.

A large number of national statistical offices are now collecting SWB measures either on an experimental basis or as part of their core programs. Among OECD countries, 32 out of 34 national statistical offices collect data on life satisfaction along the lines recommended by the *OECD Guidelines*. The United Kingdom, for example, now collects four measures of SWB aligned with the *OECD Guidelines* in its *Annual Population Survey*, providing a total sample of approximately 160,000 each year. On 3 September 2014, the UK Statistics Authority granted these four measures accredited National Statistics status, confirming them as part of the highest tier of official statistics in the UK.

With the increasing prominence of SWB in official statistics, it is useful to review their conceptual scope and to consider how they can be applied to policy. While the majority of measurement initiatives and academic research have focused on how people evaluate their lives (often, misleadingly, described as measures of "happiness"), widespread consensus has emerged that SWB has multiple distinct dimensions. This paper presents a general overview of SWB and its underlying complexity, then discusses the implications for data collection, measurement, and informing policy.

2. What is subjective well-being?

Subjective well-being encompasses several separate but interacting aspects of people's feelings, attitudes, and experiences. The construct covers a number of different aspects of a person's subjective state; however, there is debate about exactly what elements should be emphasized (Diener et al., 1999; Kahneman,

³ The *Journal of Economic Literature*, the *Journal of Economic Perspectives* and the *Journal of Political Economy* have all published papers on subjective well-being in the recent past (Frey and Stutzer, 2002; Di Tella, MacCulloch, and Oswald, 2001; Rayo and Becker, 2007, among many others).

Diener, and Schwarz, 1999). Kahneman and Krueger (2006), for example, focus primarily on experienced well-being, while Huppert et al. (2009) emphasize measures of good psychological functioning. The OECD *Guidelines* (2013) defines SWB as involving “good mental states, including all of the various evaluations, positive and negative, that people make of their lives, and the affective reactions of people to their experiences⁴.” This characterization is inclusive in nature, encompassing a broad dimensional conceptualization of SWB. In particular, the reference to good mental functioning acknowledges concepts such as interest, engagement, and meaning alongside more commonly identified notions of satisfaction and emotional state. Similarly, Diener (2006) argues that “subjective well-being is an umbrella term for the different valuations people make regarding their lives, the events happening to them, their bodies and minds, and the circumstances in which they live.”

Two definitional points are worth making here. First, SWB is narrower in scope than are self-reported measures in general which may be directed toward outcomes that have no relationship to mental states. For example, a survey may ask respondents to report income, marital status, or employment information, none of which is directly a measure of SWB (though they certainly may be correlates to it). Second, SWB is not necessarily synonymous with well-being as a whole. In the measurement of human welfare, largely non-subjective variables such as income levels, health status, knowledge and skills, environmental quality and social connections often play important roles.

In order to develop meaningful measures of SWB, it is essential to identify which of its elements is to be the central focus. Although some researchers argue in favour of a single overall construct (Stewart-Brown and Janmohamed, 2008), SWB is more commonly acknowledged to encompass three core dimensions—life evaluation, experienced or hedonic well-being, and eudaimonia (which includes concepts such as sense of purpose or meaning and locus of control)—and that each should be measured. Although both the OECD *Guidelines* (2013) and NAS (2013) recommend measuring these dimensions separately, they also recognize that they are interrelated. Many of the distinctions in SWB measurement constructs relate to their temporal characteristics which may be thought of in terms of a continuum, with essentially real-time assessments of experience, emotional state, or sensations at the shortest end of the spectrum and overall evaluations of life satisfaction, purpose, or suffering at the other end (the longest reference period). Sense of meaning or purpose may impact a respondent's assessment of either a momentary situation (why do I not mind reading Dr. Seuss to my child over and over again?) or to life evaluation (will studying 15 hours a day to become a physician lead to a better life?) (NAS, 2013, p. 15). Next, we define and review the major dimensions of SWB—life satisfaction, affect and experienced well-being, and eudaimonia.

⁴ This definition derives from Diener et al. (2006).

2.1. Life evaluation

Life evaluation is conceived of as resulting from a reflective assessment of a person's life. Such assessments entail judgement by an individual, and stand in contrast to descriptions of mood or emotional state. Pavot et al. (1991) describe the evaluative process as involving individuals constructing a "standard" that they perceive as appropriate for themselves, and then comparing the circumstances of their life to that standard. Although it is not clear whether this process of comparison is a conscious one, in practice, the relatively short response time associated with life evaluation questions in surveys suggests that respondents will typically use a heuristic to form a rating (OECD, 2013).

It is tempting to equate life evaluation with an economist's definition of utility as the criteria by which different choices are evaluated. There is a *prima facie* plausibility to the idea that people pursue goals that maximise the evaluation of their lives, a view that has found significant empirical support (Clark, 2001; Clark, Frijters, and Shields, 2008; Frijters, 2000; Helliwell and Barrington-Leigh, 2010). However, there are also strong reasons to be cautious in treating measures of life evaluation as measures of utility. First, although economists traditionally assume (at least implicitly) that the remembered utility on which people base their decisions is equivalent to the sum of momentary utilities associated with moment-by-moment experience, SWB measurement has revealed this to not always be the case. Life evaluations are based in part on how people remember their experiences, which can differ significantly from how they actually experienced things at the time (Kahneman et al., 1999). For example, the so-called "peak-end rule" states that a person's evaluation of an event is based disproportionately on the most intense (peak) and last (end) emotions experienced during the event, rather than the average or integral of emotional experiences over time. A second critique of the view that life evaluation measures utility focuses on the observation that people are prepared to trade off life satisfaction in order to achieve other outcomes. If measures of life evaluation fully captured utility as conceived of by economists, the notion of people accepting reduced levels of life evaluation in order to gain some other goal would make little sense.⁵ Despite these concerns, measures of life evaluation remain of high interest for two reasons. First, although life evaluation is probably not measuring an economist's conception of utility, other approaches to analysing utility also have limitations. Measures of life evaluation can therefore add to the sum total of knowledge without themselves being perfect measures of utility. Second, regardless of whether life evaluations measure utility, how people feel about their lives is an important consideration in its own right. Life evaluations may provide insights into people's well-being more generally, even if they do not align perfectly with some over-arching view of the concept.

⁵ Benjamin, Kimball, Heffetz, and Rees-Jones (2013) examine such trade-offs in the context of stated preferences in the residency choices of medical students.

As a reflective construct of a respondent's subjective state, life evaluation is usually measured through one or more survey questions. Perhaps the best validated measure in the psychological literature is the 5 question Satisfaction With Life Scale developed by Diener and Pavot (1993). However, the two most widely used measures are single item survey questions: the Self Anchoring Striving Scale (more commonly known as the Cantril Ladder) and the *World Values Survey* Satisfaction With Life question. The Cantril Ladder is used in the *Gallup World Poll* and is thus the basis of much recent research on the drivers of life evaluation across countries. Until recently it was believed that the Cantril Ladder and the Satisfaction With Life question collected slightly different information with the former being the more purely evaluative of the two. Recently, however, evidence from the Gallup World Poll and research based on split sample surveys in the UK *Household Opinion Survey* has provided convincing evidence that the two questions are closely comparable (Helliwell, Layard, and Sachs, 2013; ONS, 2011). In part due to this finding, OECD (2013) recommends a life evaluation question based on the simpler World Values Survey version:

The following question asks how satisfied you feel on a scale from 0 to 10. Zero means you feel "not at all satisfied" and 10 means you feel "completely satisfied".

Overall, how satisfied are you with life as a whole these days?" [0-10]

This question, with minor variations, forms the basis of most subjective well-being data currently collected by national statistical agencies.

2.2. Affect, experienced well-being

Affect is the term psychologists use to describe a person's feelings. Affect can be thought of as particular feelings or emotional states and is typically measured with reference to a particular point in time. Such measures capture how people experience life rather than how they remember it (Kahneman and Krueger, 2006). While an overall evaluation of life can be captured in a single measure, affect has at least two distinct—so-called hedonic—dimensions associated with positives and negatives (Kahneman et al., 1999; Diener et al., 1999). Positive affect captures emotions such as happiness, joy, and contentment. Negative affect comprises the experience of unpleasant emotional states such as sadness, anger, fear, and anxiety. While positive affect is thought to be largely uni-dimensional (in the sense that positive emotions are strongly correlated with each other), negative affect is more multi-dimensional. For example, it is possible to imagine at a given moment feeling anger but not fear or sadness.

Bradburn (1969) was one of the first researchers to determine that positive and negative affect are not opposite ends of one dimension but are largely independent of one another; a person can rate highly on one state and either high or low on the other. Bradburn's findings have been replicated many times; for

example, Gere and Schimmack (2011) found that, even after correcting for measurement error and bias, positive and negative feelings were distinct. This body of research evidence led the NAS panel to conclude that:

Both positive and negative emotions must be accounted for in experienced well-being measurement, as research shows that they do not simply move in an inverse way. For example, an activity may produce both negative and positive feelings in a person, or certain individuals may be predisposed to experience both positives and negatives more strongly. Therefore, assessments of [experienced well-being] should include both positive and negative dimensions in order for meaningful inferences to be drawn (p. 39).

Other dimensions of experienced well-being such as arousal, which relate to positive and negative emotions in a range of ways, are important as well. Sensations such as pain, numbness, heat, or cold may also figure into emotional states and into hedonic assessment of those states—particularly if the context is people’s health or housing conditions. Certainly, people experiencing pain will on average report higher levels of negative well-being, all else being equal (Krueger and Stone, 2008).

The term hedonic well-being typically is used in association with the emotional (or affect) component SWB. And, although the term "experienced well-being" is sometimes treated synonymously with the affect, they are not identical. Experienced well-being is broader in the sense that it may include pain and other sensations that factor into suffering or happiness which may be omitted by the narrower hedonic focus on emotions. Even more broadly, as described below, appraisals of concepts beyond the emotional, such as meaning or purpose, may also be included in the experienced well-being construct (NAS, 2013). Measuring “experience” broadly is essential for addressing issues of long-term suffering which are of concern to policy makers. As elaborated below, these characteristics carry also implications for data collection strategies.

The characteristics of affective states also raise an interesting question about their relationship to life evaluation. Research has established that positive and negative experience track at least partially independently of life satisfaction and of each other. Kahneman et al., (1999) argue for the existence of a “good/bad” axis on which people are able to place experiences based on their emotional states at the time. In principle, this process is similar to that involved in forming life evaluations from remembered affective states. Kahneman’s point is that affective states can be compared and that one can therefore reasonably aggregate measures of current affect. For this reason, affect measures are sometimes reported in terms of affect balance, which captures the net balance between positive and negative affect (Kahneman and Krueger, 2006).

A number of measurement approaches have been used to measure affect and, more broadly, experienced well-being. Sometimes approach is dictated by the

measurement objective; sometimes it is constrained by survey (or other) data gathering instrument. The basic categories are:

- Ecological momentary assessment, which signals a person to respond in the moment. Sometimes considered the gold standard for measuring affect, in the experience sampling method (ESM), participants are prompted to record their feelings and perhaps the activity they are undertaking at either random or fixed time points, usually several times a day, throughout the study period, which can last several weeks. To maximise response rates and ensure compliance throughout the day, electronic diaries are often used to record the time of response. While the ESM produces an accurate record of affect, it is also expensive to implement and intrusive for respondents.
- Reconstructed activity-based measures; time use/day reconstruction methods (DRM) allow contextual information to be linked to measures associated with specific activities (e.g. job search, child care, commuting) and in turn to policy questions. DRM, in which respondents are questioned about events from a time-use diary recorded on the previous day, are often more practical and viable for government surveys. Research has shown that the DRM produces results comparable with ESM, but with a much lower respondent burden (Kahneman et al., 2004).
- Single day measures, which ask respondents about their experiences globally for a given day or episodes during that day. Surveys are typically administered at the end of day or the next day. A number of important survey and measurement issues arise when single day measures are used to approximate results of momentary measures (these are dealt with in the paper by Dylan Smith in this volume).

Experience Sampling, DRM, and similar methods for collecting affect data in time-use studies allow for analyses capable of associating particular affective states with specific activities. Measures of affect collected in this way thus capture well-being as reported by a person in a particular place, time, and set of circumstances as opposed to some sense of overall SWB. It is also possible to collect affect data in general household surveys via questions about a person's mood or emotional state over a particular recall period.⁶ However, although such measures capture information on a person's affective state, they cannot easily capture information linking affect to particular activities. On the other hand, it is also possible to collect information about some aspects of eudaimonia (see the next section) using similar techniques to those used to measure experienced affect. For example, the American Time Use Survey well-being module collects information on experienced "meaning and purpose" associated with specific daily episodes.

⁶ The Gallup World Poll contains a range of questions on affect during the previous day, which have been extensively tested. The UK Office of National Statistics has collected similar measures of affect in its Integrated Household Survey programme.

2.3. Eudaimonia

A substantial literature exists on the concept of good psychological functioning, sometimes referred to as “flourishing” or “eudaimonic” well-being (Huppert et al., 2009; NEF, 2009; Clark and Senik, 2011; Deci and Ryan, 2006). Eudaimonic well-being goes beyond a respondent’s reflective evaluation and emotional states to focus on functioning and the realisation of the person’s potential. In developing the questionnaire on psychological well-being for the *European Social Survey*, for example, Huppert et al. (2009) characterise the “functioning” element of well-being as comprising autonomy, competence, interest in learning, goal orientation, sense of purpose, resilience, social engagement, caring and altruism. Eudaimonic conceptions of SWB thus differ significantly from the evaluative and affective components in that they are concerned with capabilities as much as with final outcomes. Because the measurement of eudaimonia identifies a central role for people’s “needs” or “goals”, the approach represents a useful response to the criticism that the measurement of SWB is built purely on hedonistic philosophy, and also aligns itself with many people’s perceptions of what is important to value in life.

While a consensus has emerged regarding the distinction between life evaluation and affect, the conceptual structure of eudaimonic well-being is less well fleshed out. It is not clear, for example, whether eudaimonic well-being describes a uni-dimensional concept in the sense of life evaluation, or a range of different concepts. It is clear, however, that eudaimonic measures capture important aspects of people’s subjective perceptions about their own well-being that are not covered by either life evaluations or affect. For example, having children has a negligible (or even mildly negative) correlation with average levels of life evaluation (Dolan, Peasgood, and White, 2008), and child care (even of one’s own children) is associated with relatively low levels of positive affect (Kahneman et al., 2004). This conflicts with the intuitive assumption that children, at least for those who choose to have them, contribute in some way to their parent’s well-being. Indeed, people with children report higher average levels of meaning or purpose in their lives than other respondents (Thompson and Marks, 2008).

Concepts of “worthwhileness” or purpose appear crucial for understanding (and even predicting) behaviour, specifically why and when people engage in various activities or how they make decisions affecting their life course. White and Dolan (2009), for example, use day reconstructions to measure rewards associated with various daily activities. They find discrepancies between activities that people find pleasurable versus those found to be rewarding or meaningful. As noted above, activity based data indicate that time spent with children is relatively more rewarding than pleasurable, whereas time spent watching television is relatively more pleasurable than rewarding (NAS, 2013, p. 19). Similarly, people do many things that are pleasant even if they are not viewed as having much long-

term meaning or positive impact on future well-being. Either the pleasure or the purpose may be drivers of behaviour (Kahneman and Krueger, 2006).

While there is less agreement on the appropriate strategy to adopt when measuring eudaimonia than is the case for life evaluations or experienced well-being, three different measurement approaches have emerged: Economists have focused on meaning and purpose as the element of eudaimonia that most clearly captures additional information to other dimensions of SWB and which can be clearly distinguished from personality (Dolan, Layard, and Metcalfe, 2011). This has been reflected in the inclusion of a single question on meaning and purpose in the measures of SWB collected by the UK Office for National Statistics. An alternative strategy is that adopted in the well-being module of the European Social Survey (Huppert et al., 2009) where a battery of questions relating to different aspects of psychological well-being is collected, allowing for an analysis of the different concepts that are grouped together as eudaimonia. Finally, an extensive literature has emerged on a measurement of well-being grounded in mental health promotion. Mental well-being incorporates many of the elements of eudaimonia, but also combines these with measures of life evaluation and experienced well-being to provide a single index of overall psychological flourishing. A good example is the Warwick-Edinburgh Mental Well-being Scale (Tennant et al., 2007).

2.4. Relationships among the dimensions of SWB

While life evaluation, experienced, and eudaimonic well-being are all conceptually distinct, it is helpful to understand how they relate to one other. In the SWB continuum identified above, one end is demarcated by a point-in-time reference period and is purely hedonic (“How do you feel at this moment?”) while the other involves evaluation of a comparatively very long reference period (“Taking all things together, how would you evaluate your life?”). Momentary assessments of affect represent the shortest framing period while global assessments of affect over the past day or even several weeks are at the longer end for experience measures. As the reference and recall periods lengthen, a measure is less dominated by actual experience and is more influenced by personality and/or cognitive reflection. Specification of the reference period has a determinative impact on the results of a survey and, indeed, on what nature of what is being measured (NAS, 2013, p. 29).

Figure 1 below provides a simple model of the different elements of a SWB measurement framework. The model emphasises three dimensions involved in the measurement of SWB. These are: (1) the measurement concept; (2) the sub-components of well-being; and (3) their determinants. The list of determinants and sub-components in the figure is illustrative rather than exhaustive – the model is intended to serve as an organising framework for thinking about the scope of SWB.

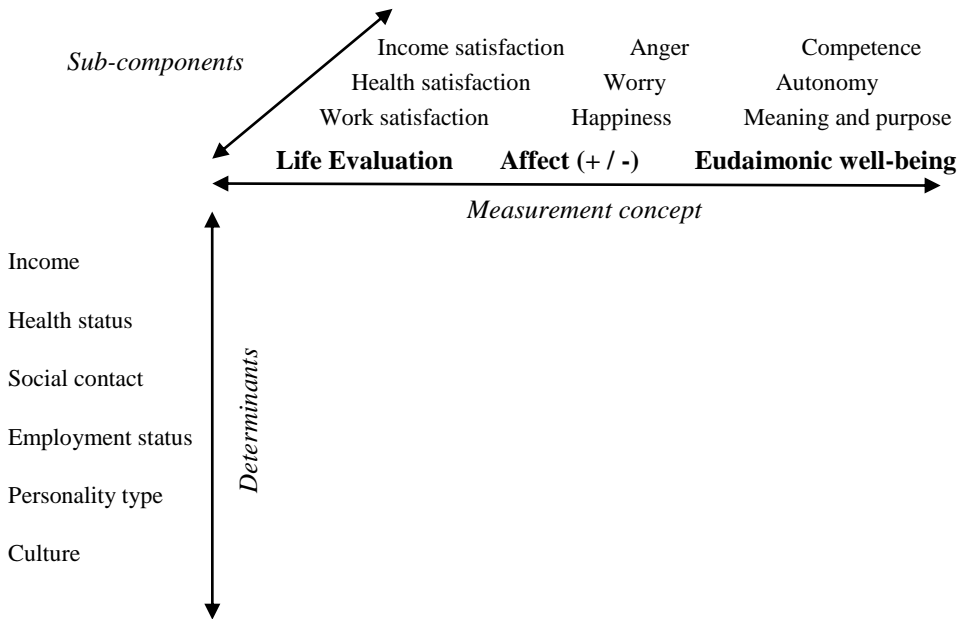


Figure 1. A simple model of subjective well-being

Source: OECD, 2013.

Empirically, there is extensive evidence about the relationship between measures of affect and overall measures of life evaluation. Diener, Kahneman, Tov, and Arora (in Diener, Helliwell, and Kahneman, 2010) reveal a high correlation (0.82) across countries between the most commonly used average measures of life evaluation, but a much lower correlation (0.55-0.62) between average affect balance and either of two life evaluation measures (life satisfaction and the Cantril Ladder). Similarly, at the individual level, Kahneman and Krueger (2006) report only a moderate correlation (0.38) between life satisfaction (an evaluative measure) and net affect.

Above, we have already hinted at how eudaimonia relates to experienced and evaluative well-being. For measurement, it may not make much difference whether sense of purpose contributes directly to positive or negative emotions or is positioned alongside but separate from them as a distinct sentiment. What matters is that the adjectives for purpose (e.g. fulfilment) are distinct from those used for pleasure (e.g. fun) and that a range of good feelings, emotions, or sentiments contributes to overall well-being.

A body of evidence exists on the empirical relationship between eudaimonic well-being and other dimensions of SWB suggesting that the correlation is smaller than is the case between affect and life evaluations. Clarke and Senik (2011) report a correlation between life satisfaction and four different aspects of eudaimonic well-being of between 0.25 and 0.29. Diener et al. (2009) report a correlation of 0.62 ($N=563$, $p < .001$) between their Psychological Well-Being

Scale and the evaluative Satisfaction with Life Scale, and correlations of 0.62 and 0.51 respectively between the Psychological Well-Being Scale and the positive and negative subscales of the Scale of Positive and Negative Experience ($N=563$, $p < .001$ in all cases). Huppert and So (2009) found a correlation of 0.32 between flourishing and life satisfaction in European Social Survey data. Among the European Social Survey sample overall, 12.2% met the criteria for flourishing, and 17.7% met the criteria for high life satisfaction, but the percentage for both flourishing and high life satisfaction was 7.2%.

Table 1 below gives the correlations between individual measures of life evaluation derived from the Gallup World Poll (life satisfaction), positive affect, negative affect and eudaimonic well-being (purpose) across 362 000 respondents in 34 OECD countries. The correlation is highest between the two measures of affect, at -0.3855, and lowest between purpose and negative affect, at -0.091. Life satisfaction has a correlation of about 0.23 with both measures of affect, and of 0.13 with purpose. While all the coefficients in Table 1 show the expected sign and all are significant at the 0.1% level, none of the measures have a correlation near 1, indicating that the different measures capture different underlying phenomena.

Table 1. Correlation coefficients among purpose, life satisfaction, positive affect, and negative affect at the individual level, 2006-2010

	Purpose	Life satisfaction	Positive affect	Negative affect
Purpose	1.000			
Life Satisfaction	0.134	1.000		
Positive Affect	0.142	0.229	1.000	
Negative Affect	-0.091	-0.231	-0.3855	1.000

Note: The precise measures used are the so-called “Cantril Ladder” for life satisfaction, an “important purpose” in life for purpose, and the sum of “yes” responses to smiled yesterday, experienced joy yesterday, and was well rested yesterday for positive affect and an equivalent index based on experience of sadness, worry and depression for negative affect.

Source: Gallup World Poll.

Because dimensions of SWB are distinct, and cover different reference periods, they can go in different directions. For example, studying hard for years to become a surgeon or working in devastated areas of the globe to alleviate poverty may not be immensely pleasurable but may ultimately yield high life satisfaction or reported sense of purpose. Individuals who have a longer-term focus and are more "achievement oriented," may at times sacrifice daily experience for longer term objectives and anticipated life satisfaction in the

future. The fact that people exhibit high and low discount rates means that they do not all have same focus. Individuals who focus primarily on daily experiences—due to low expectations, lack of agency, or imposed social norms—may have less incentive to invest in the future.

Relative to life satisfaction, experienced well-being is more directly related to the environment and context of people's lives. Using data from the Gallup World Poll, Deaton (2012) found, for example, that health state correlates more strongly with experienced well-being (though it is also important for evaluative well-being) as are marital status and social time (see also Boarini et. al., 2012). Other aspects of daily behaviour, such as the nature of a person's commute to work and the nature of a person's social networks, are reflected in positive and negative affective states (separable aspects of experienced well-being). The quality of people's daily experiences is also linked to health status and other outcomes via channels such as worry and stress on the one hand and pleasure and enjoyment on the other. Evaluative well-being, while also sometimes influenced by these factors, is more likely to reflect people's longer-term outlook about their lives as a whole. It may also be related to, and reflected in, longer-term behaviours such as investments in health and education. These distinctions make experience measures ideal for assessing emotions as they fluctuate from moment to moment and in response to day-to-day events and activities. In contrast, life satisfaction is more likely to reflect general, long-lasting factors such as unemployment, income, or a happy marriage, although it is easy to see how these circumstances could directly impact emotions on a day to day basis as well (NAS, 2013, p. 92).

These nuances and interactions led the NAS panel to conclude that: "To make well-informed policy decisions, data are needed on both [experienced well-being] and evaluative well-being. Considering only one or the other could lead to a distorted conception of the relationship between SWB and the issues it is capable of informing, a truncated basis for predicting peoples' behaviour and choices, and ultimately compromised policy prescriptions" (p. 93). A similar view is expressed in the *OECD Guidelines on Measuring Subjective Well-being*, which recommends that measures of affect and eudaimonia be collected alongside measures of life evaluation because they capture different aspects of SWB (with a different set of drivers) and because the different measures are affected in different ways by cultural and other sources of measurement error.

3. The role of national statistics in SWB measurement

3.1. Principles of official statistics

Official statistics are produced to meet the needs of policy-makers in planning and assessing the impact of policy decisions, and to inform the general public about the state of society. Academics and the media are also important users of official statistics, contributing to a better understanding of society and informing the public and decision-makers.

The principles of official statistics generally reflect the view that information is collected only when there is good reason and for a clear purpose. The OECD framework for data quality identifies relevance as the first of the seven key dimensions of quality. Relevance implies that the value of data “is characterised by the degree to which that data serves to address the purposes for which they are sought by users” (OECD, 2013). Similarly, the *United Nations Fundamental Principles of Official Statistics* asserts that the role of official statistical agencies is to compile and make available “official statistics that meet the test of practical utility... to honour citizens’ entitlement to public information.”

There are sound ethical and practical reasons why official statistical agencies insist on having a clear understanding of the uses of any proposed statistical measures. Many official statistical agencies have the power to compel responses from respondents. That is, respondents are legally required to provide information when approached by a national statistical agency. The corollary of such authority is the requirement for national statistical offices to use data responsibly. From an ethical standpoint, only information that is sufficiently important to justify the intrusion into respondents’ lives should be collected. The International Statistical Institute’s *Guidelines on Professional Ethics* notes that:

Statisticians should be aware of the intrusive potential of some of their work. They have no special entitlement to study all phenomena.

Over and above this ethical concern is a practical one. Even if compliance is legally mandated, the quality of resultant data depends heavily on preserving a good relationship between respondents and the official statistical agency. This is undermined if the statistical agency cannot articulate why the data being collected is important and how it will be used. Additionally, statistical agencies must be careful not to over-burden respondents and jeopardise the good will on which high-quality responses depend. Because of this, collecting measures of SWB will have an opportunity cost in terms of other data that will not be collected in order to produce such measures. If SWB measures are to be included in official statistics, therefore, it is essential to be clear about how they will be used.

3.2. Comparative advantages/disadvantages of government surveys

The fact that NSOs have historically led the way in the development of population surveys—both general and specialized (e.g. health interview surveys, time use, neighbourhood environment) – for research purposes puts them at a comparative advantage for collecting data on some dimensions of SWB, and perhaps at a comparative disadvantage for others. Traditional government surveys work especially well for large, cross-sectional formats. Life satisfaction or global yesterday questions of the type developed by the UK ONS are easily incorporated. Cross-sectional surveys are most often used to address group differences—for example, in the SWB context, are older people happier than younger people? Are females more stressed than males? Or, do people in high income countries report higher life satisfaction than those in low income

countries? A prime attraction of including SWB questions in large government surveys is their ability to accurately detect these group differences in a minimally burdensome way.

Among the key strengths of large scale surveys run by national statistical agencies are large sample sizes, high response rates, and ability to spread the enumeration out over a long period of time. These factors are important as they compensate for the weaknesses inherent in many research surveys, including those forming the basis of much of the SWB literature such as the *World Values Survey* and the *Gallup World Poll*. In both these cases, small national samples and short periods of enumeration make the error terms associated with SWB measures large, and raise the risk that transient events such as the weather, holidays, or news stories will impact on how people respond in unintended ways.⁷ By way of contrast, a large sample size reduces the error term associated with SWB measures and a long (ideally annual) enumeration period will largely eliminate measurement bias due to one-off events.

More generally, surveys carried out by government statistical agencies generally collect higher quality information on potential covariates – such as income, labour force status, or education – than is possible in smaller unofficial surveys. Because the quality of SWB analysis depends not only on the quality of the SWB measure, but also on the quality of the other measures used in the analysis, surveys from national statistical agencies offer the opportunity for analysis not possible otherwise. For example, the lack of high quality income measures in surveys that include subjective well-being questions has been a factor limiting research in a number of areas. The relationship between income and subjective well-being has been a subject of interest since 1974 when Richard Easterlin identified the so-called “Easterlin paradox”: that higher income is associated with higher happiness both between individuals and across countries, but there is no evidence that average happiness increases as average income increases over time (Easterlin, 1974). Understanding the causes of the Easterlin paradox is a high research priority because of the implications the paradox has for a range of policies. On a more technical level, one of the main policy uses for measures of subjective well-being is estimating the value of non-market outcomes. This involves obtaining precise measures of the impact of people’s own income on their subjective well-being and comparing this to the impact of marginal change in the non-market outcome in question on subjective well-being.

For both better understanding the Easterlin paradox and estimating the value of non-market outcomes, the quality of income measures in surveys is at least as important as the quality of subjective well-being measures. While national statistical offices collect high quality information on household income, and are increasingly collecting measures of subjective well-being, there are currently few data sources that bring the two together. Those surveys – both official and non-

⁷ For some measurement objectives, usually associated with experienced well-being, it is appropriate to factor in such influences. For life evaluation, it typically is not.

official – that include measures of subjective well-being typically collect income only in broad bands, and in the case of non-official surveys, often also have very high item-specific non-response rates for the income question. Filling this gap is a priority for the near future.

Another key strength of collecting data through large scale government surveys is the ability to conduct high-quality experimental trials to establish the impact of different methodological issues. For example, a split-sample randomized trial using experimental national data conducted by the UK Office for National Statistics (ONS) reported an effect of question order on multiple-item positive and negative emotion questions (Office for National Statistics, 2011). Asking negative emotion questions first produced lower scores on some positive emotion items for the adjectives “relaxed,” “calm,” “excited,” and “energized.” When positive emotion questions were asked first, the mean ratings for negative emotion questions were generally *higher*—except in the case of “pain”—and the increase was statistically significant for the adjectives “worried” and “bored” (OECD, 2013, p. 87). Similarly, when the order of positive and negative adjectives was varied, Krueger et al. (2009) observed higher ratings of positive emotions in a positive-to-negative order and lower ratings of negative emotions in a negative-to-positive order.

3.3. Measurement hurdles

Measuring SWB faces a range of methodological challenges. Almost all of these are shared with other survey measures, including those of notionally “objective” outcomes, but there is reason to believe that some of these issues may be more significant for SWB measures than for many other subjects⁸. Among these are context effects (such as the weather at the time of the interview, sports news on the day), framing effects (such as question order), mode effects (how the survey was carried out) and potential cultural bias⁹. These factors can affect answers to questions on life evaluation, affect, or eudaimonia.

Survey Mode is also a major methodological concern (see OECD, pp. 102-108). Dolan and Kavetsos (2012) investigated the differences between interviewer-administered and telephone-administered responses to the UK Annual Population Survey. The authors examined (a) the impact of survey mode on SWB reports and (b) the determinants of SWB by mode, using the April-September 2011 pre-release of the survey data. Their analysis found large differences by survey mode; in fact, mode effects in the data swamped all other effects.

Although the methodological challenges associated with collecting information on SWB are real, it is important not to overstate them. For some

⁸ While the range of measurement issues – to do with survey context, question ordering, survey mode and many other factors – are briefly touched on in this paper, they are discussed in greater detail in papers by Lucy Tinkler and Paul Allin in this volume.

⁹ For a comprehensive review, see Schwarz and Strack (1999).

questions, researchers may want to include the influence of context. For example, when looking at experienced well-being while using public parks, the context is not a contamination of the measure: it is the object of study. In some cases buffer and transition questions that precede and follow SWB question modules may help reduce or eliminate context or framing effects. For example, Deaton (2010) shows how including a buffer or transition question between political questions and life-evaluation questions largely eliminates a previously detected item-order effect. More generally, when the goal is to draw conclusions about a population, only influences that affect the sample but not the population as a whole, undermine the purpose of assessment. Provided that surveys are conducted in a consistent fashion (i.e., without changes to mode or question wording) and are enumerated over a long period of time, these issues are not generally significant.

Cultural bias is potentially a more difficult form of error to address. Taken in a general form, cultural bias can be thought of as differing response styles across different groups in the population of interest. In this case, even use of the same survey methodology at the same time will not eliminate sources of bias. A particular concern in this respect is the comparison of average levels of SWB between countries as there is *prima facie* evidence that response styles do vary between countries (OECD, 2013), and this will have an important impact on the inferences drawn from the data.

Several special challenges arise when measuring experienced well-being. A number of national and international surveys have used single-day assessments to measure experienced well-being—that is, assessments that target affect or broader experience for a single day. In the US, for example, the Health and Retirement Study, the Disability and Use of Time supplement to the Panel Study of Income Dynamics, and the Gallup-Healthways survey employ single-day hedonic assessments; as do the English Longitudinal Survey of Ageing and the surveys on well-being of the UK Office for National Statistics (ONS) (NAS, 2013, p. 52). Typically, these surveys ask respondents about their experiences from the previous day. The NAS (p. 55) report concluded:

Global-yesterday measures represent a practical methodology for use in large population surveys. Data from such surveys have yielded important insights—for example, about the relationships between experienced well-being and income, age, health status, employment status, and other social and demographic characteristics. Research using these data has also revealed how these relationships differ from those associated with measures of evaluative well-being. Even so, there is much still to be learned about single-day measures.

One practical limitation of *end-of-day* – as opposed to global yesterday, which are often the default for large surveys) measures, and a reason that they have not been used more by statistical agencies – is that large population surveys often depend on telephone interviews conducted throughout the day, not just at the end

of the day. Because of the survey timing requirement, end-of-day instruments have typically been less practical for use in general surveys. However, newer technologies, such as use of interactive mobile phone assessments, may offer solutions to some of the data collection constraints associated with end-of-day methods (NAS, 2013, p. 53). There has been little systematic research into how the recall and contextual influences act differentially between end-of-day and global-yesterday measures, and how well either correlates with averages from momentary readings¹⁰.

Additionally, global-yesterday measures are limited in terms of creating a more detailed understanding of the drivers of experienced well-being over the course of the day (e.g. variation at the individual level). For this level of analysis, momentary assessments or, at the least, time-use or activities-based data – for example, data generated by day reconstruction methods (DRM) – are needed (NAS, 2013, p. 55). For some research and policy questions, contextual information about activities engaged in specific behaviours and proximate determinants is essential. For example, to investigate how people feel during job search activities, while undergoing medical procedures, or when engaged in child care, something more detailed than a global daily assessment is needed. Activity based measures attempt to fill this measurement need (NAS, 2013, p. 59).

An attractive feature of DRM is its capacity to combine time-use information with the measurement of affective experiences. Capturing the time-use and activity details of survey respondents enhances the policy relevance of experienced well-being measures by embedding information about relationships between emotional states and specific activities of daily life (NAS, 2013, p. 66). Additionally, for large surveys, DRM can be administered with less intrusion and lower burden than momentary assessment tools while still gathering fairly rich and detailed information. By asking participants to first recall the events of their day and then provide ratings associated with them, DRM exploits the fact that, while memories of ongoing experiences such as pain and mood are flawed, Memory for discrete events is more accurate (Robinson and Clore, 2002) (NAS, 2013, p. 60). For some questions (e.g. predicting consumer behaviour or whether or not a person is likely to repeat a medical procedure), a reconstructed assessment of experienced well-being may be more relevant than EMA; it may also be better at predicting a policy's impact on people's choices, but worse at assessing a policy's impact on experience.

This kind of data collection has already been successfully developed by statistical agencies. In the United States, ATUS has, since 2010, included a module asking respondents about feelings (pain, happiness, stress, sadness, tiredness) during specific episodes of the day. The ATUS SWB module is an

¹⁰ Though Christodoulou et al. (2013) compared to DRM – see Dylan M. Smith in this volume. The validity of different measures addressed in a paper by Paul Allin in this volume.

abbreviated version of a DRM approach (NAS, 2013, p. 23). Regarding the DRM, the NAS report concluded:

Preliminary assessment of DRM measures of mood and physical symptoms suggests that they reasonably approximate summary measures created from EMA protocols. An attractive feature for survey objectives is that the DRM approach goes beyond simply addressing who in the surveyed population is happy to identifying when they are happy. Additionally, it appears that the DRM is less burdensome on respondents than experience sampling, and it might reduce memory biases that are inherent in global recall of feelings. The DRM is thus a promising method for assessing feelings, mood, and physical symptoms that accompany situations and activities more efficiently than with EMA methods and with greater specificity and accuracy than traditional recall-based methods (NRC, 2013, p. 63)¹¹.

Similarly, INSEE (the French national statistical agency) has collected data on experienced well-being through the French time use survey – the *Enquete Emploi du Temps 2010*. This survey used a different approach to the DRM strategy adopted by the ATUS SWB module. Rather than collecting detailed information on multiple different affective states for just three episodes in each diary day, the *Enquete Emploi du Temps* requires respondents to rate each activity in the time use diary on a 7 point scale from very unpleasant to very pleasant. This collects a far more comprehensive picture of the activities sampled at the price of less detail on each activity. The relative strengths and weaknesses of the French and American approaches is an area for further research.

For obvious reasons, surveys do not work easily for momentary assessment. The required instruments are difficult to scale up to nationally representative surveys and impose a high respondent burden. This said, monitoring and survey technologies are changing rapidly and the ways in which government agencies administer surveys will surely evolve alongside and new measurement opportunities will come on line. Considered in terms of comparative respondent burden, it may become less intrusive to respond to a smartphone beep than to fill out a long-form survey. So, while EMA may not now be practical for flagship population surveys, real-time analyses may become so. As technology advances, such modes could become feasible, even for large-scale surveys at reasonable cost. Large-scale (more general) surveys could build in the possibility of mapping the data from single-day measures with the data from more detailed studies for a subset of the sample. Experiences in real time, because they are especially relevant to health, have been incorporated into health examination surveys, so there is precedent. It is also possible to monitor blood pressure and other physical signals related to affect in real time (NAS, 2013, p. 51).

¹¹ Smith, et al. (2012). A Test Comparison of EMA and DRM estimates supports above conclusion.

3.4. Current state of play

When the report of the Commission on the Measurement of Economic Performance and Social Progress (Stiglitz et al.) was released in 2009, the national statistical offices of only three OECD countries collected regular measures of life evaluation (Canada, Israel, New Zealand) and, of these, only Statistics Canada conformed to current best practice. No OECD country was regularly collecting measures of affect or eudaimonia. In the six years since then, this situation has transformed. Among the 34 OECD countries, 32 now collect measures of SWB – mostly life evaluation – through their national statistical agencies in a way that is broadly comparable.

Table 2. Subjective well-being measures in official statistics

EU-SILC coverage (2013 ad-hoc module; life sat, affect and eudaimonia – freq. tbc.)		European countries with additional collections	Other OECD countries with SWB measurement initiatives	OECD countries with no current NSO data collection
OECD countries (25): Austria Belgium Czech Republic Denmark Germany Estonia Finland Greece Ireland Spain France Iceland Italy Luxembourg Hungary Netherlands Norway Poland	Portugal Slovenia Slovakia Sweden Switzerland Turkey UK Non-OECD countries (7): Bulgaria Croatia Cyprus FYROM Latvia Lithuania Malta	<ul style="list-style-type: none"> • Austria (life sat, 2004-2012) • France (life sat*, in 2011; affect in 2010 time-use survey; freq. tbc.) • Italy (life sat*, from 2012, yearly) • Netherlands (life sat and happiness, from 1974) • Poland (life sat, in 2011; freq. tbc.) • UK (life sat*, affect* and eudaimonia*, from 2011, yearly) 	<ul style="list-style-type: none"> • Australia (life sat**, from 2016; every 4 years) • Canada (life sat*, from 1985; yearly) • Israel (life sat, from 2006; life sat*, affect* and eudaimonia* from 2013, yearly) • Korea (life sat*, affect* and eudaimonia*, from 2013, yearly) • Mexico (life sat* and affect, in 2012, freq. tbc.) • New Zealand (life sat** and eudaimonia**, from 2014, every 2 years) • United States (affect and experienced eudaimonia in 2011 time-use survey; freq. tbc.) 	<ul style="list-style-type: none"> • Chile • Japan

*Questions broadly in line with OECD Guidelines; **Questions planned to be in line with OECD Guidelines

The largest data collection exercise is that of the UK ONS which, since April 2011, has included a set of four questions on the core of its Annual Population Survey (n=160,000) covering life evaluation, momentary emotional state, and worthwhileness:

- Overall, how satisfied are you with your life nowadays? [evaluative well-being]
- Overall, to what extent do you feel the things you do in your life are worthwhile? [eudaimonic well-being]
- Overall, how happy did you feel yesterday? [experienced well-being]
- Overall, how anxious did you feel yesterday? [experienced well-being]

One of the most important developments identified in Table 2 is the European Union inclusion of a well-being module as an add-on to the main EU survey of

living conditions (EU-SILC). This module includes a question on life evaluation directly comparable to the OECD primary measure and a eudaimonic question that is very close to the one in the OECD core measures. As EU-SILC covers 27 EU countries as well as Croatia, Iceland, Norway, Switzerland, and Turkey, this extends the available data to the majority of the OECD, albeit with data updated only when the well-being module is run every six years. More importantly, although the decision has not been finalised, Eurostat (the EU Statistical Agency) has indicated that it is also considering including the primary life evaluation measure in the core of EU-SILC from 2020. This will make high quality annual data on life satisfaction available for the majority of the OECD on an ongoing basis.

3.5. Strategies and priorities

The nature of the policy or research question being asked dictates the appropriate SWB construct to measure and may suggest an approach to data collection. For example, if the dimension of interest is known to be sensitive on a very short time frame and responds to daily activities and events but is somewhat stable over long periods, a cross-sectional data collection conducted every 2 years may not be useful. In such cases, a high-frequency approach (even if it involves a much smaller sample) might be most informative. Similarly, if a measure varies a great deal from individual to individual on a given day but does not react very much to exogenous events (financial shocks, changes in employment rates, etc.) and tends to wash out at high aggregate levels, it may not be a particularly insightful construct to track at national levels over time (NAS, 2013, p. 16).

The Commission on the Measurement of Economic Performance and Social Progress (Stiglitz et al., 2009) concluded that, where feasible, inclusion of SWB questions on the largest population surveys will produce useful information. However, because it would be beneficial to have information about different sets of covariates for different applications, it is unlikely that an identical module could be simply plugged into different surveys to suit the many envisioned purposes for SWB data. If harmonized modules were developed that were short enough, they could in principle be included in a range of surveys. However, for surveys with a specific orientation (e.g. understanding the conditions of retirees or the time use of individuals) it would typically be preferable to tailor questions to research objectives. For example, the CPS (in which the American Time Use Survey module resides) is designed to optimize employment measures at specific levels of geographical specificity.

This diversity in the research landscape in which SWB is relevant suggests a multidimensional approach to data collection. Large-scale population surveys—such as the four-question module in the UK Integrated Household Survey or the Gallup World Poll—make up one component of a comprehensive measurement program. Data from these surveys, typically drawn from global-yesterday measures of experienced well-being and from life-evaluation questions,

provide the large sample sizes essential for repeated cross-sectional analyses capable of identifying and tracking suffering or thriving subgroups and for research on special populations such as the unemployed for whom life expectancy is falling.

The second prong of a comprehensive measurement program is smaller or more specialized data collections. One option is to construct experiments or pilots within existing large survey programs. The advantage of targeted studies and experimental modules is that they can be tailored to address specific questions of interest to researchers and policy makers—whether about health care, social connectedness of the elderly, city planning, airport noise management, or environmental monitoring.

The third prong of an ideal data infrastructure would consist of panel studies designed to document changes in SWB over time. The inclusion of SWB measures in *Understanding Society* (formerly the *British Household Panel Survey*) and the *German Socio-economic Panel* has already contributed greatly to understanding issues relating to both causality and to adaptation over time (e.g. Lucas, 2007; Lucas et al., 2004). How individuals' experienced well-being and life satisfaction change over time and in reaction to events and life circumstances cannot be fully understood without longitudinal information, which may also help to make progress on causality questions (e.g. does getting married make people happier, or are happier people more likely to get married?). The policy relevance of monitoring SWB changes over time is clear where, for example, it is important to know the full impact on people of new legislation or on outcomes of experiments such as the Oregon Health Care Study (NAS, 2013, p. 107). Schuller et al. (2012) reviews the contribution of longitudinal data in analyzing SWB responses for a range of key well-being domains, such as relationships, health, and personal finance.

A final prong of an ideal data collection is information on experienced well-being. As described above, momentary sampling methods have been central to SWB research but largely out of practical reach for adoption by national statistical offices. However, rapid changes in technology and in the way the public exchanges information have brought the world to a point where momentary assessment techniques may now be on the horizon for national statistics. Regardless of developments in EMA, collecting experienced well-being data through a DRM approach in nationally representative time use surveys has been demonstrated to be feasible both through the *ATUS* and the *Enquete Emploi du Temps*. The UNECE Guidelines for Harmonising Time Use Surveys (UNECE, 2014) and the *OECD Guidelines on Measuring Subjective Well-being* both recommend that national statistical offices should move to collect experienced well-being data in time use surveys. Recently Statistics Canada has become the first national statistical office to move in this direction following the examples of the Bureau of Labor Statistics in the USA and INSEE in France. Precisely knowing how people are doing emotionally and what they are doing in the

moment can shed light on the effects of commuting, air pollution, child care, and a long list of areas with clear ties (NAS, 2013, p. 108).

The Mappiness project (mappiness.org.uk), designed to investigate well-being effects to the public associated with open green space in the London area, allows monitors to look at individual-level variation for people located in different outdoor environments. This project provides a clear example of the emerging methods to capture SWB in the context of EMA measures and the role of portable recording—in this case the use of cellphones and global positioning system (GPS) tracking. The British Millenium Cohort Study is considering use of geospatial cellphone responses as a post-survey supplement. There are still major unresolved data quality and representativeness issues in this world of new data and big data. For instance, the sampling properties are largely unknown for data generated by social media, phone records, Internet usage, and the like. Much more will need to be learned about distributional characteristics of various underlying subpopulations.

Social media data and other kinds of unstructured data (those, such as administrative records or company-maintained information, produced initially as a by-product of non-statistical purposes) may become increasingly useful for shedding light on trends in people's emotional states. Word mining exercises have been used to show patterns in emotional states—for example, a Facebook happiness index showed the standard weekend and holiday effects and expected changes associated with major events, such as disasters. The words people use on social media such as Twitter, Facebook, and Google search queries are a rich, if imperfect, source of information about their personality and psychological state. Additionally, analyses of data generated by social media and other Internet activities will produce insights relevant to public policy (see the discussion below of relevance to understanding social or political movements such as the Arab Spring).

4. How do different dimensions of SWB link with policies?

Informing policy—or at least the potential to do so—is a critical criterion for deciding whether it is worth the time and cost of measuring SWB in national flagship population surveys or in more focused domain-specific surveys. It is clear that different kinds of SWB measures inform different kinds of policies. For example, optimizing end-of-life care decisions may give greater weight to short-term concerns—minimizing day to day suffering—and therefore suggest a need for experience based measures. Education and employment policies may focus more on life satisfaction or even eudaimonic concerns, for which evaluative measures are highly relevant. In either case, assessment is needed about the extent to which SWB adds analytic content beyond the existing “objective statistics” such as those we have come to rely on in such research and policy areas as poverty (e.g. income data) and health (e.g. vital statistics).

The unique policy value of SWB measures may lie not in assessing how income or other variables relate to an aggregate-level tracking of emotional states or life satisfaction, but in discovering actionable relationships that might otherwise escape attention in order to better understand the full of impact of commuting patterns, accessibility of child care, exercise programs, interaction and connectedness with neighbours and friends, the presence of neighbourhood amenities and other city planning issues, divorce and child custody laws, and the like (NAS, 2013, pp. 88-89).

The intended use for measures of SWB also affects judgements about the validity of such measures. In the remainder of this section, we outline the major uses of SWB measures: (1) complementing objective measures of the economy, health, and society; (2) to better understand the drivers of well-being at the level of the individual; (3) for policy evaluation and cost benefit analyses; and (4) for identifying potential policy issues.

4.1. The role of SWB as a complement to objective economic, health, and social measures

SWB measures offer significant potential for complementing conventional economic, social, and health metrics by providing an alternative yardstick of progress that is grounded in people's experiences or evaluations. Traditional market-based measures alone cannot provide an adequate portrayal of quality of life, which suggests a need to shift some portion of the measurement focus from economic production toward people's well-being. The underlying argument is that national policies should better balance growth in market production with nonmarket dimensions of well-being that cannot be captured well by conventional measures. In particular, being grounded in peoples' experiences and judgements on multiple aspects of their life, SWB measures provide information about the net impact of changes in social and economic conditions on the perceived well-being of respondents, reflecting differences in tastes and preferences among individuals. An example of how these measures can change perceptions about progress is provided by Box 4.1, in respect of the "Arab Spring."

In addition to information on aggregate trends, SWB measures can also provide a picture of which groups in society are most (dis)satisfied or experience the best or worst lives that reflect, among other things, the impact of tastes, aspirations, and life circumstances. Migrants, for example, may be more motivated than the rest of the population by income relative to other factors (Bartram, 2010), as this is a primary motive for their decision to move abroad. This heterogeneity makes assessing overall migrant well-being compared to the rest of the population challenging. However, because SWB measures incorporate the impact of different weights that people attach to aspects of their quality of life, they have the potential to add an important dimension to analyses in situations involving comparisons between population groups.

Box 4.1. Subjective well-being, GDP growth and the "Arab Spring"

For policy-makers, measures of SWB are valuable as indicators of progress when they can alert them to issues that other social and economic indicators might fail to identify. One recent example where measures of SWB demonstrated their ability to capture important elements of well-being not captured by more traditional measures was the decline in country-average measures of SWB that occurred in Egypt and Tunisia in the years leading up to 2011, a decline that contrasts with the much more favourable evolution of GDP data. For example, Tunisian real GDP per capita increased from USD 8,891 in 2008 to USD 9,489 in 2010, a real gain of around 7%. However, the proportion of the population indicating a high level of satisfaction with their life as a whole fell from 24% to 14% over the same period (Gallup, 2011). Egypt showed a similar pattern from 2005 to 2010, with a real gain in GDP per capita of around 34% and a decline in the share of respondents classified as "thriving" by almost half. This illustrates how subjective perceptions can provide information on very significant outcomes in societies that other conventional indicators such as GDP growth do not provide.

An additional use of SWB measures is for monitoring progress in aggregate cross-country comparisons, such as those included in *How's Life?* (OECD, 2011). Because controlled experiments are typically impossible, cross-country comparisons of SWB outcomes are one way to learn about the strengths and weaknesses of different policies. When SWB measures are sensitive to a different range of drivers than are other social and economic indicators, they provide additional information about the consequences of a particular policy. A crucial issue in using SWB in this way, however, is the degree to which cross-cultural comparisons of such measures are valid.

Interest by the general public and the media in using measures of SWB as complements of measures of progress represents another valid rationale for public data collection. Of particular interest to these users is the question of whether things are getting better or worse overall, and for whom. As in the policy realm, SWB measures used for general public information purposes should be viewed as one set in the much broader array of indicators through which populations are monitored and insights about societal progress or deterioration are drawn.

4.2. The role of SWB in better understanding the drivers of people's well-being

A second major use of SWB measures is to contribute to a better understanding of the drivers of well-being at an individual level. If it can be established that SWB measures accurately capture the concepts that they claim to – an overall evaluation of life or the experienced moods and emotions of an individual over a period of time – they can be used to provide information about the relative contribution of different factors and circumstances to a person's well-being. The quality of the information will be tempered by measurement error and by the fact that a person's subjective perception of their well-being is not necessarily quite the same thing as their overall well-being (see Dolan, Peasgood,

and White, 2008; Helliwell and Wang, 2011; Boarini, Comola, Smith, Manchin, and De Keulenaer, 2012).

Subjective measures can be used to test specific hypotheses about what aspects of policy are most important to people. Halpern (2010), for example, refers to an instance in which the Merseyside police, in the United Kingdom, used data on how satisfied members of the public were with the service provided by the local force, alongside more traditional performance measures on crimes committed and offence resolutions. In contrast to the expected hypothesis – which was that minimising the response time from the police was of crucial importance for public satisfaction – the evidence showed that it was much more important that police arrived when they said they would. For minor issues not involving safety, what mattered was the punctuality rather than the speed of the response.

Going beyond simply identifying what matters to people, SWB measures can provide the basis for developing a better understanding of trade-offs when policy options involve comparisons of fundamentally different types of outcome (see box 4.2 below). Dolan and White (2007) note that this issue characterises many attempts to encourage “joined-up government,” where costs and benefits of a particular intervention must be considered not just based on the outcome of concern to one agency, but also in terms of how choices affect the outcomes of other agencies.

Measures of SWB can potentially capture the combined effect on an individual’s perception of their well-being of a range of different changes in life circumstances. For example, Ferrer-i-Carbonell and Frijters (2004) compare the magnitude of the impact of health satisfaction versus housing satisfaction on overall life satisfaction.¹² Similarly, Di Tella, MacCulloch, and Oswald (2003) investigate inflation, unemployment trade-off in terms of the effect on life satisfaction. While the so-called “misery index” weights the unemployment rate and inflation rate equally as indicators of the negative impact of macro-economic outcomes, Oswald and Macculloch’s analysis suggests that the impact of unemployment on SWB is significantly greater than that of inflation.

Box 4.2. Using measures of subjective well-being to value life events

Measures of SWB provide a relatively straight-forward way of comparing the relative impact of fundamentally different life events in a quantitative way and, based on this, assigning such events a monetary value. Clark and Oswald (2002) present a method for valuing life events and, although the literature on using measures of SWB to value life events has expanded significantly since 2002, the basic methodology remains largely unchanged. Consider the results below from Boarini et al. (2012). The coefficients for the (base two) logarithm of household income, being married, and being unemployed are

¹² Consideration of initial sample variance in each measure is important here: if the sample has uniformly high levels of health satisfaction, but variable levels of housing satisfaction, housing satisfaction may look more important in a regression analysis, simply because it has more variation to associate with variation in the outcome measure.

shown, and express the change in life satisfaction (on a scale of 0 to 10) associated with a doubling of income, being married, or being unemployed, respectively, holding all else constant.

Event	Coefficient
Log Household Income	0.1482
Married	0.2584
Unemployed	-0.4643

Using these coefficients, it is possible to calculate the relative impact of being married compared to being unemployed on life satisfaction as $0.2584 / 0.4643 = 0.5565$. Or, put more simply, being unemployed has almost twice the impact on life satisfaction as does being married.

Going beyond this, the monetary value of being married or being unemployed can be calculated by comparing the relevant coefficients with that associated with the coefficient for household income. Using the values presented above, the coefficient on being married is $0.2584 / 0.1482 = 1.7435$ times larger than the impact of a doubling of household income. For a person with a household income equal to the average OECD per capita household disposable income (\$17 286 at PPP, 2008), this is equivalent to $1.7435 \times \$17\,286 = \$30\,138$. For unemployment the comparable value is $2.930 \times \$17\,286 = \$50\,647$.

These values are intended to illustrate the techniques involved, and need to be treated with caution. In particular, it would be preferable to use panel data which might better capture a causal relationship (as do Clark and Oswald) rather than just correlation; potential biases in the data as well as appropriate model specification also must be evaluated (Fujiwara and Campbell, 2011).

4.3. The role of SWB in policy evaluation and cost benefit analyses

A third use of SWB measures is to assist in the evaluation of policies. This includes both the direct use of measures of SWB in formal policy evaluations as well as the more indirect – but possibly more important – role that they can play in cost-benefit analysis. For some initiatives – where the impact on subjective experiences of the population is the main object of the program – measures of SWB may even be suitable as the primary metric for assessing its success.

Many policy evaluations already include subjective measures of client satisfaction that gauge respondents' perceptions about what elements of a program are most valuable. More general measures of overall SWB, however, have some significant advantages over and above these more focused measures. Most importantly, measures of SWB provide information about the impact of an initiative on the respondent's SWB, rather than the impact that the respondent consciously identifies. These values can differ because people's judgements about the impact of a program may be influenced by their participation (i.e., they might be more prone to assign the cause of any recent changes in their well-being to the program rather than to other factors, knowing that this is what he/she is being asked about). Also, people may not be aware of all of the various feedback loops via which a policy programme affects them. For example, in evaluating an active

employment program, respondents might consider the direct effect on their well-being of both having a job and gaining additional income, but not the flow on well-being that would stem from changes in their time-use due to longer commuting. Because measures of SWB can capture the overall impact of a change on life circumstances, without requiring a cognitive judgement by the respondent on which causal pathways are being asked about, such measures provide useful additional information on the overall impact of a programme.

In some cases, measures of SWB can be better than conventional cost-benefit analysis at treating non-monetary outcomes. Examining the relative costs and benefits of a proposal is relatively straight-forward when the proposal is aimed at strictly economic outcomes, and the costs and benefits of the proposal can be obtained from the relevant market prices. However, where the aim of a proposal is to achieve outcomes that do not have an obvious market price, it is much more challenging to obtain meaningful values for analysing the relevant costs and benefits. Because much government policy is concerned with market failures, many government policies are correspondingly focused on achieving non-market outcomes.

The traditional economic approaches to cost-benefit analysis for non-market outcomes depend on either revealed preference or contingent valuation techniques to estimate “prices” for such outcomes. A revealed preference approach involves calculating values based on the shadow prices implied by observed behaviour, while contingent valuation techniques calculate values based on the “willingness to pay” for the outcome in question, as expressed by respondents to a hypothetical question in a survey. Clark and Oswald (2002) note that measures of SWB can provide the framework for such valuations by comparing the impact of a particular outcome on SWB with the impact of a change in income on SWB. By making such a comparison, it is possible to calculate the amount of money required to achieve the same increase or decrease in well-being as that caused by the outcome under assessment.

There is good reason to believe that, in some circumstances, measures of SWB have advantages over both revealed preference and contingent valuation for the purposes of cost-benefit analysis (see box 4.3 below). An obvious advantage is that many measures of SWB – such as overall life satisfaction – are relatively easy and cheap to collect. However, there are also more substantive methodological advantages that may be associated with using measures of SWB in this way. Revealed preference relies on strong assumptions about people’s ability to know how an outcome will affect them in the future, and on the assumptions that markets are in equilibrium. Diener, Lucas, Schimmack, and Helliwell (2009) note that for market prices for houses to reflect the disutility of airport noise accurately would require that house purchasers are able to forecast how much the noise will impact them before buying the house. Similarly, in this example, it is difficult to disentangle the differences in house prices due to noise from differences in other aspects of house quality.

Box 4.3. *The Green Book* and life satisfaction

The Green Book is the formal guidance from the Treasury of the United Kingdom to other UK government agencies on how to appraise and evaluate policy proposals. The current edition of *The Green Book* dates to 2003, and provides advice on how officials should provide justification for a proposed government intervention, set objectives for the proposal, appraise the various options, and evaluate the effectiveness of the final action that results. In July 2011, *The Green Book* was updated to reflect the results of a review of valuation techniques for social cost-benefit analysis jointly commissioned by the Treasury and the Department for Work and Pensions (Fujiwara and Campbell, 2011). The review specifically focuses on the contribution that can be played by measures of SWB – particularly life satisfaction – alongside more traditional approaches to cost-benefit analysis. In summarising the conclusions of the review, *The Green Book* states (p. 58):

A newer, “subjective well-being approach” has been gaining currency in recent years. The “life satisfaction approach” looks at people’s reported life satisfaction in surveys such as the ONS’s Integrated Household Survey, which began including questions on respondents’ subjective well-being in April 2011. The life satisfaction approach uses econometrics to estimate the life satisfaction provided by certain non-market goods, and converts this into a monetary figure by combining it with an estimate of the effect of income on life satisfaction.

At the moment, subjective well-being measurement remains an evolving methodology and existing valuations are not sufficiently accepted as robust enough for direct use in Social Cost-benefit Analysis. The technique is under development, however, and may soon be developed to the point where it can provide a reliable and accepted complement to the market based approaches outlined above. In the meantime, the technique will be important in ensuring that the full range of impacts of proposed policies are considered, and may provide added information about the relative value of non-market goods compared with each other, if not yet with market goods.

While the amendment to *The Green Book* stops short of fully endorsing the use of life satisfaction measures for use in formally evaluating government programmes, the decision to make an interim amendment in itself signals strongly the importance that UK central agencies attach to obtaining improved measures of the value of non-market outcomes.

Contingent valuation also relies strongly on people’s ability to make accurate judgements about how something will make them feel in the future. Dolan and Peasgood (2006) observe that people have difficulty imagining how good or bad different circumstances are actually going to be. Indeed, the “willingness to pay” surveys commonly used for contingent valuation are, to a large degree, measures of the SWB associated with a hypothetical scenario. Using measures of SWB to calculate the costs based on the actual impact of different life circumstances on SWB removes the hypothetical element from the equation. In addition, contingent valuation surveys tend to produce very different estimates of the value of outcomes for people at different points on the income distribution. This tends to result in either weighing the desires of the rich more heavily than the poor when assessing the costs and benefits associated with the proposal under consideration

or taking account of the marginal utility of income in calculating the final cost. The latter approach is difficult in the absence of robust estimates of the marginal utility of income (Dolan and White, 2007).

4.4. The role of SWB in identifying potential policy issues

An important feature of SWB measures is their ability to provide insights into human behaviour and decision-making. In particular, measures of SWB can help researchers better understand the difference between the *ex ante* beliefs that people hold about their future well-being (which form the basis for decisions) and the *ex post* outcomes that people achieve in terms of their SWB. A better understanding of these issues is important both for policy-makers and for the broader public. Policy-makers have an interest in understanding why people make the decisions that they do, because much public policy involves dealing with the consequences of systematic poor decision-making by individuals. Similarly, businesses and the general public have an interest in understanding how people's SWB shapes their behaviour.

One example of how subjective measures are useful to businesses and the broader public is the information they provide about the characteristics of good places to live and work and in turn how that predicts future behaviour. Clark (2001) has shown that measured job satisfaction predicts the probability of an employee leaving their job. Thus businesses might well have an interest in the measured job satisfaction of their employees and in understanding the determinants of job satisfaction.

Measures of SWB can also help shed light on various biases in the way people make decisions. Although people are generally able to predict whether events are likely to be pleasant or unpleasant, Wilson, Gilbert, and colleagues have described ways in which affective forecasting can be biased or faulty, particularly with regard to the intensity and duration of emotional reactions to future events (e.g. Wilson, Wheatley, Meyers, Gilbert, and Axsom, 2000; Wilson and Gilbert, 2006). Kahneman et al. (2006) show that people are prone to over-estimate the impact of income gains on their life satisfaction relative to other factors. Commuting, for example, has been found to have a strong negative impact on both measures of affect (Kahneman et al., 2006) and life evaluations (Frey and Stutzer, 2008). This suggests that people may be prone to over-estimating the positive impact of, for example, a new job with a higher salary but a longer commute.

There are also direct policy applications for better understanding the human decision-making process and the various biases and heuristics involved in it. Consider the case of policy options that incorporate a "default" option – for example, workplace retirement schemes that are set up on a basis of either "opt in" clauses, where a new employee does not join the scheme unless he/she ticks a box to join, or "opt out" clauses, where the reverse is the case. The fact that people respond differently depending on which default is selected – despite the fact that in neither case is there any compulsion – has raised policy interest in the

idea of “libertarian paternalism”, which focuses on achieving better outcomes by setting policy defaults to influence people’s behaviour in positive directions. Dolan and White (2007) note that information on SWB can be used to help set policy default options more optimally, by indicating which default options contribute most to SWB.

While a full accounting of SWB applications to research and policy is beyond scope here, the following examples hint at their diversity and potential (NAS, 2013, p. 89):

- Kahneman and Deaton (2010) and Stevenson and Wolfers (2013) used data collected in the Gallup-Healthways Well-Being Index to estimate the impact of income and income-normalized effects on life satisfaction and experienced well-being. Understanding the relationship could prove useful for informing tax and social program policies.
- Oswald and Wu (2009) used data from the Behavioral Risk Factor Surveillance System to rank the US states based on hedonic analyses of regional variation in such factors as precipitation, temperature, sunshine, environmental greenness, commuting time, air quality, and local taxes; all suggesting a role for SWB data in assessing regional and city policies.
- Diener and Chan (2010) argue that people’s emotional states causally affect their health and longevity, concluding that the data are compelling, though “not beyond a reasonable doubt” (NAS, 2013, pp. 87-88).
- Robert Sampson’s Chicago neighborhoods study (Sampson and Graif, 2009) reveals the importance of connectedness to the well-being of neighbourhoods. One of many examples is the variation, even among relatively poor areas, in the resilience of different neighbourhoods to the 1994 heat wave in the city. Sampson’s findings suggest the value of data on people’s trust in neighbours, interactions, connectedness, as well as mechanisms whereby the built environment can promote SWB (though there is the alternative hypothesis being that happier people tend to have more autonomy over where they choose to live).
- Krueger and Mueller (2011) found that the SWB of the unemployed declines with the duration of unemployment spells; they also found that the time spent involved in job search is particularly unhappy and the unhappiness increases with the time spent in job search (measured both with life-satisfaction and sadness variables). These effects on the unemployed provide an example of how low experienced well-being related to the process could in the end undermine individuals’ incentives to persist, ultimately reducing their capacity to achieve higher levels of evaluative well-being in the future.

From company policies that improve well-being – and possibly, in turn, improve productivity and lower absenteeism – to community or regional planning policies, SWB measures would appear most valuable when costs and benefits must be weighed in the absence of market or easily quantifiable elements.

Government consideration of spending to redirect an airport flight path to reduce noise pollution, funding alternative medical care treatments when more is at stake than maximizing life expectancy, or selecting between alternative recreational and other uses of environmental resources are possible examples.

Across these policy applications, experienced and evaluative dimensions of well-being may have very different implications (Diener, 2011; Graham, 2011; Kahneman et al., 2006). For example, actions aimed at enhancing longer-term opportunities may actually impart negative short term effects on daily experience. A policy designed to enhance living quality at the end of life, for example, focuses on the hedonic dimension (which is at least one of the objectives of palliative care, that is, relieving suffering), while a policy aimed at increasing educational opportunities of youth focuses on life evaluation (NAS, 2013, p. 91). Optimization of short-term versus long-term well-being (both at individual and aggregated levels) may imply different policy actions. A program to reduce fat intake or smoking may reduce experienced well-being in the short run but increase it (via the health covariate) over the long run.

The distinction between positive and negative affect and between suffering and happiness are also important with, arguably, minimization of the negative being more relevant to public policy.¹³ The U.S. General Social Survey (GSS) registers exposure to negative circumstances and events experienced by people (e.g. hospitalization, death of a family member, eviction, crime victimization), and was designed to report “objective experiences that disrupt or threaten to disrupt an individual’s usual activities, causing a substantial readjustment in that person’s behavior” (Thoits, 1983). As described by Smith (2005), this approach has been used extensively not only to account for differing levels of reported well-being among individuals or groups but also for understanding and predicting individual illness (both psychological and physiological); in so doing, it provides “factual data for the formulation of public policies to deal with these problems” (NAS, 2013, p. 38).

Self-reports of SWB are likely to add useful information in instances where medical interventions have a desired outcome that is something other than merely an increase in life expectancy, where reflections of successful treatment and support extend beyond signs and symptoms and into domains such as functioning and social integration, and where parties other than the patients are affected by treatment and symptoms (care givers, family members, and others). See the article by Richard Frank article in this volume.

Future directions

Thinking in terms of a harmonized approach for national statistics offices to follow, the OECD *Guidelines* mark an important step forward in the measurement

¹³ Dolan and Metcalfe (2011) surveyed people to ask whether government policy should seek to (1) improve happiness or (2) reduce misery, and there was more support for the second option.

of subjective well-being, but do not provide the ‘final word’ on the subject. Although some aspects of the measurement of subjective well-being – such as questions about overall satisfaction with life – are well understood, other potentially important measures currently draw on weaker evidence. It is expected that the evidence base will continue to develop rapidly over the next few years. In particular, to the extent that national statistical offices start regularly collecting and publishing data on SWB that researchers can exploit, many methodological questions are likely to be resolved, and an increasing body of knowledge will accumulate on the policy uses of these data.

National statistics offices face two issues in particular: (1) the need to pursue experimental techniques to push the state of the art forward; and (2) the need to collect high quality covariate data alongside SWB measures. Regarding the first, national statistics offices have long histories developing survey methods through systematic experiments and so are well positioned to contribute to the evolution of SWB measurement. While the OECD *Guidelines* were being drafted, the UK Office for National Statistics (ONS) was in the process of developing and collecting its first official measures of SWB. Typically, national statistical offices invest considerable methodological research upfront before collecting data for a new measure, but then implement collection in a homogenous way. In developing their measures of subjective well-being, the ONS deviated from this process significantly. Although the ONS did invest in methodological work before proceeding to measurement, rather than standardise on a single measure immediately, an experimental approach was taken by splitting the sample in their Household Opinion Survey and using this to test different questions, question order, and other methodological points. The experimental approach adopted by the ONS has had an important impact with respect to knowledge of the validity and reliability of subjective well-being measures and best practice with respect to question design.

On the second point, part of the experimentation process involves figuring out which subject matter domains (e.g. health, time-use, environment, city planning) benefit most from adding SWB content to existing surveys. A key advantage of many surveys carried out by government statistical agencies is that they generally collect higher quality information on potential covariates – such as income, labour force status, or education – than is possible in smaller unofficial surveys. Because the potential for insightful inferences to be drawn from SWB analyses depends not only on the quality of the SWB measure, but also on the quality of data on a range of other factors, surveys from national statistical agencies offer an opportunity to advance the field in a way that may not be possible elsewhere.

REFERENCES

- BARTRAM, D., (2010). Economic migration and happiness: Comparing immigrants' and natives. Happiness gains from income. *Social Indicators Research* 103(1): 57–76.
- BENJAMIN, D., KIMBALL, M., HEFFETZ, O., REES-JONES, A., (2013). Can Marginal Rates of Substitution be Inferred from Happiness Data? Evidence from Residency Choices, NBER Working Paper 18927.
- BOARINI, R., COMOLA, M., SMITH, C., MANCHIN, R., DE KEULENAER, F., (2012). What Makes for a Better Life? The determinants of subjective well-being in OECD countries: Evidence from the Gallup World Poll. STD/DOC(2012) 3, OECD.
- BRADBURN, N., (1969). *The Structure of Psychological Well-being*. Chicago: Aldine.
- CHRISTODOULOU, C., SCHNEIDER, S., STONE, A. A., (2013). Validation of a brief yesterday measure of hedonic well-being and daily activities: Comparison with the Day Reconstruction Method. *Social Indicators Research*.
- CLARK, A. E., (2001). What really matters in a job? Hedonic measurement using quit data. *Labour Economics* 8: 223–242.
- CLARK, A. E., OSWALD, A. J., (2002). A simple statistical method for measuring how life events affect happiness. *International Journal of Epidemiology* 31: 1139–1144.
- CLARK, A. E., SENIK, C., (2011). Is Happiness Different from Flourishing? Cross-country Evidence from the ESS. Working Paper 2011-04, Paris, School of Economics.
- CLARK, A. E., FRIJTERS, P., SHIELDS, M., (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature* 46(1): 95–144.
- DEATON, A., (2010). Income, aging, health, and well-being around the world: Evidence from the Gallup World Poll. In *Research Findings in the Economics of Aging*, D. A. Wise (ed.), pp. 235–263. Chicago: University of Chicago Press.
- DEATON, A., (2012). The financial crisis and the well-being of Americans. *Oxford Economic Papers* 64(1): 1–26.
- DECI, E., RYAN, R., (2006). Hedonia, eudaimonia, and well-being: An introduction. *Journal of Happiness Studies* 9: 1–11.

- DI TELLA, R., MACCULLOCH, R., OSWALD, A. J., (2001). Preferences over inflation and unemployment: Evidence from surveys of happiness. *The American Economic Review* 91(1): 335–341.
- DI TELLA, R., MACCULLOCH, R., OSWALD, A. J., (2003). The Macroeconomics of Happiness. *The Review of Economics and Statistics* 85(4): 809–827.
- DIENER, E., (2006). Guidelines for national indicators of subjective well-being and ill-being. *Applied Research in Quality of Life* 1(2): 151–157.
- DIENER, E., (2011). The Validity of Life Satisfaction Measures. Unpublished manuscript, Department of Psychology, University of Illinois at Urbana-Champaign.
- DIENER, E., CHAN, M., (2010). Happier People Live Longer: Subjective Well-Being Contributes to Health and Longevity. Unpublished manuscript, Department of Psychology, University of Illinois at Urbana-Champaign.
- DIENER, E., PAVOT, W., (1993). Review of the satisfaction with life scale. *Psychological Assessment* 52(2): 164–172.
- DIENER, E., HELLIWELL, J. F., KAHNEMAN, D. (eds.), (2010). *International Differences in Well-Being*. Oxford University Press.
- DIENER, E., LUCAS, R. E., NAPA SCOLLON, C., (2006). Beyond the hedonic treadmill: Revising the adaptation theory of well-being. *American Psychologist* 61(4): 305–314.
- DIENER, E., LUCAS, R. E., SCHIMMACK, U., HELLIWELL, J. (eds.), (2009). *Well-Being for Public Policy*. Oxford University Press.
- DIENER, E., SUH, E. M., LUCAS, R. E., SMITH, H. L., (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin* 125(2): 276–302.
- DIENER, E., KAHNEMAN, D., TOV, W., ARORA, R., HARTER, J., (2009). Income's differential influence on judgments of life versus affective well-being. In *Assessing Well-Being*, E. Diener (ed.), pp. 233–246. Oxford, UK: Springer.
- DOLAN, P., KAVETSOS, G., (2012). *Happy Talk: Mode of Administration Effects on Subjective Well-Being* (CEP discussion paper, no. 1159). London, UK: Centre for Economic Performance, London School of Economics and Political Science.
- DOLAN, P., METCALFE, R., (2011). *Comparing Measures of Subjective Well-Being and Views About the Role They Should Play in Policy*. London, UK: Office for National Statistics.

- DOLAN, P., PEASGOOD, T., (2006). Valuing non-market goods: Does subjective well-being offer a viable alternative to contingent valuation. Imperial College Working Paper. Imperial College, London.
- DOLAN, P., WHITE, M., (2007). How can measures of subjective well-being be used to inform policy? *Perspectives on Psychological Science*. 2(1): 71–85.
- DOLAN, P., LAYARD, R., METCALFE, R., (2011). *Measuring Subjective Well-being for Public Policy*. Office for National Statistics, UK.
- DOLAN, P., PEASGOOD, T., WHITE, M., (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology* 29: 94–122.
- EASTERLIN, R., (1974). Does economic growth improve the human lot? Some empirical evidence. In *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz*. P.A. David, and M.W. Reder (eds.), pp. 89–125. New York, Academic Press Inc.
- FERRER-I-CARBONELL, A., FRIJTERS, P., (2004). How important is methodology for estimates of the determinants of happiness? *The Economic Journal* 114: 641–659.
- FREY, B. S., STUTZER, A., (2002). What can economists learn from happiness research? *Journal of Economic Literature* 40(2): 402–435.
- FREY, B. S., STUTZER, A., (2008). Stress that doesn't pay: The commuting paradox. *Scandinavian Journal of Economics* 110(2): 339–366.
- FRIJTERS, P., (2000). Do individuals try to maximize general satisfaction? *Journal of Economic Psychology* 21: 281–304.
- FUJIWARA, D., CAMPBELL, R., (2011, July). *Valuation techniques for social cost-benefit analysis: Stated preference, revealed preference, and subjective well-being approaches, a discussion of the current issues*. United Kingdom: HM Treasury and Department for Work and Pensions.
- GERE, J., SCHIMMACK, U., (2011). A multi-occasion multi-rater model of affective dispositions and affective well-being. *Journal of Happiness Studies* 12(6): 931–945.
- GRAHAM, C., (2011). *The Pursuit of Happiness: An Economy of Well-Being*. Washington, DC: The Brookings Institution Press.
- HALPERN, D., (2010). *The Hidden Wealth of Nations*. Cambridge: Polity Press.
- HELLIWELL, J. F., BARRINGTON-LEIGH, C. P., (2010). *Measuring and Understanding Subjective Well-being*. NBER Working Paper, No. 15887, National Bureau of Economic Research.

- HELLIWELL, J. F., WANG, S., (2011). Trust and well-being. *International Journal of Well-being* 1(1): 42–78.
- HELLIWELL, J. F., LAYARD, R., SACHS, J. D., (2013). *The World happiness Report*. The Earth Institute, Columbia University.
- HUPPERT, F. A., SO, T. T. C., (2009). What Percentage of People in Europe are Flourishing and What Characterises Them? Well-Being Institute, University of Cambridge, mimeo prepared for the OECD/ISQOLS meeting on Measuring subjective well-being: An opportunity for NSOs? Florence, 23/24 July.
- HUPPERT, F.A., MARKS, N., CLARK, A., SIEGRIST, J., STUTZER, A., VITTERSO, J., WAHRENDORF, W., (2009). Measuring well-being across Europe: Description of the ESS well-being module and preliminary findings. *Social Indicators Research* 91: 301–315.
- KAHNEMAN, D., DEATON, A., (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences of the United States of America* 107(38): 16489–16493.
- KAHNEMAN, D., KRUEGER, A. B., (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives* 20(1): 3–24.
- KAHNEMAN, D., DIENER, E., SCHWARZ, N., (1999). *Well-being: The Foundations of Hedonic Psychology*. Russell Sage Foundation, New York.
- KAHNEMAN, D., KRUEGER, A. B., SCHKADE, D., SCHWARZ, N., STONE, A. A., (2004). A survey method for characterizing daily life experience: The Day Reconstruction Method (DRM). *Science* 306(5702): 1776–1780.
- KAHNEMAN, D., KRUEGER, A. B., SCHKADE, D., SCHWARZ, N., STONE, A. A., (2006). Would you be happier if you were richer? A focusing illusion. *Science* 312(5782): 1908–1910.
- KRUEGER, A.B., MUELLER, A., (2012). Time use, emotional well-being, and unemployment: Evidence from longitudinal data. *American Economic Review* 102(3): 594–599.
- KRUEGER, A. B., SCHKADE, D., (2008). The reliability of subjective well-being measures. *Journal of Public Economics* 92(8–9): 1833–1845.
- KRUEGER, A. B., STONE, A. A., (2008). Assessment of pain: A community-based diary survey in the USA. *Lancet* 371(9623): 1519–1525.
- KRUEGER, A.B., KAHNEMAN, D., SCHKADE, D., SCHWARZ, N., STONE, A. A., (2009). National time accounting: The currency of life. In *Measuring the Subjective Well-Being of Nations: National Accounts of Time Use and Well-Being*, A. B. Krueger (ed.), pp. 9–86. Chicago: Chicago University Press.

- LUCAS, R., (2007). “Long-Term Disability Is Associated With Lasting Changes in Subjective Well-Being: Evidence From Two nationally Representative Longitudinal Studies.” *Journal of Personality and Social Psychology*, Vol. 92, (4), pp. 717–730.
- LUCAS, R. A., CLARK, Y., GEORGELLIS, E., DIENER, (2004). “Unemployment alters the set point for life satisfaction”, *Psychological Science*, 15, pp. 8–13.
- NATIONAL RESEARCH COUNCIL, (2013). *Subjective Well-Being: Measuring Happiness, Suffering, and Other Dimensions of Experience*. Panel on Measuring Subjective Well-Being in a Policy-Relevant Framework. A.A. Stone and C. Mackie, Editors. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- NEW ECONOMICS FOUNDATION, (2009). *National Accounts of Well-being*.
- OECD, (2011). *How’s Life? Measuring Well-Being*. Paris, OECD Publishing.
- OECD, (2013). *OECD Guidelines on Measuring Subjective Well-being*. Paris: OECD. Available:
<http://www.oecd.org/statistics/Guidelines%20on%20Measuring%20Subjective%20Well-being.pdf>.
- ONS, (2011). *Initial Investigation into Subjective Well-being from the Opinions Survey*.
- OSWALD, A. J., WU, S., (2009). *Well-Being Across America: Evidence from a Random Sample of One Million U.S. Citizens*. Unpublished manuscript, University of Warwick, UK. Presented at the IZA Prize Conference, October 22, Washington, DC.
- PAVOT, W., DIENER, E., COLVIN, C.R., SANDVIK, E., (1991). Further validation of the satisfaction with life scale: Evidence for the cross-method convergence of well-being measures. *Journal of Personality Assessment* 57(1): 149–161.
- RAYO, L., BECKER, G. S., (2007). Evolutionary efficiency and happiness. *Journal of Political Economy* 115(2): 302–337.
- ROBINSON, M. D., CLORE, G. L., (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin* 128(6): 934–960.
- SAMPSON, R.J., GRAIF, C., (2009). Neighborhood social capital as differential social organization: Resident and leadership dimensions. *American Behavioral Scientist* 52(11): 1579–1605.

- SCHULLER, T., WADSWORTH, M., BYNNER, J., GOLDSTEIN, H., (2012). The Measurement of Well-being: The Contribution of Longitudinal Studies. Report prepared for the Office for National Statistics. London. UK: Longview. Available: <http://www.longviewuk.com/pages/documents/Longviewwellbeingreport.pdf>.
- SCHWARZ, N., STRACK, F., (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In *Well-Being: The Foundations of Hedonic Psychology*, D. Kahneman, E. Diener, and N. Schwarz (eds.), pp. 61–84. New York: Russell-Sage.
- SMITH, T. W., (2005). *Troubles in America: A Study of Negative Life Events Across Time and Sub-Groups* (Russell Sage Foundation Working Paper Series).
- STEVENSON, B., WOLFERS, J., (2013). Subjective well-being and income: Is there any evidence of satiation? *American Economic Review* 103(3): 598–604.
- STEWART-BROWN, S., JANMOHAMED, K., (2008). *Warwick-Edinburgh Mental Well-being Scale (WEMWBS). User Guide Version 1*. Warwick and Edinburgh: University of Warwick and NHS Health Scotland.
- STIGLITZ, J., SEN, A., FITOUSSI, J. P., (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. Available: http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.
- TENNANT, R., HILLER, L., FISHWICK, R., PLATT, S., JOSEPH, S., WEICH, S., PARKINSON, J., SECKER, J., STEWART-BROWN, S., (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes* 5(63).
- THOITS, P., (1983). Multiple identities and psychological well-being. *American Sociological Review* 48(2): 174–182.
- THOMPSON, S., MARKS, N., (2008). *Measuring Well-being in Policy: Issues and Applications*. New Economics Foundation.
- WHITE, M. P., DOLAN, P., (2009). Accounting for the richness of daily activities. *Psychological Science* 20(8): 1000–1008.
- WILSON, T. D., GILBERT, D. T., (2006). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science* 14(3): 131–134.
- WILSON, T. D., WHEATLEY, T., MEYERS, J. M., GILBERT, D. T., AXSOM, D., (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology* 78(5): 821–836.

The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 373–396

THE OFFICE FOR NATIONAL STATISTICS EXPERIENCE OF COLLECTING AND MEASURING SUBJECTIVE WELL-BEING

Lucy Tinkler¹

ABSTRACT

The UK Office for National Statistics (ONS) started measuring subjective well-being in 2011 as part of the ONS Measuring National Well-being programme. The aim of the Measuring National Well-being programme is to measure the quality of life and progress of the UK. This article explores the development of the ONS subjective well-being measures, data collection methods, data presentational considerations, overview of findings, and latest developments. It discusses the way in which user engagement has been key to the development of the ONS subjective well-being statistics.

Key words: subjective well-being, the Office for National Statistics, evaluative, eudemonic, experience, affect, user engagement.

1. Introduction

The UK Measuring National Well-being programme was launched in 2010 with the aim of developing measures of national well-being and progress of the country. Through this programme, the UK government demonstrated its commitment to developing better measurement of quality of life, and to develop policies based on what matters most to people. At the commencement of this initiative, the Office for National Statistics (ONS) undertook a national debate on ‘what matters to you?’ between 26 November 2010 and 15 April 2011. ONS received over 7,900 responses in the form of completed questionnaires, including more than 50 responses from organisations. As part of the debate, ONS also established a national well-being website. During the debate this site generated almost 17,700 visits and comments from 1,200 people (ONS 2011a). Consultation with users has been a key part in the development of measures of subjective well-being as well as the Measuring National Well-being programme as a whole.

As part of the Measuring National Well-being programme the National Statistician announced that the (ONS) were planning to collect data on ‘subjective

¹ Office for National Statistics. E-mail: lucy.tinkler@ons.gsi.gov.uk.

well-being' i.e. asking individuals to provide their own assessment of their well-being (ONS, 2011a).

This development occurred at a time when there was increasing international emphasis towards encompassing subjective well-being measures in official Statistical systems. The report of the Commission for the Measurement of Economic Performance and Social Progress (CMEPSP) stated that:

"It is possible to collect meaningful and reliable data on subjective as well as objective well-being. Subjective well-being encompasses different aspects (cognitive evaluations of one's life, happiness, satisfaction, positive emotions such as joy and pride, and negative emotions such as pain and worry)... [subjective well-being] should be included in larger-scale surveys undertaken by official statistical offices" (CMEPSP, 2009).

Subjective well-being can be defined as an approach which allows the individual to decide what is important when making an assessment about how they think and feel about their lives (Hicks, Tinkler, Allin, 2013).

The collection of subjective well-being data is now established within the UK ONS Measuring National Well-being programme as one of the key components of National Well-being, alongside other measures of society, economy, the environment, and the sustainability of well-being into the future (ONS, 2014d).

2. The ONS approach to measuring subjective well-being

Before selecting subjective well-being questions ONS identified the three broad approaches associated with the measurement of subjective well-being; 'evaluative', 'experience', and 'eudemonic'.

The evaluative approach requires respondents to make an appraisal or cognitive reflection of their life (Diener, 1994). Respondents can be asked to provide an assessment of their overall life satisfaction or certain aspects of their life such as satisfaction with their health, job, or relationships. An alternative evaluation question is known as the Cantril ladder of life in which respondents rate their current life on a ladder scale for which 0 is 'the worst possible life for you' and 10 is 'the best possible life for you'. Other measures include general happiness measures that are not specific to a particular point in time. The evaluation approach to measuring well-being has been the most prevalent both in national and international surveys. These type of questions have also been seen by policy makers as useful sources of information for some time (Donovan and Halpern, 2002).

Experience (or affect) measures aim to provide an assessment of the emotional quality of an individual's experience in terms of the frequency, intensity and type of affect or emotion at any given moment, for example, happiness, sadness, anxiety or excitement. This can be collected via diary based methods such as the Day Reconstruction Method (DRM), and the Experience Sampling Method (ESM), where respondents report feelings at different times of the day while carrying out different activities. It is also possible for this

information to be collected via more general social survey questions through asking respondents questions about their feelings over a short reference period, for example, ‘Overall, how anxious did you feel yesterday?’ (Dolan, 2010; Hicks, 2011). Experience measures can pick up both positive emotions, such as happiness, joy or contentment, and negative ones, such as anxiety, worry, pain, or anger (Tinkler and Hicks, 2011).

The eudemonic approach is based on the theory that people have underlying psychological needs for their lives to have meaning, to have a sense of control over their lives and to have connections with other people (Ryff, 1989). This approach to subjective well-being is also sometimes described as the ‘functioning’ or ‘psychological’ approach to well-being. Eudemonic measures aim to capture a range of factors that can be considered important, but are not necessarily reflected in evaluative or experience measures and can include autonomy, control, competence, engagement, good personal relationships, a sense of meaning, purpose and achievement. These types of measures are also sometimes known as measures of ‘flourishing’ (Tinkler and Hicks, 2011).

2.1. Choice of ONS subjective well-being questions

Before developing the ONS four subjective well-being questions a review was undertaken of existing subjective well-being questions both in the UK and abroad (ONS, 2010). ONS also sought academic advice from Prof. Paul Dolan (LSE), Prof. Lord Richard Layard (LSE), Dr Robert Metcalfe (Oxford University) and Prof. Felicia Huppert (Cambridge University) for the development of the ONS four subjective well-being questions. The subject was also discussed at the ONS National Statistician’s Advisory Forum and the ONS Technical Advisory Group. These groups consisted of a range of specialists including representatives from the Organisation for Economic Co-operations and Development (OECD), Eurostat (the statistical office for the European Union), other UK government departments and academics.

From April 2011, ONS introduced four subjective well-being questions onto its largest household survey covering evaluative, eudemonic and experience measures of well-being. The aim was to develop a balanced set of subjective well-being questions which took account of the different approaches to measuring Subjective Well-being. The four questions are as follows:

- overall, how satisfied are you with your life nowadays? (experience),
- overall, to what extent do you feel the things you do in your life are worthwhile? (Eudemonic),
- overall, how happy did you feel yesterday? (positive affect),
- overall, how anxious did you feel yesterday? (negative affect).

All questions use a 0 – 10 scale. A copy of the questionnaire showing the four questions can be found in Appendix 1.

Evaluative

The ONS life satisfaction question was selected to reflect the evaluative approach to measuring subjective well-being which is widely used and established both within the UK and internationally. Similar questions have been asked on many other surveys, for example the World Values Survey, European Social Survey and the ‘Understanding Society’ survey.

Across different surveys in the UK and internationally there are instances when evaluative measures do not have an explicit time frame of assessment. The use of time frame constraint is an important difference as some respondents may find it difficult to evaluate their life satisfaction when no specific time frame is provided in the question (Waldron, 2010; Dolan et al., 2011). Without a time frame the immediate context in which the question is being asked may also have more of an influence on responses. ONS, therefore, made the decision to use the term ‘nowadays’ in the life satisfaction question. This has been used in other surveys and, although this leaves the respondent to make a judgement about how to interpret the time frame, it does limit the reference period to more recent times rather than encouraging the respondent to consider their life as a whole (Tinkler & Hicks, 2011).

Eudemonic

For the eudemonic measure ONS selected a question asking respondents the extent to which they feel the things they do in their lives are worthwhile. This question was selected as one that provides information on how much meaning and purpose people get from the things that they do in their lives. Developing an overall question to measure the eudemonic approach was more challenging because often a larger set of questions are used to pick up the different dimensions of this approach. The question that ONS selected was adapted from the European Social Survey (ESS) which asks about what people ‘do in their lives that is valuable and worthwhile’. The question was reworded to fit with a 0-10 scale, and a similar question was tested by ONS prior to inclusion. ONS decided not to include the words ‘valuable and worthwhile’ but to instead use only the term ‘worthwhile’ as advice from cognitive testing experts from within ONS and some members of the Technical Advisory Forum suggested that respondents could see these concepts as distinct rather than complementary. ONS received advice on the precise wording from Felicia Huppert from Cambridge University, an expert in eudemonic measures (Tinkler and Hicks, 2011).

Experience or affect

In terms of an experience measure ONS concluded that it was necessary to include not only a positive but also negative ‘experience’ question after receiving comments from the Measuring National Well-being Technical Advisory Group. ONS also decided to use the time frame of ‘yesterday’ in order to approximate to the Day Reconstruction Method (DRM) time use approach. The adjective ‘happiness’ was chosen as it is commonly used for positive affect questions and

has been used both in the DRM and Gallup-Health Ways data. ‘Anxious’ is widely used as an indicator for poor mental well-being, for example it is used in the European Social Survey and in the European Quality of life-5 Dimensions (EQ-5D) well-being measure (Dolan 2011).

2.2. Response scales

Different response scales for subjective well-being questions have been used on different surveys (Stevenson and Wolfers, 2008). For example, on the ‘Understanding Society’ survey respondents are asked ‘How dissatisfied or satisfied are you with.....your life overall’ with responses on a scale of 1 to 7 where 1 is ‘Not satisfied at all’ and 7 is ‘Completely satisfied’. However, the subjective well-being questions on the Euro barometer Survey use a four point scale: ‘On the whole are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with the life you lead?’ Differences in scaling and labelling of scale points require careful consideration as a different scale may affect how people respond. Additionally, the use of different scales in different surveys means that it is more difficult to compare across different sources of data.

ONS decided that an 11 point scale from 0-10 where 0 is ‘not at all’ and 10 is an absolute value such as ‘completely’ should be used for the ONS subjective well-being questions. The reason for this decision was to ensure that the scales between the questions are consistent in order to help respondents answer the questions more easily and also to aid analysis across the separate questions. Further to this, 11 point scales of this nature are commonly used across other surveys of interest, particularly internationally, and using the same type of scale will aid comparisons with these estimates.

3. Implementing subjective well-being questions into ONS data collection

Once the four subjective well-being questions had been decided on, ONS had to consider how the questions would be introduced on ONS surveys; including which surveys to introduce the questions on, and placement of questions within surveys.

3.1. Question placement

Feedback from some members of the Measuring National Well-being Technical Advisory Group highlighted the likely impact the placement of the questions would have on the estimates. For example, placing after questions relating to health or the labour market may impact the answers that respondents make. Prior to April 2011, ONS carried out small scale cognitive testing of the placement of the subjective well-being questions on ONS surveys. It was decided that the placement of the overall monitoring questions should be close to the

beginning of the questionnaire after the basic questions on household and individual demographics. This placement allows time for rapport to be built up between the interviewer and the respondent by the time the subjective well-being questions are asked without allowing later questions, such as those on employment, to influence response to the subjective well-being questions.

3.2. Choice of ONS surveys for the subjective well-being questions

Due to resource constraints and ONS consideration of respondent burden, the four subjective well-being questions were added to existing ONS surveys rather than creating a new ONS well-being survey. The four ONS subjective well-being questions were added to the ONS Annual Population Survey (APS), the UK's largest household survey from April 2011.

The APS Personal Well-being Dataset contains approximately 165,000 directly questioned adults per year. The APS carries a wide range of objective variables including: sex, age, socio-economic class, ethnicity, national identity, country of birth, religious affiliation, smoking, health, income, employment status and qualification levels. This has allowed ONS to analyse a wide range of determinants of subjective well-being as well as conduct analysis at low levels of geography. The APS is constructed from the quarterly Labour Force Survey, and it therefore offers a stable platform for the subjective well-being questions to be asked on as it is unlikely to be subject to any major cuts, changes or revisions. It is important to note, however, that although the APS collects data on individual earnings it does not collect data on household income. Furthermore, the APS is not designed for longitudinal analysis which means that an individual's well-being cannot be monitored over time.

Additionally, ONS has been trialling a number of different subjective well-being questions on the monthly Opinions and Lifestyle Survey (OPN) which is discussed in more detail in section 4 of this article (Hicks, Tinkler and Allin, 2013).

In order to increase policy impact of the four subjective well-being questions, in addition to the APS and Opinions and Lifestyle survey the questions have now been introduced on the following ONS surveys:

- Crime Survey for England and Wales;
- Wealth and Assets Survey;
- Living costs and food survey.

The Crime Survey for England and Wales is an annual survey consisting of a sample size of approximately 50,000 households. The survey measures the extent of crime in England and Wales by asking people whether they have experienced any crime in the past year. The four subjective well-being questions were introduced onto this survey in 2012-13 to allow analysis of the impact that being a victim of different types of crime has on subjective well-being.

The Wealth and Assets Survey (WAS) is a longitudinal survey that covers Great Britain (England, Wales, and Scotland). Respondents to wave one of the

survey are invited to a follow-up interview two years later. The sample in wave one is approximately 30,000 households, and the sample in wave two is approximately 20,000 households. WAS collects information on the level of household assets, savings and levels of debt. There is strong demand from both policy makers and academics to analyse an individual's well-being over time; the longitudinal element of WAS will allow this to be possible. Analysing subjective well-being data combined with wealth and debt is also key for the use of subjective well-being data within policy.

The Living Costs and Food Survey (LCF) collects information on household spending patterns, cost of living, and household spending patterns. It has an annual sample size of 6,000 respondents. This is data not held on the APS and will make an important contribution to the understanding that household income and spending patterns have on subjective well-being.

The importance of the well-being of children and young people emerged as an important theme of the ONS National Debate. As most ONS surveys are administered to respondents aged 16 and above, ONS worked closely with the Children's Society and other organisations to ensure that the subjective well-being of children and young people is also measured. The four questions were cognitively tested on children aged 8 to 15 years old. Only the questions on life satisfaction, worthwhile and happiness yesterday were introduced on surveys for children, as cognitive testing showed that the anxiety question was not well understood. These three subjective well-being questions were introduced on the 'Understanding Society – the UK household longitudinal study' which is run by the Institute for Social and Economic Research, the sample of which includes children aged 11 to 15. Additionally, The Children's Society in the UK has conducted a regular on-line well-being survey since 2010 the sample of which includes both children and parents. In autumn 2014 ONS published an analysis of these data 'Exploring the well-being of children in the UK, 2014' (ONS 2014c).

ONS promotes the use of the four ONS subjective well-being questions for use on surveys outside the ONS within other government departments, local government, charities and the private sector. For a full list of surveys containing the ONS four subjective well-being questions see Appendix 2.

4. Testing and development of ONS subjective well-being questions

The collection of subjective well-being data is a new area for ONS and as such ONS has undertaken a thorough investigation regarding the quality of subjective well-being data. Item non-response of the four subjective well-being questions has been investigated as one indication of quality. It was found that there was very little non-response to the four subjective well-being questions from respondents taking part in the APS as there was a response rate of approximately 99%. This is very positive and indicates a general acceptance of the subjective

well-being questions by respondents, as well as the ability of respondents to answer the questions without difficulty (Tinkler and Hicks, 2011).

ONS has also investigated how the four overall subjective well-being questions are associated with each other to investigate how much extra information is gained from each of the questions and how similar or apparently different questions actually are. Analysis from both the OPN and the APS shows that although the four ONS subjective well-being questions are correlated, they appear to be picking up different concepts. In particular, it seems the experience questions (happiness and anxiety yesterday) are different from the evaluative and eudemonic questions (ONS 2011a). This indicates that collecting information from all four subjective well-being questions is a valid exercise.

Since the four questions were introduced on the APS in April 2011, ONS has been conducting quantitative testing in parallel with qualitative investigation including cognitive testing. The quantitative testing has consisted of split trials including comparing mode of interview, different orders of the four subjective well-being questions, alternative question wording, use of show-cards, and alternative preambles to the subjective well-being questions. It has also been possible to begin to look at how the day of the week and different months of the year affect subjective well-being (ONS, 2011a and ONS, 2012b).

The ONS has carried out two main phases of cognitive testing. The first phase focused on investigating the four subjective well-being questions. A combination of telephone and face-to-face interviews were carried out to reflect the modes of interview on the APS. Respondents' reactions and interpretations of the questions and scales were investigated, as well as the use of question scale show cards (ONS 2011b).

A second phase of cognitive testing was carried out in 2013 which also concentrated on the four headline subjective well-being questions, as well as investigating alternative preambles and wording of the four headline questions.

It is outside the scope of this paper to provide detailed results on all the testing that has been undertaken on the ONS subjective well-being questions; however, some of the testing is documented in the following publications (ONS, 2011a and ONS, 2012).

ONS plans to publish further results of the Personal Well-being split trial testing in early 2015.

4.1. Presentation of subjective well-being analysis

Before presenting subjective well-being results consideration was given to what would be most meaningful to the user. It is key that measures are presented effectively for their use by policy makers and the wider public. ONS consulted with experts including Prof. Paul Dolan (LSE), Prof. Lord Richard Layard (LSE) and the Organisation for Economic Co-operation and Development (OECD). This issue was also consulted on with members of the ONS Measuring National Well-being Technical Advisory Group.

ONS displays subjective well-being data in a number of ways including:

- Response distributions along the 0 – 10 scale
- Mean averages
- Grouped thresholds, where the percentage of people that fall above or below a certain threshold on the 11 point scale is displayed

The use of thresholds is an effective way of presenting subjective well-being data as the threshold groupings highlight the distribution of subjective well-being along the 0 – 10 scale. This distribution is hidden if mean averages only are used to display data. Further advantages of using threshold groupings are that they show year-on-year change more effectively than mean averages, which is of key interest for policy makers. There are however, some limitations with the use of thresholds to display subjective well-being data. For example, when displaying analysis of subjective well-being for sub-groups with many categories it is not always possible to display these data using thresholds due to small sample sizes. In addition, there is a practical difficulty of presenting graphically a large sub group by each of the four threshold groupings.

Consideration was also given to the creation of a subjective well-being index to disseminate ONS subjective well-being results in order to provide one headline measure of subjective well-being. ONS decided not to produce an aggregated measure such as this because the ONS subjective well-being questions are designed to collect information on three fundamentally different concepts and should therefore be kept separate to allow analysis of each of the distinct elements of subjective well-being.

In addition to consulting with topic experts on the presentation of subjective well-being measures ONS has also consulted on the presentation of subjective well-being measures with non-expert users. In 2013 ONS carried out focus groups with non-experts regarding personal well-being outputs. The aim of these focus groups was to investigate how members of the public responded and interpreted ONS subjective well-being outputs. This exercise led to changes in the design, use of colour and types of graphs produced to make them more appealing to look at and easier to interpret. This also led to the development of ONS referring to ‘Personal Well-being’ within ONS publications rather than ‘Subjective Well-being’ as the term ‘Subjective well-being’ was not well understood by non-expert users.

5. Key findings

ONS has published three annual analytical reports from the Annual Population Survey. Some of the most notable results were the relationship between subjective well-being and age, health and employment status. ONS also presents results by different levels of geography including the four countries of the UK, and English regions in addition to smaller local levels of geography.

Figure 1 displays results of analysis of age and subjective well-being ‘Average personal well-being, by age group’. It shows that life satisfaction, a sense that what ones does in life is worthwhile, and happiness well-being is lowest in the middle years (45-54), higher for younger age groups, and peaks in the 65-79 age group, while anxiety peaks in the 45-54 age group, and is lowest for the 65 and over age group. These findings reflect previous research into subjective well-being where the relationship between subjective well-being and age has been described as ‘U’ shaped; highest in the younger and older age groups, and lowest in the middle years.

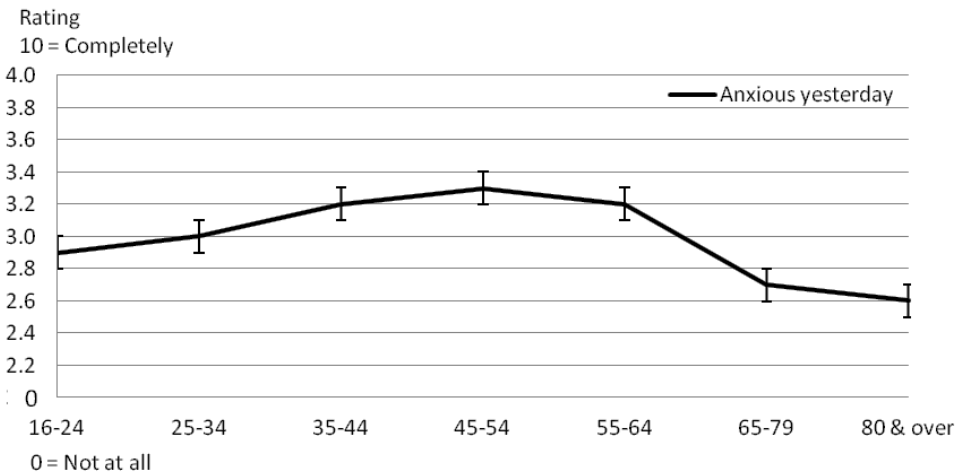
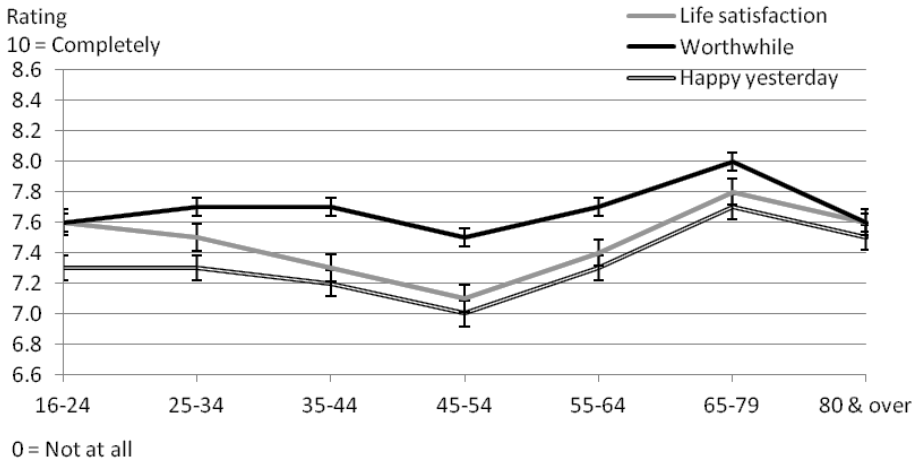


Figure 1. Average subjective well-being, by age group, 2012/2013 United Kingdom

Source: Annual Population Survey (APS) - Office for National Statistics.

Figure 2 displays an analysis which compares self-reported health with subjective well-being. The graph clearly shows that those in very good health have the highest life satisfaction, highest sense that the things they do in life are worthwhile, highest happiness, and the lowest anxiety.



Figure 2. Average subjective well-being, by self reported health, 2012/13 United Kingdom

Source: Annual Population Survey (APS) - Office for National Statistics.

Regression analysis

In order to understand the drivers of well-being in greater depth ONS have carried out a number of regression analyses. The key benefit of regression analysis is that it provides a more refined method of identifying factors which are associated most with subjective well-being compared with an analysis where the relationship between only two factors at a time is considered.

ONS has published four regression analyses using personal well-being data. Published in May 2013 ‘Measuring National Well-being – What matters most to Personal Well-being?’ (Oguz et al. 2013) analysed factors within the Annual Population Survey and considered how these are associated with subjective well-being. Of the variables available within the APS, self-reported health had the strongest association with all four measures of subjective well-being. The second strongest association was employment status, and the third was relationship status. Sense of choice and contentment also appear to be associated with

subjective well-being. For example, those who are employed but want a different or additional job have lower levels of personal well-being than employed people who are not looking for another job (Oguz et al. 2013).

Commuting and Personal Well-being was published in February 2014 (ONS, 2014a) and examined the relationship between commuting to work and personal well-being using regression analysis. The analysis found that commuters have lower life satisfaction, a lower sense that their daily activities are worthwhile, lower levels of happiness and higher anxiety on average than non commuters. The most negative effects of commuting on personal well-being were associated with journey times lasting between 61 and 90 minutes. On average all four aspects of personal well-being were negatively affected by commutes of this duration when compared to those travelling only 15 minutes or less to work (ONS, 2014a).

Income, Expenditure and Personal Well-being was published in June 2014 (ONS, 2014b) and provided new findings on the relationship between personal well-being and household income and expenditure. The Effects of Taxes and Benefits on Household Income dataset was analysed which was derived from the Living Costs and Food Survey (LCF). The analysis found that individuals in households with higher incomes report higher life satisfaction and happiness, and lower anxiety. Higher household income, however, was not significantly related to the question measuring eudemonic subjective well-being. An increase in the proportion of household income from cash benefits received from the state such as housing benefits and Jobseeker's allowance was associated with lower well-being across all four measures, with the effects stronger for men than for women. This effect remained even when taking differences in household income into account.

Interestingly, household expenditure had a stronger relationship with people's life satisfaction, sense that the things they do in life are worthwhile and happiness than household income. There was no significant relationship between higher household expenditure and lower anxiety (ONS 2014b).

ONS has also published a working paper 'Exploring Personal Well-being and Place' which analysed the relationship between personal well-being and location of residence.

For a full list of ONS subjective well-being publications see Appendix 3.

6. National Statistics status

In September 2014 ONS Personal Well-being was granted National Statistics status by the UK Statistics Authority. The UK Statistics Authority is an independent body directly accountable to Parliament. The Authority's statutory objective is to promote and safeguard the production and publication of official statistics. The Authority provides independent scrutiny in the form of a rigorous assessment of all official statistics produced in the UK. If the official statistics are

deemed to be of high quality across a number of dimensions they are granted 'National Statistics' by the UK Statistics Authority.

To meet National Statistics requirements ONS Subjective Well-being statistics underwent the UK Statistics Authority's assessment process for official statistics. ONS statistics were assessed against the Statistics Authority 'Code of Practice' for Official Statistics.

The code of practice relates to several dimensions of statistical quality including:

- meeting user needs,
- impartiality and objectivity,
- sound methods and assured quality,
- confidentiality,
- proportionate burden,
- resources,
- frankness,
- accessibility.

(UK Statistics Authority 2009).

See Appendix 4 for more details.

In order for data to gain National Statistics status, each of the eight principles of the code of practice had to be adhered to.

ONS had to provide the most recent subjective well-being publications as well as details of analysis and quality assurance procedures. In addition, evidence regarding the ways that the statistics are used by members of the public, and by government, were documented and provided to the Statistics Authority.

After receiving these documents The Authority submitted them to the Assessment Committee. The Assessment Committee asked ONS to fulfil a number of requirements and suggestions for improvement which ONS has a set time frame to address including:

- to publish plans on further public engagement,
- to add more information in the methodology section of statistical reports regarding the impact of the different modes of data collection, and the strengths and limitations of the estimates,
- Changes to the commentary and the way that the data was presented graphically within statistical reports.

ONS documented in detail how each requirement would be addressed. Additionally, where applicable, ONS made the suggested amendments to a statistical report and also submitted this to the statistics authority for evidence. ONS subjective well-being statistics were then granted National Statistics.

This development was an important achievement for the ONS subjective well-being statistics, and will help to embed these statistics firmly into policy. This has also helped to achieve one of the important aims of the ONS National Well-being

programme, which is to develop a trusted set of statistics to measure the well-being of the nation.

7. Latest developments and next steps

Key to the acceptance, legitimacy and success of ONS subjective well-being measures has been the focus on stakeholder and citizen engagement. This started with the national debate and has continued around the collection, analysis and dissemination of subjective well-being data. ONS will continue to engage with these groups to ensure that ONS subjective well-being statistics continue to meet user needs. For example, in response to user demand ONS has produced an aggregated three year personal well-being dataset, to provide more robust local area and sub-group estimates. The dataset was released in October 2014 and it is planned that a rolling three-year personal well-being dataset will be produced annually. In March 2015 ONS published interactive maps to engage users with the three year dataset, in addition to producing more detailed analytical reports based on these data.

ONS plan to build on existing established relationships with policy makers to ensure that use of the measures of subjective well-being become fully embedded within government policy at all levels. From the outset of the collection of subjective well-being data ONS has been working with the Social Impacts Taskforce which was set up in August 2010 with the aim of developing a cross Government approach to understanding and embedding social impacts into policymaking. Members of the taskforce include UK civil service departments as well as the devolved administrations in Scotland, Wales and Northern Ireland.

In October 2014 the ‘What Works Centre for Well-being’ (WWCW) was set up. WWCW is an independent centre dedicated to making policy and services work for well-being and will commission universities to research the impact that different interventions and services have on well-being. The aim is that the results of this research will help government, councils, health and well-being boards, charities and businesses make decisions regarding the well-being of people, communities and the nation as a whole. ONS will be working closely with the WWCW to ensure that the ONS approach to measuring subjective well-being and the results of ONS research in this area is widely promoted.

ONS is aware that data from ONS household surveys only captures information from respondents living in private households and excludes people living in communal establishments (a diverse set of premises including hotels, guest houses and nursing homes for example). Although this is a relatively small part of the population, approximately 1.8 per cent, ONS is very much aware of this issue and of the challenges in surveying people in non-household populations (Tinkler and Hicks, 2011).

UK work on measuring national well-being is highly regarded internationally and ONS is a member of several international working groups as part of the ‘GDP

and beyond' agenda. ONS sees it as a priority to work collaboratively with international partners to achieve, where possible, consistency with international standards and concepts in order to increase the value of these statistics. ONS will continue to work with international organisations such as the Organisation for Economic Co-operation and Development (OECD) and Eurostat.

Specifically regarding subjective well-being ONS contributed to the OECD handbook on 'Measuring Subjective Well-being' and regularly take part in various OECD-led high level meetings on measuring well-being. ONS were part of the Eurostat taskforce to create an ad-hoc quality of life module in 2013 on the European Survey of Income and Living Conditions (EU SILC). In 2015 ONS will participate in another task force to develop a six yearly module on 'Quality of life and social and cultural participation' which will include subjective well-being.

Following the introduction of the four subjective well-being questions on the ONS 'Wealth and Assets Survey' ONS plan to conduct a regression analysis of these data in 2015. This will include variables on household income and debt.

ONS will continue to engage with a wide range of stakeholders including international organisations, policy makers as well as citizen users to ensure that ONS subjective well-being analysis is relevant. This will help to ensure that the ONS subjective well-being questions continue to be one of the key evidence bases for government policy related to well-being.

REFERENCES

- AKED, J., MARKS, N., CORDON, C., THOMPSON, S., (2008). Five ways to well-being: The evidence. A report presented to the Foresight Project on communicating the evidence base for improving people's wellbeing. London: The New Economics Foundation (NEF). Available at: <http://www.neweconomics.org/publications/five-ways-well-being-evidence>.
84 S. Hicks et al.
- ALLIN, P., (2014). Measuring Wellbeing in Modern Societies. In P.Y.Chen and C.L. Cooper (eds.) *Work and Wellbeing: Wellbeing: A Complete Reference Guide, Volume III*. John Wiley & Sons, New York, 409–465.
- CAMERON, D., Prime Minister of the UK, (2010). 25 November. Prime Minister Speech on Well-being. Available at: <http://www.number10.gov.uk/news/pm-speech-on-well-being>.
- CLERC, M., GAINI, M., BLANCHET, D., (2010). Recommendations of the Stiglitz-Sen-Fitoussi Report: A few illustrations, *L'économie française*—2010 edition, Insee, Available as pdf at http://www.insee.fr/en/publications-et-services/default.asp?page=dossiers_web/stiglitz/performance_eco.htm.
- Commission, European, (2009). *GDP and beyond: Measuring progress in a changing world*. Luxembourg: European Commission.
- DEATON, A., (2011). The financial crisis and the well-being of Americans, *Oxford Economic Papers*, May 2011, 1-26 OEP Hicks Lecture (November 2011). Available at: http://www.princeton.edu/*deaton/papers.html.
- DIENER, E., (1994). Assessing subjective well-being: Progress and opportunities. *Social Indicators Research*, 31, 103–157.
- DIENER, E., (2005). Guidelines for national indicators of subjective well-being and ill-being, December 2005, Available at: http://internal.psychology.illinois.edu/*ediener/Documents/Guidelines_for_National_Indicators.pdf.
- DOLAN, P., LEYARD, R. Notes on the measurement of well-being. Available at www.ons.gov.uk/wellbeing/technical-advisory-group/notes-on-the-measurement-of-well-being-11-april-2011.pdf.
- DOLAN, P., WHITE, M. P., (2007). How can measures of Subjective well-being be used to inform public policy? *Perspectives on Psychological Science*, 2, 71–85.
- DOLAN, P., METCALF, R., (2011). Comparing measures of subjective well-being and views about the role they should play in policy: Office for National Statistics.

- DOLAN, P., LAYARD, R., METCALF, R., (2011). Measuring subjective wellbeing for public policy; recommendations on measures: Special paper No. 23. Centre for Economic Performance. Available at: <http://cep.lse.ac.uk/pubs/download/special/cepsp23.pdf>.
- DOLAN, P., PEASGOOD, T., WHITE, M. P., (2006). Review of research on the influences on personal well-being and application to policy making. London: DEFRA Publications.
- DOLAN, P., PEASGOOD, T., WHITE, M., (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology*, 29, 94–122.
- DONOVAN, N., HALPERN, D., (2002). Life satisfaction: The state of knowledge and implications for Government. London: Prime Minister's Strategy Unit.
- FRANCIS, P., TIWANA, H., (2004). Unpaid household production in the United Kingdom, 1995–2000. Available at: <http://www.ons.gov.uk/ons/rel/elmr/economic-trends-discontinued-/no-602-january-2004/unpaid-household-production-in-the-uk-1995-2000.pdf>.
- FUJIWARA, D., CAMPBELL, R., (2011). Valuation techniques for social cost-benefit analysis: Stated preference, revealed preference and subjective well-being approaches. Available at: http://www.hm-treasury.gov.uk/d/green_book_valuationtechniques_250711.pdf.
- GOV.UK., (2014). New what works centre for well-being: <https://www.gov.uk/government/news/new-what-works-centre-for-wellbeing>.
- HALPERN, D., (2010). *The Hidden Wealth of Nations*. Polity Press.
- HARSANYI, J. C., (1982). Morality and the theory of rational behaviour. In A. Sen & B. Williams (Eds.), *Utilitarianism and beyond*, pp. 39–63, Cambridge: Cambridge University Press.
- HELLIWELL, J. F., (2003). How's life? Combining individual and national variables to explain Subjective wellbeing. *Economic Modelling*, 20, 331–360.
- HELLIWELL, J. F., BARRINGTON-LEIGH, C. P., (2008). Measuring and Understanding Subjective well-being. NBER Working Paper Series, 15887.
- HELLIWELL, J. F., (2011). How can subjective well-being be improved? In F. Gorbet, & A. Sharpe (Ed.) *New directions for intelligent government in Canada: Papers in Honour of Ian Stewart*, Centre for the Study of Living Standards, pp. 283–304, 2011. Available at: <http://ideas.repec.org/h/sls/ndphis/12.html>.

- HICKS, S., (2011a). Spotlight on: Subjective well-being: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/social-trends-rd/social-trends/spotlight-on-subjective-well-being/index.html>.
- HICKS, S., (2011b). The measurement of subjective well-being; Paper for Measuring National Well-being.
- Technical Advisory Group [pdf]. Newport: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/guide-method/user-guidance/well-being/advisory-groups/well-being-technical-advisory-group/index.html>. Adapted from: Dolan, Layard & Matcalfe, 2011 and Smith, 2011. OECD, unpublished.
- HICKS, S., TINKLER, L., (2011). Draft Testing and Development plan for Subjective well-being questions on ONS Surveys. Newport: Office for National Statistics. Available at: www.ons.gov.uk/well-being/technical-advisory-group/draft-testing-and-development-plan-11-april-2011.pdf.
- HICKS, S. TINKLER, L. ALLIN, P., (2013). Measuring Subjective Well-being and its potential role in policy: Perspectives from the UK Office for National Statistics. Newport: Office for National Statistics.
- INGLEHART, R., FOA, R., PETERSON, C., WELZEL, C., (2008). Development, freedom, and rising happiness. *Perspectives on Psychological Science*, 3, 264–285. Available at: <http://pps.sagepub.com/content/3/4/264>.
- KAHNEMAN, D., KRUEGER, A., (2006). Developments in the measurement of Subjective well-being. Available at www.krueger.princeton.edu/PDF%20of%20Kahneman%20Krueger%20paper.pdf.
- KAHNEMAN, D., KRUEGER, A. B., SCHKADE, D. A., SCHWARZ, N., STONE, A. A., (2004). A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. *Science*, 306, 1776–1780.
- LAYARD, R., (2005). *Happiness: Lessons from a new science*. London: Penguin.
- NEF, (2008). Five ways to well-being. Available at: www.neweconomics.org/publications/five-ways-to-wellbeing.
- NEF, (2011). ‘Measuring our progress’ Available at: www.neweconomics.org/publications/measuring-our-progress.
- NISSEL, M., (1970). *Social trends*, Vol. No. 1, Central Statistics Office, Her Majesty’s Stationary Office.
- OECD, (2011). *Better life initiative: Measuring well-being and progress*. Paris: OECD.
- ONS, (2011a). *Measuring what matters; National statistician’s reflections on the national debate on measuring national well-being*. London: Office for National Statistics.

- ONS, (2011b). Initial investigation into subjective well-being from the opinions survey. Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-subjective-wellbeing-inthe-uk/investigation-of-subjective-well-being-data-from-the-ons-opinions-survey/index.html>.
- ONS, (2011c). Subjective well-being: A qualitative investigation of subjective well-being questions. December 2011.
A working paper for the measuring national well-being advisory group. Available at: <http://www.ons.gov.uk/ons/guide-method/user-guidance/well-being/advisory-groups/well-being-technical-advisory-group/index.html>.
- ONS, (2012a). Analysis of experimental subjective well-being data from the annual population survey, April to September 2011. Newport: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-subjective-wellbeing-in-the-uk/analysis-of-experimental-subjective-well-beingdata-from-the-annual-population-survey-april-september-2011/index.html>.
- ONS, (2012b). First ONS annual experimental subjective well-being results. Newport: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-subjective-wellbeing-in-the-uk/firstannual-ons-experimental-subjective-well-being-results/index.html>.
- ONS, (2012c). Summary of results from testing of experimental Subjective Well-being questions [PDF]. Newport: Office for National Statistics.
- ONS, (2013a). Measuring Subjective Well-being in the UK, Differences in Well-being by Ethnicity. Newport: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-subjective-wellbeing-in-the-uk/differences-in-well-being-by-ethnicity/index.html>.
- ONS, (2013b). Measuring National Well-being, What matters most to Personal Well-being? Newport: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/what-matters-most-to-personal-well-being-in-the-uk-/index.html>.
- ONS, (2013c). Measuring National Well-being, Personal Well-being in the UK, 2012/13. Newport: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/personal-well-being-in-the-uk-2012-13/index.html>.
- ONS, (2013d). Measuring National Well-being, Personal Well-being Across the UK, 2012/13. Newport: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/personal-well-being-across-the-uk-2012-13/index.html>.

- ONS, (2014a). Measuring National Well-being, Commuting and Personal Well-being, 2014. Newport: Office for National Statistics.
Available at <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/commuting-and-personal-well-being-2014/index.html>.
- ONS, (2014b). Measuring National Well-being, Income, Expenditure and Personal Well-being, 2011/2012. Newport: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/income-expenditure-and-personal-well-being/index.html>.
- ONS, (2014c). Exploring Personal Well-being and Place. Titchfield: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/rel/regional-trends/regional-economic-analysis/exploring-personal-well-being-and-place-in-the-uk/index.html>.
- ONS, (2014d). Insights across society, the Economy and the Environment, Reflections on Measuring National Well-being, July 2014. Newport: Office for National Statistics. Available at:
<http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/reflections-on-measuring-national-well-being-may-2014/index.html>.
- ONS, (2014e). Measuring National Well-being, Personal Well-being in the UK, 2013/14. Newport: Office for National Statistics.
Available at: <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/personal-well-being-in-the-uk-2013-14/index.html>.
- ONS, (2014f). Measuring National Well-being, Exploring the Well-being of Children in the UK, 2014. Newport: Office for National Statistics.
Available at:
<http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/exploring-the-well-being-of-children-in-the-uk-2014/index.html>.
- PUDNEY, S., (2010). An experimental analysis of the impact of survey design on measures and models of subjective well-being. Institute for Social and Economic Research University of Essex.
Available at: www.iser.essex.ac.uk/publications/working-papers/iser/2010-20.pdf.
- RYFF, C. D., (1989). Happiness is everything, or is it? Explorations on the meaning of psychological wellbeing. *Journal of Personality and Social Psychology*, 57, 1069–1081.
- SCHWARZ, N., STRACK, F., KOMMER, D., WAGNER, D., (1987). Soccer, rooms and the quality of your life: Mood effects on judgements of satisfaction with life in general and with specific life – domains. *European Journal of Social Psychology*, 17, 69–79.

SEN, A., (1999). *Development as freedom*. Oxford: Oxford University Press.

SNAPE, D., (2014). *UK's Methodological approach to measuring personal well-being*. Newport: Office for National Statistics.

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

STIGLITZ, J. E., SEN, A., FITOUSSI, J. P., (2009). *Report by the commission on the measurement of economic performance and social progress*. Available at: http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.

The Government Office for Science, (2008). *Foresight mental capital and wellbeing project. Final project report*. London. Available at: *Mental Capital and Wellbeing: Making the most of ourselves in the 21st century, 2008*.

TINKLER, L., HICKS, S., (2011). *Measuring subjective well-being*. Newport: Office for National Statistics. Available at: <http://www.ons.gov.uk/ons/index.html>.

WALDRON, S., (2010). *Measuring subjective wellbeing in the UK*. Office for National Statistics. Available at: <http://www.statistics.gov.uk/cci/article.asp?ID=2578>.

UK Statistics Authority. *Producer handbook for the assessment process*. Newport: UK Statistics Authority. Available at: <http://www.statisticsauthority.gov.uk/assessment/assessment/guidance-about-assessment/producer-handbook-for-the-assessment-process.html>

APPENDICES

Appendix 1.

The four ONS subjective well-being questions

SUBJECTIVE WELLBEING

68. *Satis*

UK

Next I would like to ask you four questions about your feelings on aspects of your life. There are no right or wrong answers. For each of these questions I'd like you to give an answer on a scale of nought to 10, where nought is 'not at all' and 10 is 'completely'.

Overall, how satisfied are you with your life nowadays?

Interviewer instruction: where nought is 'not at all satisfied' and 10 is 'completely satisfied'

Scale from 0 to 10

Asked if: PersProx = IN PERSON, DVAge = 16+

69. *Worth*

UK

Overall, to what extent do you feel that the things you do in your life are worthwhile?

Interviewer instruction: where nought is 'not at all worthwhile' and 10 is 'completely worthwhile'

Scale from 0 to 10

Asked if: PersProx = IN PERSON, DVAge = 16+

70. *Happy*

UK

Overall, how happy did you feel yesterday?

Interviewer instruction: where nought is 'not at all happy' and 10 is 'completely happy'

Scale from 0 to 10

Asked if: PersProx = IN PERSON, DVAge = 16+

71. *Anxious*

UK

On a scale where nought is 'not at all anxious' and 10 is 'completely anxious', overall, how anxious did you feel yesterday?

Scale from 0 to 10

Asked if: PersProx = IN PERSON, DVAge = 16+

Thank you, that is the end of this section of questions.

Appendix 2.

Surveys the 4 ONS personal well-being questions have been included on:

- Crime Survey for England and Wales - Office for National Statistics - previously conducted by Home Office
- Civil Service People Survey United Kingdom - Civil Service
- Wealth and Assets Survey - Office for National Statistics

- Life Opportunities Survey - Department for Work and Pensions and Office for National Statistics
- The National Study of Work-search and Well-being findings - Department for Work and Pensions - NatCen
- Survey regarding population of employees - Department for Work and Pensions
- Armed Forces Continuous Attitude Survey - Ministry Of Defence
- Families Continuous Attitude Survey - Ministry Of Defence
- Impact of Further Education Learning Survey - Business Innovation and Skills
- The National Survey for Wales - Welsh Government
- Community Life Survey - Cabinet Office
- Labour Force Survey/Annual Population Survey - Office for National Statistics
- Opinions and Lifestyle Survey - Office for National Statistics
- Living Costs and Food Survey - Office for National Statistics
- Taking Part Survey - Department Culture Media and Sport
- National Citizenship Service evaluation - Cabinet Office
- English Longitudinal Study of Ageing
- English Housing Survey - Department Communities Local Government commission from NatCen
- Quarterly National Household Survey - Central Statistics Office - Ireland
- Monitor of Engagement with the Natural Environment: The Natural Survey on People and the Natural Environment - Natural England

Appendix 3.

Office for National Statistics publications on Measuring Subjective/Personal Well-being:

- Measuring Subjective Well-being in the UK (September 2010)
- Measuring Subjective Well-being (July 2011)
- Initial investigation into Subjective Well-being from the Opinions Survey (December 2011)
- Analysis of Experimental Subjective Well-being Data from the Annual Population Survey, April to September 2011 (February 2012)
- First ONS Annual Experimental Subjective Well-being Results (July 2012)
- Measuring National Well-being: Programme Overview, the place of Subjective Well-being and Recent Findings (December 2012)
- Differences in Well-being by Ethnicity (April 2013)
- Personal Well-being in the UK, 2012/13 (July 2013)
- Personal Well-being Across the UK, 2012/13 (October 2013)
- Commuting and Personal Well-being, 2014 (February 2014)
- Income, Expenditure and Personal Well-being, 2011/2012 (June 2014)

- Exploring Personal Well-being and Place (June 2014)
- Personal Well-being in the UK, 2013/14 (September 2014)
- 3 year dataset, 2011/2014 and Smoking and Personal Well-being in Bristol (October 2014)

Appendix 4.

UK statistics Authority Code of Practice:

Principle 1: Meeting user needs	The production, management and dissemination of official statistics should meet the requirements of informed decision-making by government, public services, business, researchers and the public.
Principle 2: Impartiality and objectivity	Official statistics, and information about statistical processes, should be managed impartially and objectively.
Principle 3: Integrity	At all stages in the production, management and dissemination of official statistics, the public interest should prevail over organisational, political or personal interests.
Principle 4: Sound methods and assured quality	Statistical methods should be consistent with scientific principles and internationally recognised best practices, and be fully documented. Quality should be monitored and assured taking account of internationally agreed practices.
Principle 5: Confidentiality	Private information about individual persons (including bodies corporate) compiled in the production of official statistics is confidential, and should be used for statistical purposes only.
Principle 6: Proportionate burden	The cost burden on data suppliers should not be excessive and should be assessed relative to the benefits arising from the use of the statistics.
Principle 7: Resources	The resources made available for statistical activities should be sufficient to meet the requirements of this Code and should be used efficiently and effectively.
Principle 8: Frankness and accessibility	Official statistics, accompanied by full and frank commentary, should be readily accessible to all users.

(UK Statistics Authority 2009)

For further detail see: http://intranet/Images/code-of-practice-for-official-statistics_tcm67-70155.pdf

The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 397–408

OFFICIAL STATISTICS ON PERSONAL WELL-BEING: SOME REFLECTIONS ON THE DEVELOPMENT AND USE OF SUBJECTIVE WELL-BEING MEASURES IN THE UK

Paul Allin¹

ABSTRACT

This paper draws on experience of the UK Office for National Statistics (ONS) programme to measure national well-being, particularly the high-profile element of the programme in which subjective well-being measures have been collected and published since April 2011. We reflect on drivers of the ONS work and on how these have given rise to interest both in national well-being – the Beyond GDP agenda – and in the use of subjective well-being measures (self-reported, personal well-being) in public policy. Although we touch briefly on measurement and analysis issues, we mainly concentrate on user requirements, the international context, and political, policy, public and business use of well-being data.

Key words: beyond GDP, measuring national well-being, subjective well-being, user requirements, well-being and policy, cost-benefit analysis.

1. Introduction

When British writer John Berger met the Brazilian photographer Sebastião Salgado he found someone who had trained as an economist and one day “asked himself whether pictures might not reveal as much or more than statistics” (Berger, 2013, p169). Salgado photographed people in many different parts of the world, including Rwanda, Mozambique and Kosovo, and concluded that what he saw was “not the proper way” for humans to live, that “we have a responsibility in the time we are living in to provoke a discussion, to provoke a debate, to ask questions. A debate everybody should participate in and have a responsibility for. If we want to survive as a species we must find a proper direction to go, we must choose another way” (Berger, 2013, p.176).

¹ Imperial College London. E-mail: p.allin@imperial.ac.uk.

The challenge to those working in official statistics is to produce statistical ‘pictures’ that are recognisable, trusted and relevant. Moreover these statistical pictures should contribute to – and perhaps even stimulate – debate, discussion and questions about the state of our nations, and about the sustainability of our current well-being into future generations.

In the words of the United Nations Fundamental Principles for *Official Statistics*, “*Official statistics provide an indispensable element in the information system of a democratic society, serving the government, the economy and the public with **data about the economic, demographic, social and environmental situation**. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information*” (from <http://unstats.un.org/unsd/methods/statorg/FP-English.htm>, downloaded 7th January 2015, emphasis added).

In many countries there is indeed a plethora of official statistics on which users could draw for measures or indicators of well-being and progress. Collections such as *Social Trends* in the UK were a legacy of the social indicators movement of the 1960s. One aim then was to develop social statistics, which “had long tended to drag behind economic statistics in priority and quality” (Moser, 2000). The more recent rise of sustainable development indicators was another way in which progress was to be assessed, using official statistics to describe the economy of a country, its demographic profile and many aspects of society and of the natural environment.

However, the issue remains that, in the view of many people, we are mismeasuring our lives by continuing to select from the range of statistics to focus on economic data, and on the headline measure of Gross Domestic Product (GDP) in particular. (For two publications that helped prompt UK developments discussed in this paper, see Commission of the European Communities, 2009, and Stiglitz et al., 2010). There is a need to sum up how our nation is doing, and how the EU or the World as a whole is doing, more broadly than by whether or not GDP is increasing. There are many calls for action to change the way in which we assess well-being and progress. The call is to go ‘Beyond GDP’ and to live our lives taking account not only of economic performance but also in terms of social progress and with reference to the state of the natural environment, both now and for future generations.

There are currently many such wider measurement initiatives around the world. David Hand and the author have counted over 200 national, regional and local measures of well-being and progress (listed at Allin and Hand, 2014, p.258). Only one of these, the UK measurement of national well-being, is discussed in this paper but a disjoint that we observe in the UK - between producing measures

and using measures - appears to be widespread. Perhaps it is still early days, but there are only a few examples of countries in which politicians, policy makers, businesses and the population appear to have moved away from a focus on GDP growth as the way in which they direct their well-being and progress.

In this paper we briefly summarise the UK work before turning to look at why the programme was established, including the emerging need for measures of subjective well-being. We then look at the use to be made of the measures that are now produced, including a section on well-being and policy, before offering a few, tentative conclusions.

2. Measuring national well-being in the UK

In November 2010, the UK Prime Minister and the UK National Statistician launched the Measuring National Well-being Programme of the Office for National Statistics (ONS). Prime Minister David Cameron said *“From April next year we will start measuring our progress as a country not just by how our economy is growing, but by how our lives are improving, not just by our standard of living, but by our quality of life”* (Cameron, 2010). The ONS work aims to produce a series of measures, reflecting ‘what matters’ to people, although the concept of national well-being is not explicitly defined. Rather, the measures are meant to provide a broader assessment of ‘how the country is doing’ than economic statistics, especially GDP, provide (see the many outputs of the programme, and on-going statistical presentations, at <http://www.ons.gov.uk/ons/guide-method/user-guidance/well-being/index.html> downloaded on 2 February 2015).

Personal, or subjective, well-being is at the heart of the ONS framework of domains, indicators and dimensions for measuring national well-being. Lucy Tinkler, in a paper in this volume, explores the development of the ONS subjective well-being measures, including data collection methods and considerations of how to present subjective well-being data, along with an overview of findings to date (Tinkler, 2015). She makes the point that this is still work in progress and she notes that user engagement has been key to the development of the ONS subjective well-being statistics.

Although the media sometimes characterise the ONS work as ‘Mr Cameron’s Happiness Index’, the measures of national well-being include much more than subjective well-being (which is itself more than whether or not you were happy yesterday). The framework for measuring national well-being was constructed following a ‘national debate’ – a consultation exercise that far exceeded most consultations of the users of statistics (summarised in Tinkler, 2015). The debate

“showed that the well-being of the individual is central to an understanding of national well-being. There are a number of factors that are thought to particularly influence individual well-being and so should be included in providing a picture of well-being in the UK” (ONS, 2011, Introduction). These factors, which we summarise in the following paragraph, reflect the core tenet of the ONS work, that “individual well-being is best understood by relating it to areas that directly affect overall individual well-being, as well as to more contextual domains that are important but contribute less directly to individual well-being. The overall effect of these different factors varies for different individuals, raising important considerations for analysis and policy beyond looking at individual well-being ... therefore ... equality, fairness and sustainability issues are part of national well-being measures” (ONS, 2011, Introduction).

To measure national well-being, the ONS therefore not only draws on its new measures of subjective well-being but also on many other statistics. These include greater analysis of the national economic accounts, especially to understand household income, expenditure and wealth, and further accounts linked to the national accounts, including the UK Environmental Accounts, valuing household production and 'human capital'. Other domains address what are sometimes called 'quality of life' issues - mental and physical health, relationships, personal finance, education and skills, what people do and where they live. In addition to seeing the economy as a context for measuring national well-being, the ONS framework also recognises the natural environment and governance (involvement in democracy and trust in how the country is run) as part of the context. The well-being of children and young people is also seen as part of national well-being. With such a wide range, there is clearly a need for the headline indicators to summarise national well-being and the progress we are making as a society. The ONS has so far resisted calls to produce a single index of national well-being, in addition to the picture painted by the forty or so indicators that feature in its on-line summaries of national well-being. And although personal well-being is at the heart of national well-being, the personal well-being measures are not presented as *the* overall measure: there is at present no single measure of UK national well-being.

3. Drivers for UK work on measuring national well-being

As we have mentioned, the ONS programme was created to respond to increasing interest in measures of well-being and progress that go beyond GDP. GDP was designed as a headline measure of economic performance but has also been taken as measure of well-being, social progress and even sustainability. The

reasons it does not measure those things, and how they could be measured, have been well documented in a number of places, including by the Commission on the Measurement of Economic Performance and Social Progress (Stiglitz et al., 2010). Note that the Stiglitz Commission's report includes measuring the state of the environment and sustainable development, in addition to covering the economy and quality of life as suggested by its title.

National statistical offices were among the many audiences that the Commission wished to address. In their case, the Commission was aiming to direct them to "areas where further developments might be particularly valuable" (Stiglitz et al., 2010, p.7). In particular, the Commission concluded that measures of both objective and subjective well-being were needed to understand people's quality of life, recommending that statistical offices "should incorporate questions to capture people's life-evaluations, hedonic experiences and priorities" in their surveys (Stiglitz et al., 2010, p.18). This is one of twelve main recommendations for new measures (or for making better use of existing data). The Commission was echoing the Istanbul Declaration of 2007 in which international organisations, including the Organisation for Economic Cooperation and Development and the European Commission (the UK belongs to both of these), committed to "measuring and fostering the progress of societies in all their dimensions and to supporting initiatives at the country level" (<http://www.oecd.org/site/worldforum/49130123.pdf> accessed 7 February 2015). The declaration urged statistical offices, in cooperation with others, "to work alongside representatives of their communities to produce high-quality, facts-based information that can be used by all of society to form a shared view of societal well-being and its evolution over time".

There were various ideas in play here. There was the specific view that GDP was not a good metric, exemplified by the difference between aggregate economic growth and how people saw their own condition. Some have claimed that this contributed to a wider loss of trust in government and institutions. There were also views that GDP growth was the means to the goal of well-being, not the overall goal of public policy, and others felt that GDP growth itself was not viable. So, despite the fact that the output of the ONS, like other national statistical offices, was considerably more than just the GDP figures, it was difficult to avoid the conclusion that something needed to be done to change what is counted as the measure of national well-being and progress, with the expectation that this would change what counts. There was also some unravelling to be done, to separate GDP as a valid statistical measure from GDP as a policy goal. If ONS is to meet the full set of requirements envisaged in the UN Fundamental Principles, then it must provide statistics that enable a number of different narratives about well-being and progress, and what to do about them. It

may be a subtle change of wording, but the work came to be seen as ‘GDP and Beyond’, to allow for these different readings.

This all was (and is) quite a challenge. There was clearly a need to draw together existing data, and to develop new measures, that overcome the shortcomings of using GDP as the sole measure of well-being and progress. Using the language of the Stiglitz Commission, this covered ‘classical GDP issues’, such as looking at the distribution of income, consumption and wealth. It also related to aspects of quality of life that matter to people beyond income, and again how this is distributed across population groups. Last but not least, the work needed to address the environment and the sustainability of well-being, for example the extent to which current activity is drawing down the stock of natural resources.

It was also not just about better measurement, but crucially about ensuring that new measures are used. This is not a task only for national statistical offices, but what comes out of the Stiglitz Commission report, like the Istanbul Declaration, is that statistical offices should work in conjunction with others on getting the new measures used.

In a public statement shortly after the Istanbul Declaration, ONS announced that it was starting to analyse “societal wellbeing” and emphasised the need “to understand more fully the requirements for measures beyond GDP” (Allin, 2007, p.46). At that stage, much of the discussion about ‘beyond GDP’ seemed aspirational, with “little detail of how new measures would be used, and what would be done differently”, including the role of existing sustainable development indicators published by Defra, the government department then responsible for sustainability (Allin and Hand, forthcoming).

In parallel with this, it was also apparent that there was policy interest in personal well-being, not just in health (increasing bracketed with well-being) but in other domains. The then head of the UK civil service suggested that “In applying sophisticated approaches to measuring the impact on people’s quality of life through cost-benefit analysis and impact assessments, the UK leads the world in the systematic application of analysis to developing policy. From health to social care, from education to preventing crime, policy methods are developed using methods which draw in the best evidence and analytical thinking from across government and academia” (O’Donnell, 2010). This added weight to the need for personal well-being measures to be regularly collected in ONS’s national surveys, and in ways in which these could be associated with subjective well-being assessments carried out in specific policy areas.

The election for a new UK government in 2010 gave the opportunity to see what the main political parties were saying about well-being and about quality of life as a purpose for government (see Allin and Hand, 2014, Chapter 7). Not

everyone in the political sphere was signed up to this. The Institute of Economic Affairs, for example, remained unconvinced of the role of ‘happiness’ in economics and public policy (Johns and Ormerod, 2007, p.14). The overall conclusion, though, was that ONS needed to take forward the measurement of national well-being, including subjective well-being. Working within the incoming UK government, ONS secured funding for a programme of work and the Budget Statement of June 2010 recorded that “The Government is committed to developing broader indicators of well-being and sustainability, with work currently underway to review how the Stiglitz, Sen and Fitoussi report should affect the sustainability and well-being indicators collected by Defra, and with the ONS and the Cabinet Office leading work on taking forward the report’s agenda across the UK” (HM Treasury, 2010, p.10).

We have not listed in detail all of the drivers and specific initiatives that led to the ONS programme of work on measuring national well-being, but we can summarise in terms of the UN Fundamental Principles for Official Statistics referred to above. First, the relevance of official statistics: the programme addresses a range of emerging requirements for data about the economic, demographic, social and environmental situation, and how this situation is changing. There are political and policy requirements to be met, as well as wider aspirations for alternative measures and data that support ‘new economics’, as well as the standard formulation of national accounts. In short, we must measure what matters. Next, the UN principles call for international coordination and statistical co-operation between countries, and we have pointed to several such initiatives above (and there are more, such as work co-ordinated by Eurostat, the statistical office of the European Union). We should also note that the principles require official statistics to be delivered cost-effectively, which in this case included recognising that the Stiglitz Commission recommendations were a good fit against existing developments and statistical resources at the ONS. There was no need to come up with another set of recommendations and the ONS could focus on implementing all the Stiglitz Commission recommendations. It could also, as Tinkler (2015) reports, draw on academic work on measuring subjective well-being in developing subjective well-being questions suitable for its surveys and for the requirements it had identified.

4. Uses of UK measures of subjective well-being

In this section we mainly discuss subjective well-being. This is not to ignore the other dimensions of national well-being but rather to recognise that subjective well-being measurement was a new departure for ONS. Also, subjective well-

being is at the heart of the ONS framework for measuring national well-being. It has also been suggested (though not by ONS) as *the* alternative measure of well-being to GDP. Layard (2005, p.234) for example, concluded he could “think of no nobler goal than to pursue the greatest happiness of all – each person counting”, although Boarini et al. (2006, p.6) found that “there are several distinct domains – such as joblessness, family and community ties – that contribute to overall life-satisfaction and their influence cannot be reduced to a single dimension of economic resources”. Other claims made for subjective well-being measures are that they reflect each respondent’s individual weighting of the influence of different domains and events, with their own priorities, and that they give a voice to ‘the silent majority’, not otherwise heard through official statistics.

The policy use of subjective well-being arises as an alternative to the long-standing assumption that individuals tend to act to maximise their utility (i.e. well-being within these terms), which is a concept related to their consumption and their income. Under this approach, identification of new policy, selecting between policy options and evaluating the effectiveness of policy interventions rests solely on economic costs and benefits (in shorthand, the impact on GDP). However, “this misses out so much of what makes life worth living” (O’Donnell et al., 2014, p.13) and one answer is to attempt to maximise subjective well-being instead, because this allows for individuals to make a much fuller assessment of their well-being. Subjective well-being measures are therefore needed, not only to include it in cost-benefit analysis but also to see overall how individuals make choices (which can lead, but beyond the scope of this paper, to a role for public policy to tackle behavioural ‘failures’, when individuals make choices that might not be in their, or society’s, long-term interest, see O’Donnell et al., 2014, p.13).

Following the availability of robust data on subjective well-being, a number of ways are being developed to incorporate these data into cost-benefit analysis. One of these is to monetarise the observed subjective well-being effects. Under this approach, the value attributed to a person’s change in well-being arising from a specified policy intervention is given by the change in their income that would produce the same change in subjective well-being (e.g. Fujiwara et al., 2014, p.7). This approach is not without criticism. O’Brien (2010, p.7) commented on this emerging methodology and observed that “The relationship between income and wellbeing is still not fully understood. The method needs more research before its findings will rival or replace existing forms of economic valuation”. However, the All-Party Parliamentary Group on Wellbeing Economics (APPGWE) has now recommended that “New policy should be routinely assessed for its impacts on well-being”, using well-being analysis in making the case for spending, setting priorities and evaluating impacts (APPGWE, 2014, p.3 and p.7). The APPGWE also indicated the breadth of possible applications of well-being in public policy.

Their report addresses well-being and the labour market, planning (including of housing, jobs and the local environments), health, education, arts and culture.

A final point to make is that we have been considering how well-being should be taken into account in politics and in policy. However, recalling that subjective well-being is framed in terms of national well-being should remind us that businesses, civic society, households and individuals are also involved (if they so wish) in taking account of well-being. The many calls to change how we measure our lives referred to earlier all see the issue as more than just making better measurements: they expect that we will make use of the measures, including by changing behaviours and patterns of consumption. As Dolan (2014, p.189) has put it, you might want “to listen more to your real feelings of happiness than to your reflections on how happy you think you are or ought to be”.

5. Concluding comments

We appear to have reached the stage where the publication of robust measures of well-being, including subjective well-being, by national statistical offices and international organisations is becoming accepted as part of their regular outputs. This is not the case everywhere and, where there are new measures, there are inevitably technical developments still under way (for example on the measurement of various capital and resource stocks, which are increasingly seen as a way of understanding current well-being and future sustainability). Well-being – or equivalent words in other languages – is not always used to describe the issue: notably, the Organisation for Economic Cooperation and Development refers instead to Better Lives, because it believes that helping people live better lives will in turn contribute to economic and social well-being. But the point is that well-being can be measured, and increasingly is being measured.

The key question is then, how are new measures of well-being and progress, including subjective well-being, to be used? The UN Fundamental Principles referred to earlier anticipate that official statistics are designed to be used: the principles and supporting material offer guidance on how national statistical offices should gather user requirements. Such procedures work well when requirements for national statistics can be clearly stated, for example for a measure of inflation that reflects consumers’ experience of changes in the price of goods and services. However, the requirement for measures of well-being and progress is less tangible. In essence, the call is for measures that help shape the picture of society, both now and in the future, and which direct policy makers and others to where action is needed, or where different routes can be taken for how we live our lives.

The ‘bottom line’ is that well-being can give a very different picture from one constructed only with economic considerations in mind. Looking at that picture should aim public policy towards improving the well-being of people, not just focussing on national income and economic growth. This is not just for politicians and government policy-makers. We can all support and take part in the process, whether we are in business, civic society or acting as individuals, by engaging with the new ways of measurement and by using these new measures to build better lives for all of us, and for the generations who will follow us.

Acknowledgements

As the former director of the ONS Measuring National Well-being programme, I thank colleagues there for their advice and expertise. (Views expressed in this paper are mine and should not be attributed to ONS). I am also grateful for my collaboration with David J. Hand at Imperial College London, and for my participation in the series of seminars on the politics of wellbeing, funded by the UK Economic and Social Research Council and organised by Ian Bache of the University of Sheffield.

REFERENCES

- ALLIN, P., (2007). Measuring societal wellbeing. *Economic & Labour Market Review*. Office for National Statistics. Vol. 1(10), 46–52.
- ALLIN, P., HAND, D. J., (2014). *The Wellbeing of Nations: Meaning, motive and measurement*. John Wiley & Sons, Ltd., Chichester.
- ALLIN, P., HAND, D. J., (forthcoming) *New Statistics for Old? - Measuring the Wellbeing of the UK*, submitted to *Journal of the Royal Statistical Society (Series A)*.
- APPGWE, (2014). *Wellbeing in four policy areas*. Report by the All-Party Parliamentary Group on Wellbeing Economics at http://b.3cdn.net/nefoundation/ccdf9782b6d8700f7c_lcm6i2ed7.pdf downloaded 13 January 2015.
- BERGER, J., (2013). *Understanding a Photograph*, Penguin Books, London.
- BOARINI, R., JOHANSSON, A., D'ERCOLE, M. M., (2006). *Alternative Measures of Well-Being*. OECD Social, Employment and Migration Working Paper DELSA/ELSA/WD/SEM(2006)2, OECD, Paris.
- CAMERON, D., (2010). PM's speech on wellbeing at <https://www.gov.uk/government/speeches/pm-speech-on-wellbeing> downloaded 7 January 2015.
- COMMISSION OF THE EUROPEAN COMMUNITIES, (2009). *Communication from the Commission to the Council and the European Parliament: GDP and beyond, Measuring progress in a changing world*. COM(2009) 433 final, CEC, Brussels.
- DOLAN, P., (2014). *Happiness by Design*. Penguin Books, UK.
- FUJIWARA, D., KUDRNA, L., DOLAN, P., (2014). *Quantifying and Valuing the Wellbeing Impacts of Culture and Sport*, at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/304899/Quantifying_and_valuing_the_wellbeing_impacts_of_sport_and_culture.pdf downloaded 13 January 2015.
- HM TREASURY, (2010). *Budget 2010 (HC61)*, The Stationery Office, London.
- JOHNS, H., ORMEROD, P., (2007). *Happiness, Economics and Public Policy*. The Institute of Economic Affairs, London.
- LAYARD, R., (2005). *Happiness: Lessons from a New Science*. Penguin Books, London.
- MOSER, C., (2000). *Foreword in Social Trends 30*, Office for National Statistics, London.

- ONS, (2011). Measuring National Well-being: A discussion paper on domains and measures at <http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/discussion-paper-on-domains-and-measures/measuring-national-well-being---discussion-paper-on-domains-and-measures.html#tab-Introduction> downloaded on 2 February 2015.
- O'BRIEN, D., (2010). Measuring the value of culture: a report to the Department for Culture Media and Sport, at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/77933/measuring-the-value-culture-report.pdf downloaded 13 January 2015.
- O'DONNELL, G., (2010). Introduction to Wellbeing Conference, 25 November 2010, HM Government.
- O'DONNELL, G., DEATON, A., DURAND, M., HALPERN, D., LAYARD, R., (2014). Wellbeing and Policy. Report commissioned by the Legatum Institute, London.
- STIGLITZ, J. E., SEN, S., FITOUSSI, J-P., (2010). Mismeasuring our lives: why GDP doesn't add up. The New Press, New York.
- TINKLER, L., (2015). The Office for National Statistics Experience of Collecting and Measuring Subjective Well-Being, *Statistics in Transition – new series*, Vol. 16, No. 3, pp. 373–396.

The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 409–428

EXPLOITING ORDINAL DATA FOR SUBJECTIVE WELL-BEING EVALUATION

Marco Fattore¹, Filomena Maggino², Alberto Arcagni³

ABSTRACT

The evaluation of subjective well-being, and of similar issues related to quality of life, is usually addressed through composite indicators or counting procedures. This leads to inconsistencies and inefficiency in the treatment of ordinal data that, in turn, affect the quality of information provided to scholars and to policy-makers. In this paper we take a different path and prove that the evaluation of multidimensional ordinal well-being can be addressed in an effective and consistent way, using the theory of partially ordered sets. We first show that the proper evaluation space of well-being is the partially ordered set of achievement profiles and that its structure depends upon the importance assigned to well-being attributes. We then describe how evaluation can be performed extracting information out of the evaluation space, respecting the ordinal nature of data and producing synthetic indicators without attribute aggregation. An application to subjective well-being in Italy illustrates the procedure.

Key words: subjective well-being, multidimensional ordinal data, partial order.

1. Introduction

The aim of the paper is to show how the evaluation of subjective well-being can be addressed in a consistent and effective way, using tools from partial order theory and overcoming the limitations of composite and counting paradigms. The topic is valuable for two main reasons. In a “beyond GDP” perspective, the measurement of multidimensional well-being is progressively gaining importance for social scientists and policy-makers. In particular, the evaluation of subjective well-being proves relevant, since it makes clear that personal satisfaction is not just a matter of “objective achievements” and reveals that quality of life eludes

¹ Department of Statistics and Quantitative Methods, University of Milan – Bicocca, Italy.
E-mail: marco.fattore@unimib.it.

² Department of Statistics, Informatics, Applications "G. Parenti" (DiSIA), University of Florence, Italy. E-mail: filomena.maggino@unifi.it.

³ Department of Statistics and Quantitative Methods, University of Milan – Bicocca, Italy.
E-mail: alberto.arcagni@unimib.it.

reductionist approaches. At the same time, available evaluation procedures are not suitable for well-being assessments; they are in fact mainly designed to deal with numerical variables and not with ordinal attributes, so common in well-being studies. Answering the needs of social scientists and policy-makers thus requires developing new statistical strategies, namely alternative procedures to exploit the information power of multidimensional ordinal data. Starting from this consideration, a new evaluation procedure has been recently proposed by the Authors (Fattore, Brueggemann and Owsinski, 2011; Fattore, Maggino and Colombo, 2012; Fattore and Maggino, 2015), capable to deal directly with ordinal data and to compute synthetic indicators without attribute aggregation. The procedure follows the classical steps of any evaluation process in the social field, namely identifying well-being attributes, choosing a well-being threshold and computing individual and overall statistics. All of these steps, however, are accomplished in “purely ordinal terms”, without introducing any artificial transformation of ordinal degrees into numerical scores. Here, we extend the procedure introducing a formal and consistent way to insert into computations exogenous information pertaining to the relevance of well-being attributes. The main focus of the paper is on the concept of evaluation space, which is not simply the set of selected evaluation dimensions, but more profoundly the mathematical structure determining which kind of computations can be consistently performed on the data. It is shown that the natural evaluation space of well-being is a partially ordered set, whose specific structure depends upon the different relevance assigned to attributes. Once the proper evaluation space is set, it is outlined how the evaluation process proceeds and which overall indicators may be computed.

The paper is organized as follows. In Section 2 we briefly discuss the statistical problem of evaluation in a multidimensional setting, mainly focusing on the composite indicator approach and on the Counting Approach of Alkire and Foster. In Section 3, we introduce partial order theory, discuss the construction of the evaluation space and how to account for attribute relevance. In Section 4 we sketch the evaluation procedure. In Section 5, we apply the evaluation procedure to real data pertaining to subjective well-being in Italy, comparing different attribute relevance patterns. Section 6 concludes. The aim of the paper is mainly methodological, so artificial and real examples are meant to introduce the procedure, rather than to provide a deep study of subjective well-being. At the same time, examples aim at showing how the methodology may be, at least in principle, straightforwardly applied, despite its not trivial mathematical foundations. The mathematics involved in the evaluation procedure may seem abstract, at first. Indeed, it is the “mathematics of order”, which is algebraic and combinatorial in nature. Working out all of the technical details behind the evaluation procedure would take too much space, so we describe it in a synthetic and rather informal way (other information can be found in cited references). Only the technical tools employed in accounting for attribute relevance are formally detailed, being the first time they are presented to scholars.

2. Evaluation in multidimensional systems of ordinal attributes

The central role of multidimensional ordinal data in current social studies neatly emerges, when inspecting the structure of national and international social surveys pertaining to well-being and quality of life. Consistently with a beyond GDP perspective, most of questionnaires' items is devoted to aspects of personal or familiar life that can be meaningfully described only in yes/no or ordinal terms. This poses statistical problems in data synthesis and, more specifically, in evaluation studies, since usual dimension reduction tools prove scarcely effective, if not inconsistent.

2.1. The composite indicator approach

Under more or less sophisticated forms, the main road to synthesis in social evaluation is the computation of composite indicators. These may be built using simple weighted averages, directly computed on variable scores, or can be obtained as outcomes of more complex procedures, e.g. of structural equation models under reflective or formative schemes, as is the case. Independently of the adopted statistical tools, an aggregative-compensative approach is followed and synthesis is basically achieved through weighted sums of original variable scores. Although not pertaining to well-being, an instructive and prototypical example of the difficulties involved in the composite indicator approach is provided by the Regional Competitiveness Index (RCI), developed by the Joint Research Centre (Annoni and Dijkstra, 2013). With the aim of producing a ranking of the economic attractiveness of 262 regional European areas, 73 different and heterogeneous indicators are aggregated in a single index. The aggregation proceeds by steps: first 11 so-called "sub-pillars" are built; these are in turn aggregated into three "pillars" (Basic pillar, Efficiency pillar and Innovation pillar); finally, RCI is computed as a weighted average of regional pillar scores (see Fattore, Arcagni and Barberis, 2014, for a graphical scheme of the aggregation process). Although the aim of comparing European regions may be sensible, it is legitimate to ask which kind of information is really gained by such a mixing of different dimensions and whether the resulting European attractiveness map is a faithful reproduction of reality. In its essence, the key problem is that socio-economic issues like well-being (or territorial attractiveness, in the case of RCI) are inherently multidimensional and complex. The complexity of a concept refers to the impossibility of capturing it through compensative approaches, ultimately based on dimensional reduction tools. A complex concept comprises many different dimensions, logically related, but not necessarily statistically correlated. Indeed, a major problem in social evaluation is that dimension reduction approaches often fail, since "logical components" of the concept are not enough correlated and thus cannot be "composed" in a satisfactory way. In the case of multidimensional well-being, the inadequacy of aggregative-compensative procedures is further made evident by the nature of the

attributes involved in the evaluation exercise. Well-being attributes are naturally expressed on ordinal scales, ruling out any consistent way to manipulate them using classical statistical analysis. The “vector space” approach (i.e. considering attributes as vectors that can be summed and multiplied by scalars) is useless and inconsistent. Even attempts to scale ordinal attributes into numerical variables are not satisfactory. If a concept is conceived in ordinal terms (e.g. naturally expressed through adverbial forms), why is one to force it into numbers? In addition, scaling procedures may lead to arbitrary and counterintuitive results (Madden, 2010), revealing that ordinal degrees should not be naively seen as rough manifestations of underlying truly continuous scores.

So we are left with an apparently unsolvable problem. On the one hand, we must produce synthetic views of well-being, out of complex systems of multidimensional attributes; on the other hand, aggregative procedures cannot be employed, for both conceptual and technical issues. Partial order theory provides a way out to this problem. Before showing how this is achieved, however, we briefly discuss the Counting Approach of Alkire and Foster (Alkire and Foster, 2011), which tries a different approach to evaluation and which has been gaining an increasing popularity, since its appearance in 2007.

2.2. The Alkire-Foster counting approach

The Counting Approach has been originally designed for deprivation measurement, but actually it provides a general framework for multidimensional evaluation studies and can be consistently applied to systems of ordinal data. It has the merit to realize that ordinal attributes cannot be handled like numerical variables and to provide a procedure that does not introduce any scaling tools. However, the Counting Approach still sticks to an aggregative paradigm and achieves consistency in the treatment of ordinal attributes at the cost of dichotomizing them, losing a great deal of information. With reference to well-being, the Counting Approach is composed of two main steps: (i) the **identification step**, where satisfied or dissatisfied (it depends upon the focus of the study) individuals are identified and (ii) the **measurement step**, where aggregate satisfaction/dissatisfaction indicators are computed. Suppose the focus is on dissatisfaction. Dissatisfied individuals are identified based on their achievements on a set of k ordinal satisfaction variables v_1, \dots, v_k . The so-called **identification function** is computed through a **dual cutoff procedure**. First, a set of k dissatisfaction thresholds c_1, \dots, c_k is exogenously selected: individuals whose achievement on the i -th attribute is equal to or less than c_i are classified as dissatisfied on v_i . Then an overall cutoff c is defined: individuals whose number of “dissatisfactions” equals or exceeds c are classified as definitely dissatisfied. The Alkire-Foster identification function is thus a 0-1 function, classifying individuals as either non-dissatisfied or dissatisfied in a crisp way. Once dissatisfied individuals have been identified, three aggregate measures can be computed. Let n be the number of individuals in the population and m the number

of dissatisfied, identified through the dual cutoff procedure. The **Head Count Ratio** H is defined as $H = m/n$, i.e. as the fraction of dissatisfied individuals within the population. The **Average Deprivation Share** is instead the average fraction of “dissatisfactions” suffered by dissatisfied individuals, in formulas, $A = T/mk$ where T is the sum of the number of “dissatisfactions” over dissatisfied units. H and A can be combined together, as $M = HA = T/nk$. M can be interpreted as the share of “dissatisfactions” over the maximum number of possible “dissatisfactions” in the population (which is achieved when all individuals are dissatisfied on all of the k attributes).

As clear by the above description, the Counting Approach reduces to computing the number of dimensions an individual is dissatisfied upon, declaring it globally dissatisfied, if this number is equal or exceeds a pre-determined threshold. In this respect, it follows an aggregative approach, where summing over attributes is made equivalent to counting dissatisfactions, by virtue of the dichotomization process. Correspondingly, it also appears that the Counting Approach is not a truly “ordinal procedure”, since ordinal information is not effectively exploited, being lost in simpler binary classifications. This leads to a crisp view of personal satisfaction/dissatisfaction, whereas well-being has a vague nature (Sen, 1992, pp. 48-49), usually accounted for by means of fuzzy set theory. Moreover, when applied to multidimensional ordinal data (which is the case of interest here), there is no natural way to introduce attribute relevance into computations. In summary, the Counting Approach does not provide a completely satisfactory solution to multidimensional ordinal evaluation, being inefficient in the treatment of ordinal data and conceptually similar to the composite indicator paradigm.

3. Partial orders in social evaluation

Any multidimensional evaluation study, and the well-being case is no exception, starts by identifying the so-called **evaluation space**, i.e. the set of dimensions against which evaluation is performed. For example, in the application illustrated later in the paper, the evaluation space comprises four subjective well-being dimensions pertaining to satisfaction on personal health, personal economic situation, familiar relationships and leisure time. Not only the input space settles the conceptual framework of the evaluation process but, even more important, it also determines which information can be extracted from the data and how. What turns a set into a “space”, in fact, is the mathematical structure put on it, which defines the set of meaningful operations that can be performed on its elements. In the Counting Approach, the focus is on attributes and the structure underlying the evaluation space is implicitly that of a vector space, with the drawbacks and inconsistencies already mentioned. Which is, instead, the proper structure of the input space associated with subjective well-being (and with multidimensional ordinal data in general)? Given k well-being

attributes v_1, \dots, v_k , with each statistical unit a **well-being** (or **achievement**) **profile** (i.e. a sequence of ordinal scores on the well-being attributes) is naturally associated. Two profiles \mathbf{p} and \mathbf{q} may be ordered in terms of well-being, when achievements of (say) \mathbf{p} are not worse than those of \mathbf{q} and at least one of them is better. In this case, \mathbf{p} is “better than” \mathbf{q} , written $\mathbf{q} < \mathbf{p}$, and the two profiles are said to be **comparable**. On the contrary, if \mathbf{p} and \mathbf{q} have conflicting scores, i.e. if \mathbf{p} is better than \mathbf{q} on some attributes and is worse on others, than they cannot be ordered and are said to be **incomparable** (written $\mathbf{p} \parallel \mathbf{q}$). So some pairs of achievement profiles may be ordered, others cannot. The input space for well-being evaluation is therefore the set of well-being profiles, naturally structured as a **partially ordered set**. Notice that the focus here is not on ordinal attributes, but on multidimensional profiles, which are the entities that actually characterize individual well-being. Correspondingly, profiles are seen as elements of a “partial order space” and not as elements of a vector space. Although partially ordered sets may seem rather “poor” mathematical structures, they in fact prove very powerful, in view of evaluation. To show this, we must first introduce some basic notions of partial order theory.

3.1. Elements of partial order theory

Given a set Q , a **partial order** \leq on Q is a **reflexive**, **antisymmetric** and **transitive** binary relation defined on it (Davey and Priestley, 2002); the pair (Q, \leq) is called a **partially ordered set** or a **poset**, for short. Let Π be the set of well-being profiles and let $\mathbf{p} = (p_1, \dots, p_k)$ and $\mathbf{q} = (q_1, \dots, q_k)$ be two elements of Π . The set of profiles with partial order defined by

$$\mathbf{q} \leq \mathbf{p} \text{ if and only if } q_i \leq p_i \text{ for each } i = 1, \dots, k \quad (1)$$

is called the **basic achievement poset**. Definition (1) is a purely mathematical condition, identifying the minimal set of comparabilities “anyone would agree upon”. This is why Π is qualified as “basic”. Posets defined on finite sets may be conveniently depicted *via* **Hasse diagrams**, a kind of directed acyclic graph written according to the following two rules: (i) if $\mathbf{q} \leq \mathbf{p}$, then the node corresponding to \mathbf{q} is put below the node corresponding to \mathbf{p} and (ii) if \mathbf{p} **covers** \mathbf{q} (i.e. if $\mathbf{q} \leq \mathbf{s} \leq \mathbf{p}$ implies either $\mathbf{q} = \mathbf{s}$ or $\mathbf{p} = \mathbf{s}$), then an edge is drawn from \mathbf{p} to \mathbf{q} . A subset of Π whose elements are mutually comparable is called a **chain**. On the contrary, a subset of mutually incomparable elements is called an **antichain**. A partially ordered set which is also a chain is called a **complete** (or **linear**) order. If each variable is conceived as a linear order over the set of its degrees, the basic achievement poset is the so-called **product order** of v_1, \dots, v_k , i.e. it is the partial order defined by (1) over the Cartesian product of (the degrees of) v_1, \dots, v_k . A subset D of Π such that “if \mathbf{p} is in D and $\mathbf{q} < \mathbf{p}$, then \mathbf{q} belongs to D ” is called a **down-set**. Analogously, a subset U of Π such that “if \mathbf{p} is in U and $\mathbf{p} < \mathbf{q}$, then \mathbf{q} belongs to U ” is called an **up-set**.

Example. A simple, yet useful, example of poset (the so-called “Cube”) and of its Hasse diagram, is given by the set of all the profiles on three binary attributes v_1 , v_2 and v_3 , partially ordered according to the product order. The attributes may be thought of as expression of satisfaction (1) or dissatisfaction (0) on some well-being dimensions. The poset comprises $2^3 = 8$ profiles and is depicted in Figure 1.

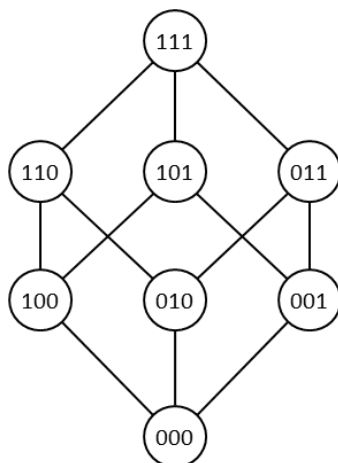


Figure 1. Hasse diagram of the product order of three binary variables (“Cube”). The diagram is to be read from top to bottom; for this reason, directed edges are drawn as simple lines and not as arrows. The subset $\{111, 110, 100\}$ is a chain, while the subset $\{110, 001\}$ is an antichain. The subset $\{011, 101, 100, 010, 001, 000\}$ is a down-set; the subset $\{111, 110, 011, 010\}$ is an up-set.

The link between partially ordered structures and well-being evaluation can be intuitively guessed: any measurement process is a comparison to a benchmark and partially ordered sets are the natural setting to perform multidimensional comparisons. With reference to Example 1, if one identifies profile 110 as dissatisfied, one can immediately conclude that profiles 100, 010 and 000 are dissatisfied as well, since they represent worse situations than 110. Similarly, one can assert that profile 111 is better than 110 and better than all of the other dissatisfied profiles. However, nothing can be said about profiles 101, 011 and 001, which are incomparable with 110. This represents a difference with respect to the Counting Approach: according to the Alkire-Foster procedure, two profiles with the same number of dissatisfactions are equivalent and two profiles with a different number of dissatisfactions may always be compared. The existence of incomparabilities, however, is deeply consistent with the intrinsic multidimensionality and vagueness of well-being (Qizilbash, 2006) and posets do account for them. The way partially ordered sets will be used to derive concrete evaluations of personal well-being, will be outlined later. What is of concern here, is to stress that (i) the input evaluation space is a partially ordered set and (ii) its

structure is the primary source of information pertaining to well-being. In fact, while achievement profiles describe the personal status of individuals, it is the network of comparabilities/incomparabilities they are embedded in, to determine the “social meaning” of their achievement configurations.

3.2. Attribute relevance and the structure of the achievement poset

The definition of the basic achievement poset does not incorporate any information about possible differences in attribute relevance. In real applications, this is not appropriate and attribute relevance should be accounted for (or, at least, the evaluation procedure should provide this opportunity), in order for the input space to better fit the system of social values under investigation. The basic idea is to inject information on attribute relevance into computations, properly modifying the structure of the basic input space. Before showing how to achieve this in practice, we must deepen the mathematical study of the basic achievement poset.

3.2.1. Decomposition of the basic achievement poset by linear extensions

Let (Π, \leq) be the basic achievement poset. An **extension** $\Pi^{\text{ext}} = (\Pi, \leq^{\text{ext}})$ of (Π, \leq) is a poset, such that $\mathbf{p} \leq \mathbf{q}$ implies $\mathbf{p} \leq^{\text{ext}} \mathbf{q}$. In practice, Π^{ext} comprises the same comparabilities of Π and adds some more; in this sense, the set of comparabilities of Π^{ext} extends that of Π . An extension λ which is also a linear order is called a **linear extension** of Π . The set of linear extensions of Π is denoted by $\Omega(\Pi)$. In view of accounting for attribute relevance, we are interested in a particular subset of $\Omega(\Pi)$, namely the set $\text{Lex}(\Pi)$ of linear extensions **lexicographically ordered**.

Definition. Given a permutation π of indices $(1, \dots, k)$, a linear extension λ_π is called lexicographically ordered along π if its order relation \leq_π is defined by (here “ \prec ” refers to the ordering of attribute degrees):

$$\mathbf{q} \leq_\pi \mathbf{p} \text{ if } q_{\pi(1)} < p_{\pi(1)}, \text{ or } q_{\pi(h)} = p_{\pi(h)}, \text{ for } h < s \text{ and } q_{\pi(s)} < p_{\pi(s)} \text{ (} s = 2, \dots, k \text{).}$$

(where $\pi(i)$ stands for the i -th component of the permuted vector of indices). In practice, the sequence of attributes is permuted and profiles are ordered in an “alphabetic” fashion, according to the permutation. Figure 2 depicts an extension, a linear extension and a lexicographic linear extension of the Cube. Linear extensions can be considered as those complete profile rankings that are compatible with the structure of the basic achievement poset, since no comparabilities of Π are violated in them. Clearly, the map L defined by

$$L(v_{\pi(1)}, \dots, v_{\pi(k)}) = \lambda_\pi$$

is a bijection between $\text{Lex}(\Pi)$ and the set of possible permutations of attributes, so that the cardinality of $\text{Lex}(\Pi)$ is $k!$. There is a deep link between $\Omega(\Pi)$ and the

partial order structure of the basic achievement poset, in fact one can reconstruct Π from its linear extensions. More precisely, it can be proved that (see Neggers and Kim, 1998):

Proposition 1. The basic achievement poset Π is the intersection of its linear extensions:

$$\Pi = \cap \Omega(\Pi).$$

Explicitly, this means that the set of comparabilities of Π coincides with all of the comparabilities common to its linear extensions. This fact is of central importance for the evaluation procedure. In view of attribute relevance, however, the following proposition is even more relevant.

Proposition 2. The basic achievement poset Π is the intersection of its lexicographic linear extensions:

$$\Pi = \cap \text{Lex}(\Pi).$$

Proof. Since $\text{Lex}(\Pi)$ is a subset of $\Omega(\Pi)$, then $\cap \Omega(\Pi)$ (which is equal to Π) is a subset of $\cap \text{Lex}(\Pi)$, i.e. $\cap \text{Lex}(\Pi)$ is an extension of Π . On the other hand, if \mathbf{p} and \mathbf{q} are incomparable in Π , then they have at least two conflicting scores in, say, position i and j . Then \mathbf{p} and \mathbf{q} are ordered differently in at least two lexicographic linear extensions whose corresponding permutations differ in the order v_i and v_j are listed. This proves that \mathbf{p} and \mathbf{q} are also incomparable in $\cap \text{Lex}(\Pi)$. Therefore, $\Pi = \cap \text{Lex}(\Pi)$. q.e.d.

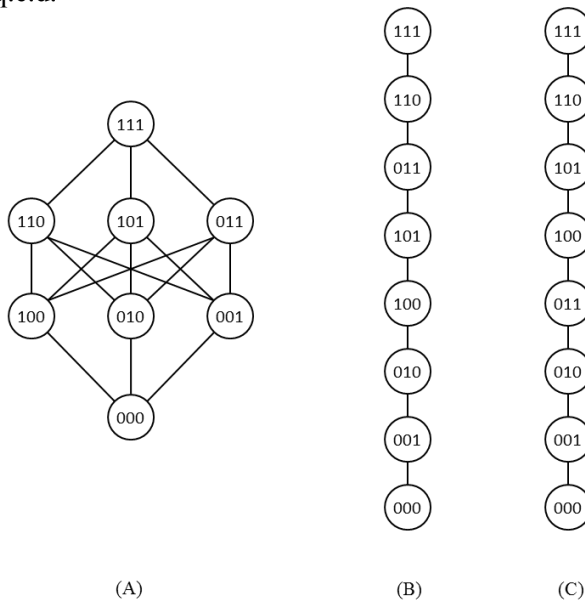


Figure 2. (A) an extension of the Cube; (B) a linear extension of the Cube; (C) a lexicographic linear extension of the Cube (along the identity permutation).

Lexicographic linear extensions are intuitively the “representatives” of attribute rankings in $\Omega(\Pi)$. Since $\text{Lex}(\Pi)$ generates Π by intersection, we see that a link can be made between attribute rankings and the structure of Π . In the next paragraph, we show how this link can be exploited to inject into the structure of the achievement poset exogenous information on attribute relevance.

3.2.2. The attribute poset

To show how lexicographic linear extensions of the achievement poset Π are involved in accounting for attribute relevance, we must first introduce a new poset (Λ, \leq_Λ) on the set Λ of well-being attributes.

Definition. Let v_i and v_j be two well-being attributes, we write $v_i \leq_\Lambda v_j$ if and only if v_j is more relevant than v_i . The set Λ partially ordered by \leq_Λ is called **attribute poset**.

Poset Λ represents a formal yet easy way to define the “relevance pattern” of well-being attributes. It is in fact more natural to assign partial orderings of relevance among attributes, rather than introducing weights (which would also be inconsistent with an ordinal setting), as if one could realistically state a set of precise “attribute equivalences”. In addition, not any partially ordered relevance pattern can be reproduced through weighting schemes. To realize this, consider the attribute poset depicted in Figure 3, on the three attributes of Example 1. According to the Hasse diagram, attribute v_1 is more important than attribute v_2 , but nothing is stated about v_3 . If we were to adopt an equivalent weighting scheme, a greater weight should be attached to v_1 than to v_2 ; but then it is impossible to find a weight to attach to v_3 which makes it “indifferent” to both of the other attributes.

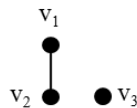


Figure 3. Attribute poset on three attributes. As it can be seen, v_3 is incomparable with both v_1 and v_2 . It is impossible to emulate this relevance pattern attaching weights to each attribute.

As any finite partial order, according to Proposition 1 the attribute poset is equivalent to the set of its linear extensions $\Omega(\Lambda)$, i.e. to the set of attribute rankings, **compatible** with \leq_Λ (or **admissible** with respect to Λ). These linear extensions, in turn, can be seen as attribute permutations. With a little notational abuse, we can thus associate lexicographic linear extensions of Λ to linear extensions of Π through the map L . The image of $\Omega(\Lambda)$ in $\text{Lex}(\Pi)$ through L identifies the set $\text{Lex}(\Pi; \Lambda)$ of lexicographic linear extensions of Π that are consistent with the attribute poset Λ (i.e. **Λ -admissible** lexicographic linear extensions). When Λ is an antichain, which means that attributes share the same relevance, $L(\Omega(\Lambda)) = \text{Lex}(\Pi)$ and all lexicographic linear extensions of Π are Λ -admissible. On the contrary, when Λ is not an antichain, some elements of

$\text{Lex}(\Pi)$ are not Λ -admissible. This implies that Π itself is not consistent with Λ . To restore consistency, Π must be turned into a new poset $\Pi^*(\Lambda) = (\Pi, \leq^*)$ such that $\text{Lex}(\Pi^*(\Lambda)) = L(\Omega(\Lambda))$. The only way to achieve this is to define $\Pi^*(\Lambda)$ as the intersection of all Λ -admissible linear extensions of Π :

$$\Pi^*(\Lambda) = \cap L(\Omega(\Lambda)).$$

$\Pi^*(\Lambda)$, as defined above, is a proper extension of Π if and only if Λ is not an antichain (in that case, it coincides with Π) and it is easily proved to be the smallest extension of Π consistent with Λ .

In summary, we have the following logic chain

$$\Lambda \rightarrow \Omega(\Lambda) \rightarrow \text{Lex}(\Pi; \Lambda) \rightarrow \cap \text{Lex}(\Pi; \Lambda) = \Pi^*(\Lambda)$$

that turns the attribute poset into the desired extension of the achievement poset, incorporating the exogenous information on attribute importance. Figure 4 gives an example of the above chain for the Cube. Two final remarks are in order. First, it is indeed expected that introducing attribute relevance leads to extending the basic achievement poset; additional information conveys in fact new criteria to resolve “comparison ambiguities”, reducing the number of incomparabilities. Secondly, notice that the extension procedure is **purely ordinal**: no numerical weights enters the computations and the nature of the attributes is fully respected. The introduction of attribute relevance concludes the process of evaluation space definition. One can thus proceed to evaluating well-being, through the evaluation procedure briefly described in the next paragraph.

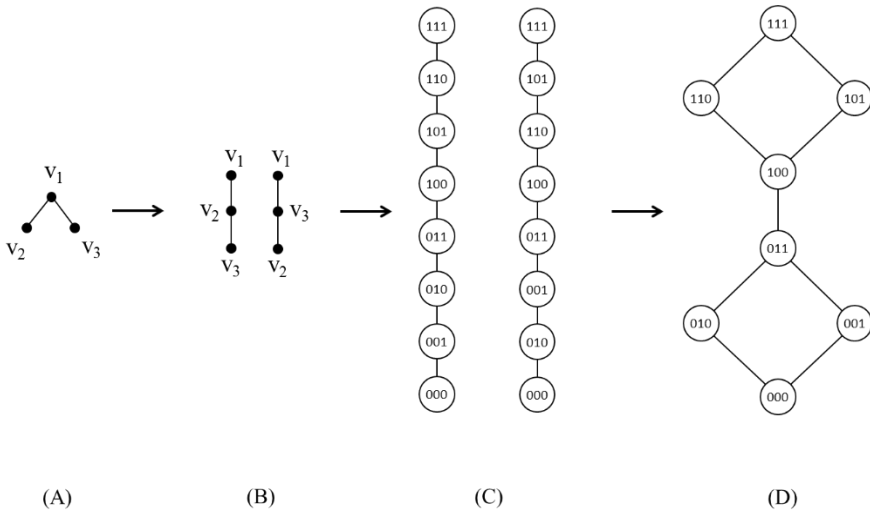


Figure 4. (A) attribute poset Λ on three binary attributes; (B) linear extensions of Λ ; (C) Λ -admissible lexicographic linear extensions of the Cube; (D) extended achievement poset $\Pi^*(\Lambda)$. The Cube has 48 linear extensions; among them, 6 are lexicographically ordered. The selection of the relevance pattern Λ reduces them to two. The final extended achievement poset has 4 linear extensions and comprises only two incomparabilities ($110 \parallel 101$ and $010 \parallel 001$).

4. Evaluating subjective well-being from the achievement poset

Previous paragraph has been devoted to show how the evaluation space of subjective well-being can be properly structured, focusing on achievement profiles and accounting for attribute relevance. Here we show how it can be used to evaluate well-being. Since the fundamentals of the evaluation procedure have been already introduced in other papers (Fattore, Brueggemann and Owsinski 2011; Fattore, Maggino and Colombo, 2012), in the following we limit ourselves to a brief outline.

The aim of the evaluation procedure is to assign subjective well-being scores to statistical units in the population. This will be achieved associating scores directly to profiles of the achievement poset; statistical units then inherit the scores of their profiles. Due to multidimensionality, however, we must distinguish between well-being **vagueness** and **intensity**. With reference to personal dissatisfaction, the former pertains to ambiguities in the classification of a statistical unit as dissatisfied or not; the latter to the severity of dissatisfaction, i.e. intuitively to the “distance” from satisfaction. Both concepts are distinctly accounted for and measured by the evaluation procedure, through a couple of evaluation functions, namely the **identification function** $\text{idn}(\cdot)$, that quantifies the ambiguity of profile classification, and the **severity function** $\text{svr}(\cdot)$, that quantifies dissatisfaction intensity. To build these functions, however, the preliminary identification of a dissatisfaction threshold is needed.

4.1. Setting the dissatisfaction threshold

The achievement poset Π^* conveys no explicit information on subjective well-being. To turn it into a direct input to evaluation, a well-being **threshold** τ must be introduced. The threshold must be conceived as a minimal set of exogenous information, which identifies profiles “on the edge of dissatisfaction”, leaving to the evaluation procedure to spread such information across the poset. Due to multidimensionality, more than one profile may be “on the edge”, so that the threshold must be in general chosen as an antichain of well-being profiles, whose elements describe **alternative dissatisfaction patterns**, to be considered as reference benchmarks. As always in evaluation studies, the choice of the threshold is a delicate step. Given the methodological aim of the paper, here we do not discuss this issue further. Notice, however, that the threshold is directly specified in terms of profiles, without any explicit reference to attribute cut-offs, as in the Counting Approach.

4.2. The identification function

Since there is no natural scale against which to assess subjective well-being, we address identification as a problem of multidimensional comparison between achievement profiles and threshold benchmarks. Differently from the unidimensional case, due to partial ordering not any well-being profile may be

unambiguously compared with the threshold. A profile whose scores are worse than those of an element of the threshold represents a dissatisfied condition (since it is worse than a “dissatisfied profile”). But in many cases, ambiguities arise and some profiles cannot be classified as below or above the threshold, due to conflicting scores. The identification function must account for such ambiguities; to this aim it is defined in such a way that:

- elements of the threshold are scored 1, i.e. they are classified as dissatisfied profiles;
- profiles below an element of the threshold in the achievement poset, i.e. profiles in the down-set generated by the threshold, are similarly scored 1;
- profiles above **any** element of the threshold in the achievement poset, i.e. profiles in the intersection of the up-sets generated by threshold elements, are scored 0 (they are classified as “non-dissatisfied”, since they represent situations that are better than any “completely” dissatisfaction pattern identified in the threshold);
- all other profiles are scored by $\text{idn}(\cdot)$ in $(0,1)$, i.e. they are scored as “ambiguously” or “partly” dissatisfied profiles.

To define the identification function in practice, we start by considering the set of linear extensions of the input achievement poset. In a linear extension, a well-being profile is either below (or coincide with) an element of the threshold or it is above all threshold elements, so that it can be unambiguously identified as “dissatisfied” or as “non-dissatisfied” in **that** linear extension. Therefore, on each linear extension λ one can define a 0-1 identification function $\text{idn}_\lambda(\cdot)$ assigning value 1 to profiles classified as dissatisfied in λ and 0 to all of the others. Linear extensions are thus seen as 0-1 classifiers. In different linear extensions, profiles classified as dissatisfied are in general different (only profiles below elements of the deprivation threshold in the input achievement poset are scored 1 in each linear extension and only profiles above all of the elements of the deprivation threshold in the input achievement poset are always scored 0). Classification ambiguities in Π^* thus reflect in different linear extensions classifying profiles differently. Counting the proportion of linear extensions where a profile is scored 1, one quantifies such ambiguities and gets a non-linear identification function assigning scores in $[0,1]$ to well-being profiles. Formally, $\text{idn}(\cdot)$ is defined as:

$$\text{idn}(\mathbf{p}) = \frac{1}{|\Omega(\Pi^*)|} \sum_{\lambda \in \Omega(\Pi^*)} \text{idn}_\lambda(\mathbf{p}).$$

From the above definition, it follows easily that $\text{idn}(\cdot)$ is **order-preserving**, i.e. if $\mathbf{q} \leq \mathbf{p}$ in Π^* then $\text{idn}(\mathbf{q}) \leq \text{idn}(\mathbf{p})$. Notice that the poset approach to subjective well-being evaluation is, in a sense, a counting approach, but differently from other counting methodologies (Alkire and Foster, 2011, Cerioli and Zani, 1990) it counts over linear extensions and not over attributes.

4.3. The severity function

A faithful picture of subjective well-being requires taking into account both its vagueness and intensity. If intensity would not be assessed, profiles with same identification scores could be considered as equivalent, while they can actually correspond to very different self-perceived situations. Obviously, dissatisfaction intensity, or **severity**, can be meaningfully assessed only on the subset of completely or partly dissatisfied profiles. Analogously to the identification function, for each linear extension and for each dissatisfied profile in it, we define a severity function $\text{svr}_\lambda(\cdot)$ and compute the severity function $\text{svr}(\cdot)$ averaging on $\Omega(\Pi^*)$. Formally, let λ be a linear extension of Π^* and let \mathbf{p} be a profile dissatisfied in it. The satisfied profile \mathbf{q} nearest to \mathbf{p} in λ is the “first” profile ranked above all the elements of the threshold, in λ . Dissatisfaction severity of \mathbf{p} in λ , i.e. $\text{svr}_\lambda(\mathbf{p})$, is thus defined as the graph distance of \mathbf{p} from \mathbf{q} in the Hasse diagram of λ , i.e. as the number of edges between \mathbf{p} and \mathbf{q} ($\text{svr}_\lambda(\mathbf{p})$ is instead set to 0 for non-dissatisfied profiles in λ). Finally, we put:

$$\text{svr}(\mathbf{p}) = \frac{1}{|\Omega(\Pi^*)|} \sum_{\lambda \in \Omega(\Pi^*)} \text{svr}_\lambda(\mathbf{p}).$$

A relative measure of profile dissatisfaction severity can be obtained dividing $\text{svr}(\mathbf{p})$ by its maximum over the achievement poset, i.e. by the severity of the bottom profile of Π^* . Like the identification function, also the severity function is order-preserving.

4.4. Synthetic indicators

Identification and severity scores assigned to profiles are inherited by statistical units. Synthetic indicators may then be computed, averaging on the population or on suitable subpopulations. Three overall indicators are of particular interest.

1. **Population dissatisfaction degree (H)**, defined as the average of $\text{idn}(\cdot)$ over the population.
2. **Specific dissatisfaction degree (D)**, defined as the average of $\text{idn}(\cdot)$ over the subpopulation of completely or partly dissatisfied statistical units (i.e. with profiles \mathbf{p} , such that $\text{idn}(\mathbf{p}) > 0$). D measures the vagueness of dissatisfaction of individuals with a non-null dissatisfaction degree. High values of this indicator may reveal that the population is polarized into two groups, the non-dissatisfied and the (almost) completely dissatisfied.
3. **Population severity degree (S)**, defined as the average of $\text{svr}(\cdot)$ over the subpopulation of statistical units with non-null identification scores. S can be turned into a relative measure S^* , dividing it by its maximum over Π^* .

Other indicators may be indeed defined and computed, starting from the distribution of evaluation scores on the statistical population; nevertheless, H, D and S suffice to provide a synthetic and comprehensive view of subjective well-being. As anticipated in the Introduction, they are computed without attribute aggregation.

4.5. Computational aspects

The evaluation procedure is combinatorial in nature and draws upon the computation of linear extensions of the achievement poset. In real cases, it is unfeasible to list all of them and one must rely on sampling algorithms, computing the evaluation functions on a subset of linear extensions. The sampling procedure and some basic functions to manipulate partial orders and to compute the identification and severity functions together with the overall indicators are implemented in the R package PARSEC (available from the Authors). Details on computational aspects and the use of the package can be found in Fattore and Arcagni (2014).

5. Subjective well-being in Italy

In this section, we provide some examples of the evaluation process on data pertaining to subjective well-being in Italy, for the year 2012. The dataset comes from the “Multipurpose survey about families: aspects of daily life”, held by the Italian National Statistical Bureau on a yearly basis. We focus on satisfaction on Health, personal Economic status, Family relationships and Leisure time. In the original dataset, satisfaction is expressed on a 4-degree scale: 1 – “very”, 2 – “enough”, 3 – “little” and 4 – “not at all”. In the following, scores have been reversed, so that 1 stands for “not at all” and 4 stands for “very”. The achievement poset is composed of 256 profiles and will not be displayed. The original dataset comprises 46464 records; among these, 6893 have missing values. Since they are not systematic and given the exemplificative purpose of this analysis, we have simply deleted them, reducing the input dataset to 39571 records. To provide examples of applications of the procedure, we consider three different attribute relevance patterns, comparing the results in terms of evaluation functions and overall indicators. Computations have been performed using the programming language R and the package PARSEC. The three attribute posets Λ_1 , Λ_2 and Λ_3 are depicted in Figure 5. In the first pattern, all attributes share the same relevance; in the second, Health, Economic status and Family relationships are equivalent and more relevant than Leisure time; in the third, Leisure time is dominated by Economic status which in turn is dominated by Health, while Family relationships is incomparable with all of the other attributes (this is the only pattern not reproducible by numerical weighting schemes). To the three relevance patterns, there correspond three achievement posets Π_1 (coinciding with the basic achievement poset Π), Π_2 and Π_3 . The threshold has been set to $\tau =$

(1223, 2123) (first score refers to Health, second to Economic status, third to Family relationships and fourth to Leisure time). The threshold is not symmetric with respect to attributes, but emphasizes the relevance of Health and Economic status. In principle, there should be consistency between attribute relevance implicit in threshold selection and the definition of the attribute poset. Given the exemplificative purpose of this section, we do not stress this aspect here.

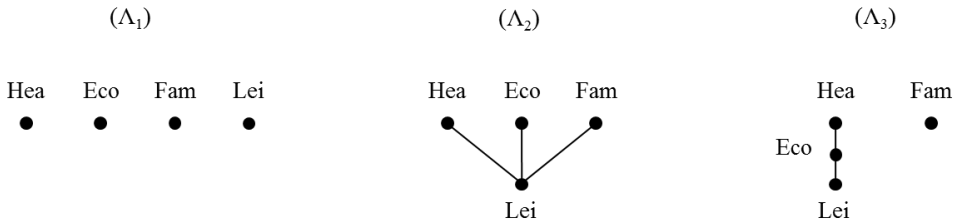


Figure 5. Three alternative attribute posets on Health (Hea), personal Economic status (Eco), Family relationships (Fam) and Leisure time (Lei). Λ_1 , Λ_2 and Λ_3 have $4!=24$, $3!=6$ and 4 linear extensions respectively.

Figure 6 depicts the identification and relative severity scores of the well-being profiles for each input achievement poset. The values of indicators H, D and S for the three cases are reported in Table 1.

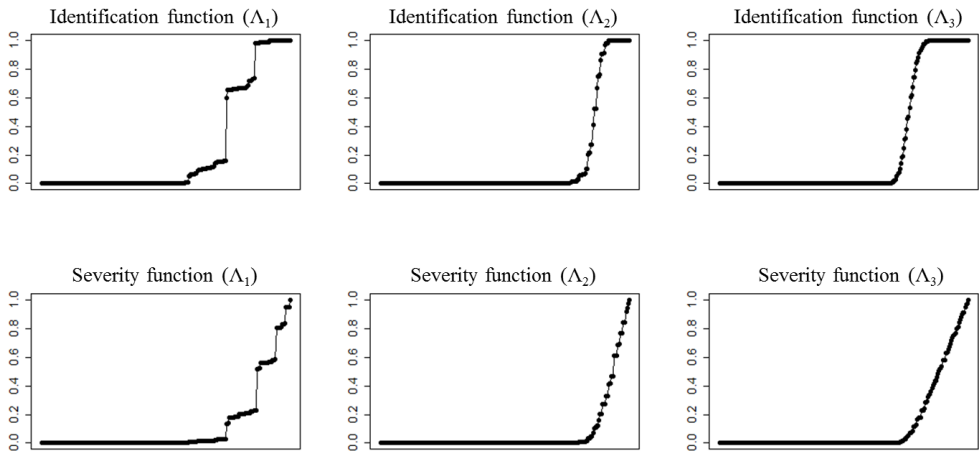


Figure 6. Evaluation functions implied by the attribute posets depicted in Figure 5. Profiles are sorted on the x-axes in increasing order of identification (upper panels) or severity (lower panels) scores (for graphical purposes, the relative severity scores of completely non-dissatisfied profiles have been set to 0, although the severity function is not defined over them).

As it can be seen, the shapes of both evaluation functions strongly depend upon the relevance patterns of the attributes. Interestingly, the identification function is not linear and when the attribute poset comprises comparabilities it assumes a sigmoid shape. The severity functions allow for dissatisfied profiles to be further distinguished in terms of dissatisfaction intensity. As for the identification scores, also these functions are non-linear, but here non-linearities tend to disappear, when the attribute poset increases the number of comparabilities. As the attribute poset gets similar to a linear order, dissatisfaction severity increases, i.e. on average partly or completely dissatisfied statistical units increase their distance from non-dissatisfaction. The Head Count Ratio, on the contrary, has a more complex behaviour, revealing that the two indicators do measure different aspects of deprivation. In fact, H is lower under Λ_1 than under Λ_2 . The Specific deprivation degree, instead, reveals that passing from Λ_1 to Λ_3 , partly/completely dissatisfied units move from non-dissatisfaction towards dissatisfaction, and Italian society would appear as basically polarized into two groups, a bigger one of non-dissatisfied individuals and a smaller one of highly dissatisfied.

Table 1. Overall indicators (expressed on a 0-100 scale) for the patterns of attribute relevance depicted in Figure 5.

Attribute poset	Population dissatisfaction degree (H)	Specific dissatisfaction degree (D)	Population Severity degree (S)
Λ_1	6.7	23.3	10.7
Λ_2	2.7	31.9	15.7
Λ_3	12.1	60.8	23.3

Much more information could be extracted from the data, e.g. considering socio-demographic covariates or comparing subjective well-being patterns at territorial level. What has been reported, however, should be enough to prove the flexibility and the effectiveness of the evaluation procedure. The shapes of the identification function show how the procedure is capable to account for the nuances of subjective well-being, distinguishing among well-being patterns that counting approaches would have scored identically. Sensitivity of final results to the choice of the attribute poset, in turn, shows how it is important to properly specify the evaluation input space, whose structure, together with the threshold, determines the final evaluation scores.

6. Conclusion

In this paper, we have outlined a comprehensive procedure to address evaluation problems in a multidimensional ordinal setting. The procedure follows the general logic of any evaluation study: identification of evaluation dimensions, addition of attribute relevance, threshold selection, computation of evaluation scores at the statistical unit and population level; noticeably, however, all of these steps are defined and performed in a purely ordinal way. Partial order theory, the mathematics of order, allows for that and makes it possible to overcome the inconsistencies of composite indicators and the inefficiency of counting approaches. We have applied the procedure to data pertaining to subjective well-being in Italy for year the 2012, comparing different patterns of attribute relevance, to show how the evaluation process is flexible, simple and straightforward to apply. A major problem with the procedure is indeed its computational burden, that currently limits its application to small or medium size posets (typically, with up to about 250 profiles). Currently, a simplification of the procedure is under development, so as to reduce the number of linear extensions to evaluate, so as to virtually remove any computational issue in real applications. The primary aim of the paper, however, is to show that a sound conceptual and formal setting can be indeed defined, where subjective well-being evaluation and, more generally, evaluation on multidimensional systems of ordinal data can be properly addressed. Hopefully, this may help social scientists to innovate their methodological and statistical toolbox, so as to capture the complexity and nuances of human experiences in a more effective and realistic way.

REFERENCES

- ALKIRE, S., FOSTER, J., (2011). Counting and multidimensional poverty measures. *Journal of Public Economics*, 95, pp. 476–487.
- ANNONI, P., DIJKSTRA, L., (2013). EU Regional Competitiveness Index - RCI 2013. RCI 2013, JRC Scientific and Policy Reports. Luxembourg: Publications Office of the European Union.
- CERIOLI, A., ZANI, S., (1990). A Fuzzy Approach To The Measurement Of Poverty. In: Dagum, C., Zenga, M. *Income and Wealth distribution, Inequality and Poverty*. Berlin Heidelberg: Springer-Verlag. pp. 272–284.
- DAVEY, B. A., PRIESTLEY, B. H., (2002). *Introduction to lattices and order*. Cambridge: Cambridge University Press.
- FATTORE, M., BRUEGGEMANN, R., OWSIŃSKI, J., (2011). Using poset theory to compare fuzzy multidimensional material deprivation across regions. In: Ingrassia S., Rocci R. and Vichi, M. *New Perspectives in Statistical Modeling and Data Analysis*. Berlin: Springer-Verlag.
- FATTORE, M., MAGGINO, F., COLOMBO, E., (2012). From composite indicators to partial orders: evaluating socio-economic phenomena through ordinal data. In: Maggino F., Nuvolati G. *Quality of life in Italy: researches and reflections*. Berlin: Social Indicators Research Series.
- FATTORE, M., ARCAGNI, A., (2014). PARSEC: an R package for poset-based evaluation of multidimensional poverty. In: Bruggemann R., Carlsen L. and Wittmann J. *Multi-Indicator Systems and Modelling in Partial Order*. Berlin: Springer.
- FATTORE, M., ARCAGNI, A., BARBERIS, S., (2014). Visualizing Partially Ordered Sets for Socioeconomic analysis. *Revista Colombiana de Estadística*, 34(2), pp. 437–450.
- FATTORE, M., MAGGINO, F., (2015). A new method for measuring and analyzing suffering – Comparing suffering patterns in Italian society. In: Anderson R. E. *World Suffering and the Quality of Life*. New York: Springer.
- MADDEN, D., (2010). Ordinal and cardinal measures of health inequality: an empirical comparison. *Health Economics*, 19, pp. 243–250.

NEGGERS, J., KIM, S. H., (1998). Basic posets. Singapore: World Scientific.

QIZILBASH, M., (2006). Philosophical Accounts of Vagueness, Fuzzy Poverty Measures and Multidimensionality. In Lemmi A. and Betti G. Fuzzy Set Approach to Multidimensional Poverty Measurement. New York: Springer.

SEN, A., (1992). Inequality reexamined, Harvard University Press.

The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 429–440

USING THE DAY RECONSTRUCTION METHOD TO QUANTIFY TIME SPENT SUFFERING AMONG OLDER ADULTS WITH CHRONIC PAIN

Dylan M. Smith¹

ABSTRACT

The Day Reconstruction Method (DRM) is a structured diary designed to measure time use in a manner that is more valid than traditional written summary measures, but less burdensome than real-time electronic diary methods. The lower respondent burden and administration costs might make it feasible to utilize this method in large national surveys. Past studies using the DRM have generally focused on subjective mood states during different types of activities. In this study, we extended the DRM to also measure suffering from negative symptoms, such as pain and fatigue, in 122 older adults, most of whom suffer from chronic pain. Results indicated that the method was well tolerated in this population, with over 98% of the sample providing interpretable responses. Chronic pain respondents reported spending a considerable proportion of their day suffering from pain, fatigue, and depression; a much higher proportion than reported by healthy controls (p 's < .01). Time spent suffering was associated with lower income, negative mood, and lower life satisfaction and quality of life.

Key words: quality of life measurement, osteoarthritis, time-weighted experience.

1. Introduction

1.1. Overview of time-use measurement

To help inform public policy, there is a growing emphasis on capturing subjective appraisals of emotional experiences, as a supplement to traditional objective economic performance measures such as income, or to epidemiological measures of health such as longevity (Kahneman et al., 2004a, Dolan, 2008, Krueger et al., 2009). In particular, it has been proposed that a better sense of the “well-being” of a nation could be achieved if surveys examined how people spend their time (e.g., time spent at work, or engaged in leisure or social activities) and,

¹ Department of Preventive Medicine, HSC Level 3, Stony Brook University, Stony Brook, NY, USA 11794. E-mail: Dylan.smith@stonybrookmedicine.edu.

importantly, how people value their experiences during those activities. However, numerous concerns have arisen regarding the accuracy of traditional self-report measures that require respondents to remember and summarize their emotional experiences over some period of time, or ask respondents to estimate their general level of satisfaction with or liking for various activities (for reviews, see (Schwarz and Strack, 1999, Diener et al., 1999, Krueger et al., 2009). These concerns have led methodologists to consider ways of capturing subjective experiences that are less reliant on participants' ability to accurately remember subjective states, and/or to aggregate these experiences into a single summary score.

Ecological Momentary Assessment (EMA) refers to a class of methods designed to capture experiences as they occur, in everyday life, and thus avoid both reliance on memory, and context effects caused by artificial environments (e.g. a laboratory). In the most paradigmatic type of EMA, respondents report their current activities and provide subjective assessments of their emotions and experiences in real time, as they go about their daily lives (Shiffman and Stone, 1998). The usual method is for the respondent to wear an electronic device for a period of time, such as a week, that prompts them at various times throughout the day to respond to a brief survey. Answers are input directly into the device. Thus, from EMA data, researchers can compute average levels of variables of interest (thus avoiding the problem of relying on participants to aggregate their own experiences), and also explore peak and diurnal experiences. Data generated from these methods allows researchers to pair affective experiences with specific activities; by also considering time spent in different activities, across the population these data could be used to generate "national time accounts" (Krueger et al., 2009).

However, the cost of this kind of EMA procedure makes it prohibitive to "scale up" this method to the level of nationally representative surveys. Devices must be provided to respondents who must be trained in their use. Given the considerable respondent burden involved, response rates may be low—especially among some vulnerable or distressed groups—and participant compensation costs are likely to be substantial. Response rates may be especially low among people who have the most difficulty using the devices (e.g., those with certain disabilities), although technological innovation may eventually make this less of an issue (Smith, 2011).

1.2. Structured diary approach to subjective assessment

The Day Reconstruction Method (DRM) was devised to assess subjective experiences in a manner that is specifically designed to avoid problems of many recall-based measures of time use while being more affordable and less burdensome than EMA methods (Kahneman et al., 2004b). In the DRM, participants follow a structured format in which they first divide a day into specific 'episodes' or events, and indicate the duration of the episode. They then describe those events in terms of the type of activity (e.g., commuting to work, having a meal, exercising), and provide a detailed rating of their affective state during the activity.

By first asking participants to recall the events of their day, and then provide ratings of those specific experiences, the DRM exploits the fact that while memory for ongoing experiences, like pain and mood, are flawed, memory for discrete events is much better (Robinson and Clore, 2002). Thus, it avoids (or at least reduces) some of the biasing factors noted above, such as the tendency to recall information that is congruent with peak or recent experiences, which are more easily remembered.

The DRM is designed to be self-administered and can be completed by most participants in a single sitting. Thus, it is much less burdensome and costly to field than the most rigorous EMA methods. Further, while the DRM was originally conceptualized as a way to pair activities with affect (positive and negative mood), its authors explicitly designed it to be adaptable to measure any type of subjectively felt experience.

1.3. Using the DRM to measure duration of suffering

In considering using measures of subjective experience in national surveys, it has been suggested that it might be important to focus not only on emotional well-being, but on suffering, with the idea that alleviation of suffering might be a more suitable target for government intervention (Stone and Mackie, 2013). The adaptability of the DRM noted above makes it a potentially good candidate for quantifying suffering (Smith et al., 2008, Krueger and Stone, 2008). In this study, we adapted the DRM to capture experiences of pain and fatigue in a group of older adults, most of whom suffer from chronic knee pain.

Specifically, we tested the feasibility of an online version of the DRM, and computed time-weighted measures of suffering along the dimensions of pain, fatigue, and depression. We examined interrelations among these variables and tested associations with demographics and previously validated measures of subjective well-being. Finally, for demonstrative purposes we paired measures of suffering with specific activities as reported on the DRM to see whether some types of activities were generally associated with more pain, fatigue and depression.

To achieve these goals, we examined data from an observational study of older adults. The study was comprised of two waves of data collection, each involving a four-day observation period, and included numerous measures, including objective performance measures of physical function, Ecological Momentary Assessments (EMA) of pain and mood, and extensive measures of quality of life using traditional survey methods, in addition to the DRM. For this set of analyses, we focus on data from the DRM, and several survey measures, all at wave 1; other results from this study are reported elsewhere (Stone and Mackie, 2013, Smith, 2011).

2. Methods

2.1. Population and sample

124 participants in total, 77 of whom had chronic knee pain (most commonly due to osteoarthritis), and 47 of whom were healthy older adults, were enrolled in this study. Participants were recruited either through an informational letter and follow-up phone call to adults who are part of the Research Participant Program (RPP) at the University of Michigan's Geriatrics Center, or through the University of Michigan Health System's online registry of ongoing research studies. Participants were included if they were aged 50 and older, cognitively intact (i.e. score of ≥ 24 on the Mini Mental Status Exam (Folstein, 1975)), and English-speaking. To be included in the knee pain group, participants had to have at least mild knee pain (score of 5 or greater on the WOMAC pain subscale; (Bellamy et al., 1988)) over the past three months. Participants were excluded if they were non-ambulatory, had medical conditions or problems (other than knee pain) that interfered with daily activity performance or cause pain and fatigue (such as cardiopulmonary problems, neurological conditions, autoimmune diseases), or if they had a joint replacement or surgery of the knee or hip in the previous six months.

Participants were compensated up to \$160 for full participation in the study. Participation was voluntary and participants were free to withdraw from the study at any time without penalty. This research was approved and supervised by the Institutional Review Board at the University of Michigan.

2.2. Procedures

We contacted potentially interested participants by phone for screening, to describe the study, and schedule an initial visit. At each of two time points, six to twelve months apart, data were collected at two separate visits. At the first visit, participants provided informed consent and demographic information, had their cognitive function assessed, and completed measures of medication use, comorbidities, and functional mobility. Participants were then electronically surveyed repeatedly over a four-day observation period to track symptoms and mood quality. On one of these days, chosen at random, participants completed an online version of the DRM (described below). Immediately after this observation period, participants returned for their second lab visit where they received payment and filled out additional questions about the observation period (including quality of life); a small number of participants who did not have internet access completed the DRM at this second lab visit.

2.3. Measures

We programmed an online version of the DRM, an instrument that provides a comprehensive assessment of experience-related affect throughout the day (Kahneman et al., 2004b). This instrument employs the aspects of a time-use study, in that respondents recall episodes of the previous day and construct a diary sequencing these episodes. Respondents then answer questions about the subjective experience associated with these episodes, as well as basic, objective questions about when and where episodes occurred, what specific tasks or activities respondents were engaged in, and with whom they were interacting. Questions probing subjective experience include four affect descriptors (happy, depressed, angry, enjoy) with response scales ranging from “not at all” (0) to “very much” (4).

Developers of the DRM designed it to be broadly adaptable to various types of self-report settings. In addition, the response scales used in the DRM can be easily modified to include domains relevant to chronic pain in older adults; thus, we added relevant items to the list of affect descriptors (pain, fatigue, physically active, physically limited).

At visit 2, positive and negative affect over the previous four days was measured using the positive affect/negative affect schedule (Watson et al., 1988). This instrument includes a list of 20 different feelings and emotions. Survey respondents are asked to indicate to what extent they have felt each of these feelings or emotions during the past week using a scale ranging from “Very slightly or Not at all” (1) to “Extremely” (5). Items include such feelings and emotions as “Interested,” “Excited,” “Nervous,” and “Uncertain about things.”

An overall QOL question asked participants to choose a number between 0 and 100 that best represents their current QOL (0 _ *the worst imaginable QOL*; 100 _ *the best imaginable QOL*). Overall life satisfaction was measured with the Satisfaction with Life Scale (Diener et al., 1985). This instrument comprises statements about respondents’ general feelings and attitudes concerning their life, such as “In most ways my life is close to my ideal” and “So far I have gotten the important things I want in life.” Respondents indicate how much they agree or disagree with these statements on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

3. Results

3.1. Response rate and demographics

Of 123 participants, 122 completed the DRM and provided affect ratings of at least three episodes (75 knee pain and 47 healthy controls). Participants divided

their previous day into an average of 8.4 episodes (ranging from 3 to 16 per participant), thus providing a total of 1043 separate episode ratings of mood and symptoms. Episodes were approximately 90 minutes in length, on average, although they ranged in length from 2 minutes to over 10 hours. The sample was 60%; the average age was 59 years. Participants were female and 71% White, 24% African American.

3.2. Proportion of time in activities while experiencing suffering

We examined “suffering” along three dimensions: pain, fatigue, and depression. To compute time weighted proportion scores for pain, we first identified any episode in which the participant reported experiencing pain. Next, we summed the length of all painful episodes for each participant, and computed the proportion of the length of that episode to the total length of all episodes for that participant (61, or 6% of the episodes had missing or non-interpretable length estimates; for these we utilized mean imputation). The same procedures were used for fatigue and depression. Finally, we computed a composite variable that estimated the proportion of episodes spent suffering at all—that is, from either fatigue, pain, or depression. As these are proportions, the possible range is from 0 to 1.

Table 1 shows the mean, observed range, and standard deviation for each dimension of suffering, and composite suffering for the total sample, and then broken down by group (chronic pain versus healthy control). Chronic pain participants reported spending more of their time in pain, feeling fatigued, and feeling depressed than did health controls. Regression analyses confirmed that these group differences were significant, after controlling for age and gender (all three p 's < .01).

Table 1. Means, standard deviations, and observed ranges of proportion of time spent suffering for the full sample and by group.

	Total sample (n=122)	Chronic pain sample (n=75)	Non-pain sample (n=47)
Proportion of time with pain (0-1)	.60 (.45)	.79 (.30)	.21 (.34)*
Proportion of time with fatigue (0-1)	.61 (.54)	.71 (.36)	.45 (.37)*
Proportion of time with depression (0-1)	.29 (.38)	.38 (.42)	.15 (.07)*
Proportion of time any symptom (0-1)	.76 (.33)	.90 (.23)	.55 (.36)*

Note. The significance tests in the table indicate whether pain participants differed from non-pain participants in regression models that controlled for age and gender.

* $p < .01$.

Next, we examined associations with demographics and previously validated measures of subjective well-being and quality of life. Table 2 depicts these correlations. As shown, time spent in pain or depressed was associated with lower income, lower life satisfaction, more negative affect, and lower overall quality of life. Time spent fatigued was associated with more negative affect, and lower quality of life. Although not depicted in Table 2, we also examined these associations within each group and did not observe notable differences in the patterns.

Table 2. Correlations between different measures of suffering and with demographics and measures of well-being (n=122).

	Proportion pain	Proportion fatigue	Proportion depressed	Age	Income	Negative Mood	Life Sat.	Quality of life
Prop. pain	--							
Prop. fatigue	.59**	--						
Prop. Depressed	-.40**	.44**	--					
Age	-.10	-.12	-.28**	--				
Income	-.26**	-.05	-.28**	.17 ^a	--			
Negative Mood	.35**	.30**	-.43**	-.10	-.19*	--		
Life Sat.	-.29**	-.14	-.45**	.29**	-.51**	-.42**	--	
Quality of Life	-.27**	-.20*	-.43**	.21*	.50**	-.41**	.67**	--

Note: Two participants failed to return for wave 2 and thus did not provide responses for mood, life satisfaction, or quality of life; for correlations involving these variables, n = 120. Three additional participants declined to respond to the income question; for correlations involving this variable, n= 117.

** = p < .01

* = p < .05

a = p < .10

3.3. Suffering by type of activity

Next, for demonstrative purposes, we paired pain and fatigue scores to specific activities listed under each episode. To examine whether a given type of activity was associated with more pain, we subtracted the participants' average pain level from the pain score reported during the activity. That is, for each individual we computed a score indicating whether a given activity was generally

associated with greater or less pain than that generally experienced by that individual (this was done to avoid confounding across participants; i.e., participants with more pain may be less likely to work).

For the total group, pain was highest while at work (average difference score = .08, $n = 50$) and while exercising (average difference score = .23, $n = 48$). Fatigue was highest while watching television (.18, $n = 91$), during childcare (.09, $n = 43$), while exercising (.09), and while relaxing (.13, $n = 96$). When we examined activities within each group, patterns were generally similar.

4. Discussion

4.1. Overview of key findings

These data provide initial support for the idea that a structured diary of time use, the Day Reconstruction Method, can be adapted to measure dimensions of suffering across a day in older adults with chronic illness. The online instrument was well tolerated by participants, 98% of whom were able to complete the instrument and provide affect ratings of at least three episodes. Findings indicated that pain was prevalent. Knee arthritis participants reported pain 85% of the time spent in their episodes throughout the day; however, pain was also reported 21% of the time in episodes reported by healthy older adults. Feeling fatigued was also common, though much more so in the knee pain group. Depression was the least prevalent among the three measures examined, occurring 15% of the time in the healthy group and 38% of the time in the pain group. Across both groups, proportion of time spent suffering was generally associated with lower income, negative affect, and lower life satisfaction and quality of life.

The DRM produces much more detailed information about time spent in different activities than do traditional summary survey measures, and there is some theoretical and empirical basis to assume that this information is more accurate as well. Past research comparing the DRM to the “gold standard” approach of repeated experience sampling has generally found relatively close agreement—closer than that observed using traditional survey methods (Kim et al., 2013, Stone and Mackie, 2013).

For descriptive purposes, we also paired ratings of pain and fatigue with specific types of activities. Few clear patterns emerged; however, this may be a function of the relatively modest sample size (compounded by the fact that not all participants were engaged in all activity types). Nonetheless, we did observe some indication that pain levels were higher when participants were at work and especially while exercising. Given that participants on average spent much more time at work than exercising, work pain may have been a bigger overall component of total time in pain. However, given the sample size we performed no significance test and made no claim regarding the replicability of the pattern we observed. In larger, more representative studies this technique could be quite

useful in identifying the types of activities that most contribute to time spent suffering. The format of the DRM also allows for weighting by the length of activity, further refining its ability to identify activities that contribute the most to the overall experience of pain.

Future studies can also expand upon the general approach pursued here. Additional dimensions of suffering can be examined (e.g., stress and anxiety, other physical symptoms including disability limitations, etc.) in different subpopulations (e.g., other health conditions, different racial/ethnic groups). In addition, the data generated from the DRM allows for other analytic approaches, including weighting by intensity of the experience of pain, for example, or utilizing a more strict criteria to define suffering (e.g., restricting to the highest pain levels, rather than including any pain level).

4.2. Caveats and limitations

The DRM is designed to be more accurate than traditional written summary measures, and it provides estimates of time spent engaged in various activities, and affect experienced during those activities; however, the time data should be interpreted with caution. Obviously, these are estimates based on potentially flawed memories of how the day went, but in addition the format of the DRM makes interpretation of these estimates somewhat ambiguous, for at least two reasons. First, participants are allowed to define how many episodes to split the day into, and they often leave gaps in time. Second, while they may report experiencing a symptom during a given episode, they may not have experienced it for the entire duration of that episode. For these reasons, rather than try to estimate how many minutes each participant reported experiencing a negative symptom, we computed a time-weighted proportion of episodes in which a symptom was experienced. This can be interpreted as how much time the participants spent in episodes with, for example, pain, as a proportion of the total time spent in all episodes the participant felt were sufficiently meaningful to list in reconstructing their day. As such, it leaves open the possibility that suffering occurred at other points in the day not listed as episodes.

Third, while the DRM is much less burdensome and expensive to administer than experience sampling methods using electronic diaries, it still requires nearly 45 minutes for most participants to complete the full instrument. Thus, it may be impractical for use in many national surveys. However, researchers have successfully fielded a brief version of the DRM in at least one large survey study, by having respondents rate only a subsample of the episodes respondents listed (Krueger and Stone, 2008).

Finally, this is a non-representative clinical sample of limited size; thus, we cannot assert that the proportion of time spent in pain observed here will generalize to broader populations. As noted, the small sample size makes it difficult to study how type of activity relates to pain and fatigue.

4.3. Conclusion

Using the Day Reconstruction Method, we observed that some older adults spend a substantial proportion of their time experiencing symptoms such as pain, fatigue, and depression, and that time spent with these symptoms is negatively associated with quality of life. Taken together, these findings highlight the adaptability of the DRM, and demonstrate its potential usefulness for capturing suffering, in addition to emotional well-being.

Acknowledgements

This research was supported by a grant from the National Institute on Aging (R21AG030177). I wish to thank Catherine Hytner, William Lopez, and Rachel Tocco for their assistance in data collection and coding. I also wish to thank Susan Murphy for her contributions to the design of the parent study that made it possible to generate these data.

REFERENCES

- BELLAMY, N., BUCHANAN, W. W., GOLDSMITH, C. H., CAMPBELL, J., STITT, L. W.,(1988). Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *The Journal of Rheumatology*, 15, 1833–1840.
- DIENER, E., EMMONS, R. A., LARSEN, R. J., GRIFFIN, S.,(1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49, 71–75.
- DIENER, E., SUH, E. M., LUCAS, R. E., SMITH, H. L.,(1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, 125, 276–302.
- DOLAN, P.,(2008). Developing methods that really do value the ‘Q’ in the QALY. *Health Economics, Policy and Law*, 3, 69–77.
- FOLSTEIN, M., FOLSTEIN S. E., MCHUGH, P. R., (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*, 12, 189–198.
- KAHNEMAN, D., KRUEGER, A., SCHKADE, D., SCHWARZ, N.,STONE, A. A.,(2004a). Toward National Well-Being Accounts. *AEA Papers and Proceedings*, 94, 429–434.
- KAHNEMAN, D., KRUEGER, A. B., SCHKADE, D., SCHWARZ, N., STONE, A. A.,(2004b). A survey method for characterizing daily life experience: The day reconstruction method (DRM). *Science*, 306, 1776–1780.
- KIM, J., KIKUCHI, H., YAMAMOTO, Y.,(2013). Systematic comparison between ecological momentary assessment and day reconstruction method for fatigue and mood states in healthy adults. *British Journal of Health Psychology*, 18, 155–167.
- KRUEGER, A. B., KAHNEMAN, D., SCHKADE, D., SCHWARZ, N., STONE, A. A.,(2009). National Time Accounting: The Currency of Life. In: KRUEGER, A. B. (ed.) *Measuring the Subjective Well-Being of Nations: National Accounts of Time Use and Well-Being*. Chicago: University of Chicago Press.
- KRUEGER, A. B., STONE, A. A.,(2008). Assessment of Pain: A community-based diary survey in the USA. *The Lancet*, 371, 1519–1525.
- ROBINSON, M. D., CLORE, G. L.,(2002). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology*, 83, 198–215.

- SCHWARZ, N., STRACK, F.,(1999). Reports of subjective well-being: judgmental processes and their methodological implications. In: KAHNEMAN, D., DIENER, E., SCHWARZ, N. (eds.) *Well-Being The foundations of hedonic psychology*. New York.
- SHIFFMAN, S., STONE, A. A.,(1998). Ecological momentary assessment: A new tool for behavior medicine research. In: KRANTS, D. S. (ed.) *Technology and Methods in Behavioral Medicine*. Mahwah: Lawrence Erlbaum Associates.
- SMITH, D. M.,(2011). Measures of time-use and well-being in survey research. *Forum for Health Economics and Policy*, 14, 1–14.
- SMITH, D. M., BROWN, S. L., UBEL, P. A.,(2008). Mispredictions and misrecollections: Challenges for subjective outcome measurement. *Disability and Rehabilitation*, 30, 418–424.
- STONE, A. A., MACKIE, C., (eds.),(2013). *Subjective Well-Being: Measuring Happiness, Suffering, and Other Dimensions of Experience*, Washington, D.C.: The National Academies Press.
- WATSON, D., CLARK, L. A., TELLEGEN, A.,(1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality & Social Psychology*, 54, 1063–1070.

The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 441–460

AN INITIAL RESEARCH ON OUTPUT WELL-BEING INDEX APPLIED TO RESIDENTS IN WEALTHY COUNTIES FROM CHINA

Zhanjun Xing¹, Xixia Qu²

ABSTRACT

Sampling from three wealthy counties in Shandong province ($n = 855$), this research examined the characteristics of an output well-being index, and the index shows good internal consistency reliability and ideal construct validity. It could be used as an instrument to evaluate the quality of life of Chinese citizens. The index was applied to an analysis of the quality of life of the residents from three wealthy counties. It shows that the level of the quality of life of the rural residents is higher than that of the urban residents, but the level of some indicators is imbalanced, and the levels of subjective and objective indicators are consistent. These characteristics of the well-being index of the rural residents from the wealthy counties are closely related with the level of the local economy and social development. It is advised that the output well-being index could be used to evaluate the degree of citizen's need that was met and the degree of human being's all-over development that was realized. It could also be used as an important policy instrument for the policy makers' good governance.

Key words: well-being index, output index, quality of life.

1 Introduction

1.1. Background

The subjective well-being indicator system is the measurement for subjective feelings. It has been much criticized because of a strong sense of subjectivity and difficulties in measuring, although the subjective well-being indicator system has been demonstrated to be valid in a variety of researches. Therefore, in order to find an index which is a real efficient reference point for policy makers, a single subjective indicators system is not adequate enough and objective quality of life

¹ School of Political Science and Public Administration, Shandong University, No. 27 South Shanda Road, Jinan, 250100, People's Republic of China. E-mail: xingzhanjun@163.com.

² School of Public Administration, Shandong University of Finance and Economics, No. 7366 East Erhuan Road, Jinan, 250014, People's Republic of China. E-mail: quxixia@126.com.

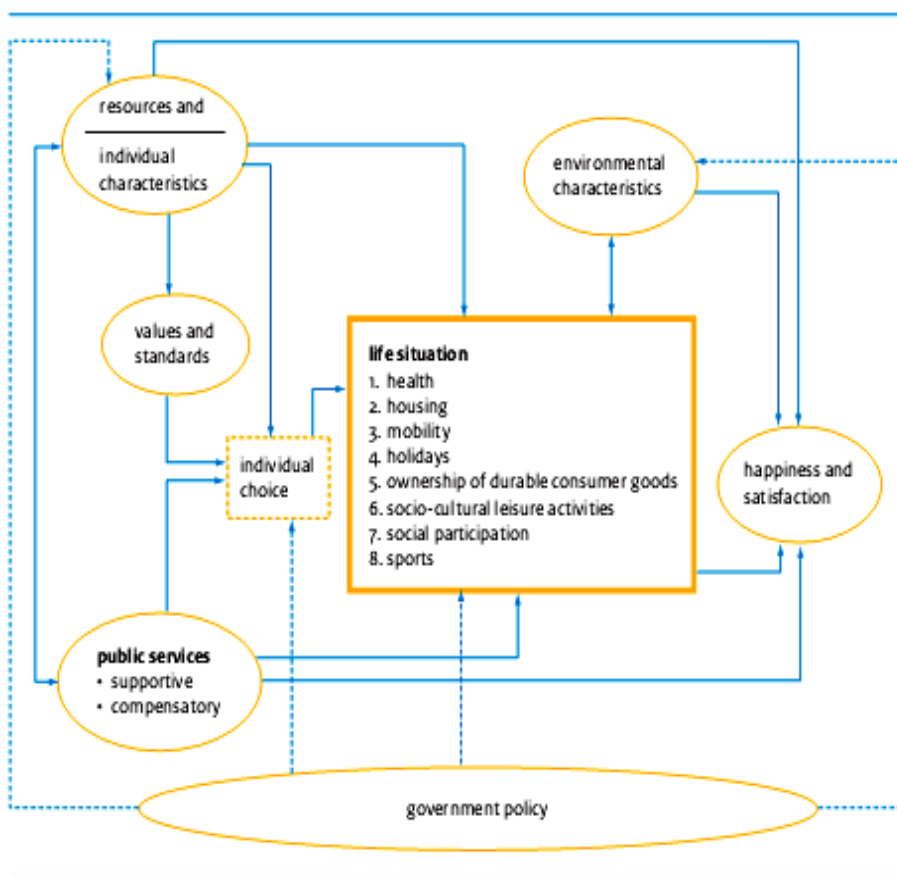
indicators are called for to set up a reasonable comprehensive index system of quality of life. Affected and inspired by researchers working on indicators in China and abroad, we proposed that a set of indicators which are closer to the citizens should be constructed. There have been many objective indicators systems constructed based on official statistical data, of which the imperfection is gradually emerging nowadays. If the first hand data which directly reflects the quality of life condition of Chinese citizens is absent, policy makers' decisions will lose their indispensable pertinence and manoeuvrability, especially to the local government, their specific rules making process will be very difficult.

"Output well-being index" is the key concept which we are talking about. The output well-being index is not a newborn concept; it has a very close relationship with the concepts of "Output Indicators" and "Input Indicators" in statistics. What some researchers in western countries are considering is emphasizing that "Input" could not be converted completely to the real quality of life of the residents, citizens, and the general public practically. That is saying that although many opportunities and conditions are offered, people might not commendably convert these favourable objective advantages into the abilities which could help to improve their quality of life. In this case, focusing on "Output" indicators is a good solution.

On this point of view, Veenhoven (1996), a researcher from the Netherlands working on life indicators, pointed out that we needed to pay attention to "Apparent quality-of-life" in the research, that is the quality of what we could see in our daily life and not just the hypothetical quality of life. Researches on quality of life should stress "flourishing" and "thriving" conditions, which stands for the relationship between ecological environment and self-development of animals and plants: if they already have obtained good environment and conditions to live, could they hence make a very good self-development? The answer is - it depends. So do our citizens. Nowadays, we stress that various kinds of public service systems should be offered to the citizens in China. This is a very important task we are facing, but whether these public service systems could ultimately convert into the real quality of life of our citizens or not is the vital question challenging our government. Noll (2002) from Germany, as a social researcher, proposed that "realized well-being" should take advantage of social opportunities. His views are generally the same as Veenhoven's. Amartya Sen (1993) proposed the concept of "Capabilities Approach" (CA) which attracted many quality of life researchers based on the view of "realized opportunity", which means the opportunities that could be realized instead of the opportunities that are offered. In Sen's Capability Approach, well-being can be defined as the freedom of choice to obtain the things in life that one has reason to value most in his or her personal life. Moreover, he stressed valuable functioning and insisted that estimating quality of life should be based on obtaining the capabilities of valuable functioning.

Life Situation Index constructed by the Social and Cultural Planning Office of the Netherlands offers an idea of an analysis framework. It was designed to monitor the differences and changes based on the concept of output quality of life.

See figure 1 below. There are 8 factors in Life Situation, they are health, housing, mobility, holidays, ownership of durable consumer goods, socio-cultural leisure activities, social participation, and sports. The other factors we thought should be monitored are outside of this framework, and happiness and satisfaction are presented as the controlled conditions. That means happiness is not a part of this Life Situation analysis framework. Beyond that, we also use Gross National Happiness (GNH) of the Kingdom of Bhutan for reference. These important points of view help us formulate our own researches in China, and they offered a very important background for us to make the discussion.



(Boelhouwer J., 2010)

Figure 1. Life Situation Index (Social and Culture Planning Office, SCP, 2008)

1.2. Structure of Output Well-being Index

There are two components that constitute the Output Well-being Index System - Life Situation and Surroundings.

The first component is Life Situation. We borrowed this concept from Life Situation Index of Netherland while giving it a different meaning and structure. Life Situation includes housing, health, sports, leisure, mobility (not the traffic, mobility refers to the travel conditions and the travel abilities, while the traffic belongs to the category of environment or surroundings), social capital, economic capability (ownership of durable consumer goods), and social participation (including political participation, we also designed political indicators in this system).

The second component is Surroundings, including security, traffic, consumption condition (surrounding consumption conditions), community condition (including communal facilities), and livable environment (emphasizing the environment). We consider such two important components as the indicators of Output Well-being Index. See Figure 2 below.

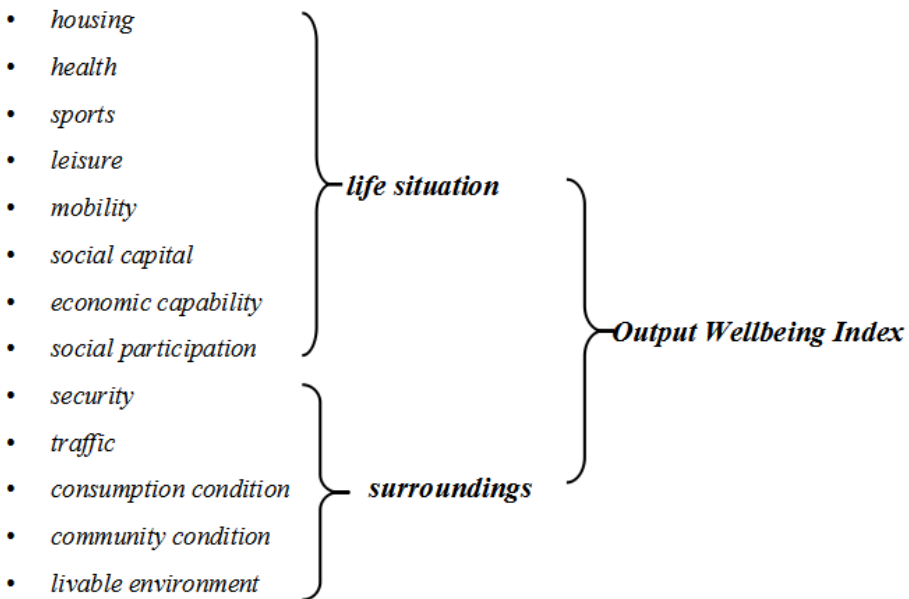


Figure 2. Output Well-being Index structure

1.3. Structure of Subjective Well-being Index

As for the Subjective Well-being Index, we stripped it away from Output Well-being Index because it was used as a detective variable. According to our definition of happiness (subjective well-being), there are three components included in the important item, life satisfaction, pleasure, and self-worth. Among

these factors, life satisfaction is very distinct and explicit, it includes seven different components which are all referring to the satisfaction in people’s life experience, they are satisfaction with housing, satisfaction with health, satisfaction with leisure, satisfaction with relationship, satisfaction with income, satisfaction with traffic, and satisfaction with environment. The seven components are corresponding with Life Situation indexes we have mentioned before. Here we use domain satisfaction instead of using overall or general satisfaction. There are two items to reflect pleasure, which is the second component of Subjective Well-being Index. The two items could stand for pleasure very well. These two items are picked from Subjective Well-being Scale of Chinese Citizens which was set up before. The third component of Subjective Well-being Index is self-worth, we also picked four items which could reflect self-worth well from our existing Subjective Well-being Scale of Chinese Citizens. Life satisfaction, pleasure, and self-worth constitute Subjective Well-being Index. See figure 3 below.

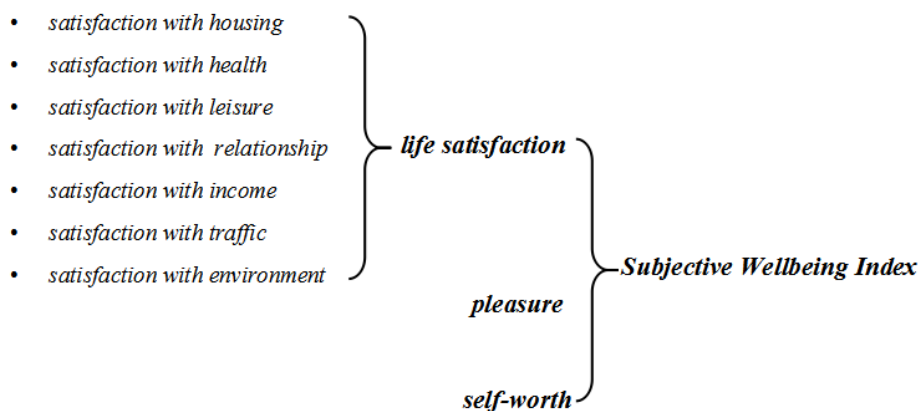


Figure 3. Subjective Well-being Index structure

2. Methods

2.1. Research objectives

The specific aims of the research are to (a) Examine the measurement properties of output well-being index. (b) Investigate the differences in quality of life (QOL) between town and rural residents. Here, we did not refer to urban but county town, which is not representative of all types of cities, therefore some of the conclusions are certainly different. In fact, the main difference between town and village residents is the census register. This means that the residents are distinguished mainly by the census register. (c) Explore the relationship between subjective and objective well-being indicators.

2.2. Participants and procedure

The investigation was carried out in May 2011. The sample was selected through the method of quota sampling because we did not consider generalizing the conclusions to the overall population and planned to do a preliminary study. In this case, according to the soliciting opinion from the experts of related disciplines, three wealthy counties in Shandong province, Zhangqiu, Shouguang and Laizhou were selected. From the historical and developmental point of view, the selected counties are quite characteristic in China. The private economy develops fast in Laizhou; Shouguang abounds in vegetables, it is honoured as the Town of Vegetables in China; Zhangqiu has developed very fast in recent years. In the past three years these three counties were monitored always within the top 50 in the comprehensive competitiveness ranking of the top 100, carried out by Chinese academy of social sciences. The geographical locations of these three counties are the east and middle area in Shandong province.

The final effective sample size is 855 people. The summary of the demographic and socioeconomic characteristics of the participants were as follows. For the gender, male, 51.5%; female, 48.5%. For the age, 24 and below, 9.7% participants; 25–34, 20.5% participants; 35–44, 25.1% participants; 45–54, 20.8% participants; 55 and above, 23.8% participants. For the education background, junior middle school educational level and below, 41% participants; senior middle school, 27.3% participants; college and university, 16.3% participants; postgraduate, 15.6% participants. For the place of residence, county town, 57.4% participants; rural area, 42.6% participants. See table 1 below.

Table 1. Demographic characteristics of participants ($n = 855$)

Variable	Meaning of variable	Frequency	Valid Percent (%)
Gender	Male	436	51.5
	Female	410	48.5
Age	24 and below	83	9.7
	25—34	175	20.5
	35—44	214	25.1
	45—54	177	20.8
	55 and above	203	23.8
Education	Junior and below	350	41.0
	Senior	232	27.2
	College	139	16.3
	Postgraduate	133	15.6
Place of residence	County town	491	57.4
	Rural	364	42.6

2.3. Measures

2.3.1. Measuring instrument

Output Well-being Index (OWI) and Subjective Well-being Index (SWI) constitute our measuring instrument. As we mentioned above, Output Well-being Index includes two domains and 13 indicators in total. Subjective Well-being Index (SWI) includes 3 domains and 11 indicators.

2.3.2. The measurement properties of OWI and SWI

The measurement properties of OWI and SWI based on investigation sample were as follows.

2.3.2.1. Reliability analysis

Using the data collected from this sample, we examined the internal consistency reliability of OWI and SWI. It has been shown that they both had favourable reliability. The Cronbach's alpha coefficient of OWI is 0.77, and the Cronbach's alpha coefficient of SWI is 0.85.

2.3.2.2. Structural validity analysis

Based on our former theoretical hypothesis and empirical exploration, we put forward OWI model and SWI model. OWI model consists of two basic components. The first one is life situation (including housing, health, sports, leisure, mobility, social capital, economic capability and social participation), and the second one is surroundings (including security, traffic, consumption condition, community condition and livable environment). According to the view of experience, subjective well-being could be regarded as an organic whole composed of life satisfaction, pleasure and self-worth. The two models were verified by the sample data. See figure 4, figure 5, and table 2 below.

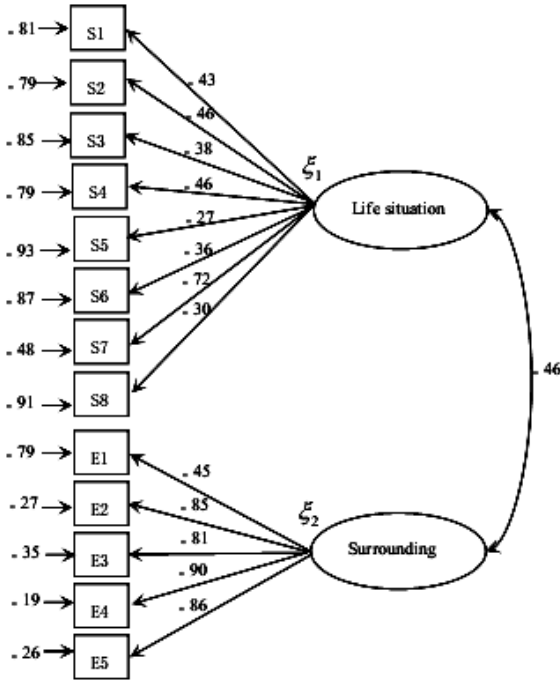


Figure 4. Estimates of OWI's parameters

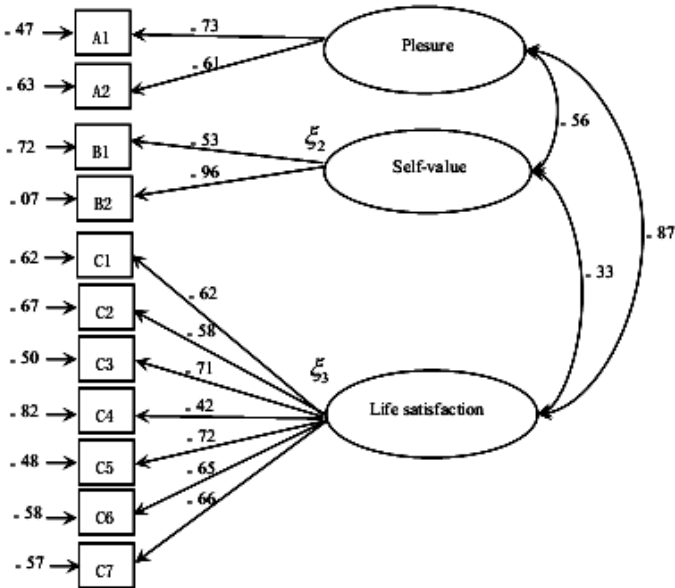


Figure 5. Estimates of SWI's parameters

Table 2. Goodness-of-fit statistics

	χ^2/df	GFI	AGFI	NFI	NNFI	CFI	IFI	RFI	RMR
OWI	1.42	0.93	0.90	0.92	0.93	0.94	0.94	0.91	0.08
SWI	5.14	0.93	0.88	0.94	0.93	0.95	0.95	0.91	0.06

3. Results

3.1. Differences in Life Situation Index between town and rural residents

Table 3 presented mean scores on Life Situation Index according to the place of residence. It showed that the rural residents scored higher than the town residents on housing, health, sports, and economic capability; there was little difference between town residents and rural residents on leisure and mobility; and the town residents group scored higher than the rural residents group on social capital and social participation. See table 3 below.

Table 3. Differences in Life Situation Index by place of residence

Life Situation Index	Place of residence	Mean	Std. Deviation
Housing	town	55.0772	22.01873
	rural	60.9040	9.94929
Health	town	65.9147	19.87696
	rural	69.5018	17.28095
Sports	town	66.3421	36.34438
	rural	73.1145	33.73265
Leisure	town	30.1009	26.37962
	rural	24.1216	24.71030
Mobility	town	50.4131	27.44961
	rural	39.9008	26.15156
Social capital	town	68.0556	18.20047
	rural	67.7794	17.29996
Economic capability	town	46.8474	29.61923
	rural	52.8307	26.80025
Social participation	town	25.3138	22.87733
	rural	25.0939	25.36471

3.2. Differences in Surrounding Index between town and rural residents

Table 4 presented mean scores on Surrounding Index according to the place of residence. It showed that the rural residents group scored higher than the town residents group on security, traffic, consumption condition, community condition, and livable environment.

Table 4. Differences in Surrounding Index by place of residence

Surrounding Index	Place of residence	Mean	Std. Deviation
Security	town	82.3410	17.83378
	rural	89.5248	12.75951
Traffic	town	60.8178	21.22735
	rural	73.7748	19.09448
Consumption condition	town	56.5824	21.07834
	rural	70.8854	16.58042
Community condition	town	64.0067	18.88507
	rural	76.2818	17.31514
Livable environment	town	57.3840	20.59211
	rural	71.6593	17.37469

3.3. Differences in Life Satisfaction Index between town and rural residents

Table 5 presented mean scores on Life Satisfaction Index according to the two groups of town residents and rural residents. It showed that for all of the indicators the rural residents group scored higher than the town residents group.

Table 5. Differences in Life Satisfaction Index by place of residence

Life Satisfaction Index	Place of residence	Mean	Std. Deviation
Satisfaction with housing	town	58.7755	29.00057
	rural	72.4100	23.21418
Satisfaction with health	town	64.2437	24.78657
	rural	69.6812	23.77614
Satisfaction with leisure	town	49.8361	26.29742
	rural	65.4444	24.46810

Table 5. Differences in Life Satisfaction Index by place of residence (cont.)

Life Satisfaction Index	Place of residence	Mean	Std. Deviation
Satisfaction with relationship	town	80.7803	14.45181
	rural	85.5220	14.95639
Satisfaction with income	town	44.5643	28.26804
	rural	58.0282	26.88224
Satisfaction with traffic	town	54.2562	27.46849
	rural	71.5254	23.08847
Satisfaction with environment	town	63.6667	25.15699
	rural	76.2319	18.13238

3.4. Differences in Well-being Index between town and rural residents

Table 6 presented the differences in Well-being Index between town and rural residents. According to the mean scores, the rural residents group scored higher than the town residents group for all of the indexes belonging to Well-being Index.

Table 6. Differences in Well-being Index by place of residence

Well-being Index	Place of residence	Mean	Std. Deviation
Life situation Index	town	49.7042	19.86471
	rural	52.5267	15.78220
Surrounding Index	town	61.8097	17.33926
	rural	76.6356	14.11514
Pleasure Index	town	60.8903	21.40875
	rural	69.7465	18.20141
Self-worth Index	town	67.4948	20.97562
	rural	70.1681	20.02736
Life satisfaction Index	town	53.6271	19.53198
	rural	68.5069	15.80865
Output Well-being Index	town	53.5221	19.03231
	rural	65.7426	12.11949
Subjective Well-being Index	town	56.4214	18.74110
	rural	66.3760	14.54758

4. Discussion

The measuring results of Output Well-being Index of town residents and rural residents above showed that rural residents presented a clearly superior level to town residents. This result is almost completely opposite to the result which was already published in the *Journal of Society* in 2006. Why is there such a difference?

4.1. Sampling strategies

Firstly, it might be affected by the sampling strategy. As mentioned above, random sampling was not used as the main sampling strategy while using the quota sampling. Quota sampling usually is used under condition that a researcher understands certain features about the overall population, and the sample size is large. It has specific advantages such as lower cost, easy to be carried out, and being qualified to meet the requirements of the population proportion. But quota sampling often covers up and hides the deviation factors that cannot be ignored. Therefore, the influence of sampling strategy will be given careful consideration in the following research stage, and a more rigorous sampling process will be taken out combining with interviewing method to further investigate this conclusion.

4.2. Income difference and quality of life

Secondly, we are considering the fact that the result might be affected by the economic conditions of the sample counties. Urban-rural income difference has an important effect on quality of life. It is usually considered to reflect the nature of negative public goods, because the larger the income difference gap the greater the negative impact people are suffering, and the lower the happiness level. Hagerty (2000) pointed that the range and skew of the income distribution in a community affected a person's happiness, as predicted by range-frequency theory, and decreasing the skew (inequality) of the income distribution in a country increases average national SWB. His studies strongly supported social comparison effects of income within a community. Graham (2006) found that inequality has negative effects on happiness in Latin America, where it seems to be a signal of persistent unfairness. Morawetz et al. (1977) found that the more unequal the income distribution the lower the individual's self-rated happiness. Takashi Oshio etc. (2010) found that individuals who lived in areas of high income inequality tended to report themselves as being less happy, even after controlling for various individual and area-level factors. According to their results the association between inequality and happiness was modestly significant,

regardless of the choice of covariates at an individual level, and stronger at a lower level of perceived happiness. Although the above studies were discussing subjective well-being from the general population without grouping urban-rural residents and without including the objective indicators of quality of life, they still illustrate some of the problems, at least the relationship between urban-rural income difference gap and subjective indicators of quality of life indeed exist, which also demonstrates the interpretation of our result is in the right direction.

Table 7. Descriptive statistics on the quality of life (monthly figures in €) (Thiess Buettner & Alexander Ebertz, 2009)

Sub-sample	Mean	S.D.	Min	Max
Rural counties (West)	170	22.7	120	245
Urban counties (West)	159	24.7	76	230
Rural counties (East)	126	12.5	98	175
Urban counties (East)	124	18.1	90	158

Table 7 summarizes the results of the research of Thiess Buettner & Alexander Ebertz (2009) on quality of life index for each of the four groups of German regions. It shows that rural counties scored higher than urban counties in both West and East Germany on quality of life index. This is the same as our result, rural residents scored higher than county town residents in Objective and Subjective quality of life.

Peng Wang (2011) reported the influence of income difference on subjective well-being in China taking the year of 2006 from CGSS database as an example. According to his study, the influence of income difference on subjective well-being presented an inverted U-shape curve. The critical point was Gini coefficient amounting to 0.4. Subjective well-being of the residents was constantly increasing as income difference gap was widening when Gini coefficient was less than 0.4; but if Gini coefficient was greater than 0.4, the widening income difference gap would lead to decrease in subjective well-being of the residents. With the enlargement of the income gap, subjective well-being presented lower level especially in the population group of city residents, non-agricultural household register residents, and the residents of higher education level. This result is opposite to the result we discussed, which implies a smaller gap between urban and rural income difference would lead to higher quality of life of rural residents. The specific situation is as follows.

The three sample places are all wealthy counties, and what is the most different from the others is the small urban-rural difference. Take the year of 2012 as example, Urban per capita disposable income (25755CNY) is 2.7 times higher

than Rural per capita net income (9446CNY) from the perspective of Shandong province, while the urban-rural income difference of these three sample counties (Laizhou, Shouguang, Zhangqiu) is obviously smaller than the average level of the entire province – (by about twice - see table 8 below). This may indicate some problems and at least at the present stage there are no identical conclusions about this topic in China and abroad because of the different research approach and techniques.

Table 8. The provincial and three counties' urban-rural difference (CNY)

	Urban per capita disposable income	Rural per capita net income
Laizhou	29485	14387
Shouguang	23130	17963
Zhangqiu	23130	13587
Shandong Province	25755	9446

4.3. Well-being and Deprivation state

Finally, we consider the existence of happiness chasm within a group as an important factor for discussion, which magnified the difference between urban and rural residents. According to Zapf's (1987) opinion, when living conditions are combined with subjective evaluations and differentiated only in terms of "good" or "bad", a 2×2 table results, which distinguished between four "welfare levels" or categories of quality of life. The combination of good/good is termed "well-being", the combination bad/bad is called "deprivation", and the two mixed responses, "dissonance" or "adaptation". (See table 9). The quality of life in a life domain is considered to be the higher the more citizens are found on the level of "well-being". The "deprived" constitute the classic target group of social policy. The "dissonants" represent the potential for protest and change. This group is also described as being in a dilemma of dissatisfaction. Those in the "adaptation" category frequently represent the reality of powerlessness and social withdrawal.

The distribution of Well-being and Deprivation states of urban and rural residents is obviously asymmetric with respect to our result presented in table 10. Does it imply existence of certain kind of a gap which leads to our result in hand? This is another research focus which will be addressed in our further studies.

Table 9. Welfare levels (Wolfgang Zapf, 1987)

Objective life conditions	Perception and Evaluation	
	good	bad
good	Well-Being	Dissonance
bad	Adaptation	Deprivation

Table 10. The happiness chasm within group

		Low LSI	High LSI	
town	Low SWB	38	2	
		11.3%	0.6%	
	High SWB	1	19	
		0.3%	5.7%	
rural	Low SWB	2	2	
		0.8%	0.8%	
	High SWB	2	7	
		0.8%	2.8%	

5. Conclusions

The major results can be summarized as below:

- The Output Well-being Index (OWI) showed ideal metric characteristics when being applied to the sample from wealthy counties;
- The quality of life of rural residents was higher than that of the county town residents, but the level of some indicators was imbalanced;
- The participants’ performance on the subjective and objective index was roughly consistent;

The OWI could be used as an important policy instrument for the policy makers.

Acknowledgments

The research was supported by Chinese National Social Science Fund (No. 12BSH051) and the Independent Innovation Fund of Shandong University (No. IFW12111).

REFERENCES

- BUETTNER, T., EBERTZ, A., (2009). Quality of life in the regions: results for German Counties. *The Annals of Regional Science*, 43, 89–112.
- BOELHOUWER, J., (2010). Wellbeing in the Netherlands: The SCP life situation index since 1974. Den Haag: Social Cultural Plan bureau.
- CAMPBELL, A., CONVERSE, P. E., RODGERS, W. L., (1976). *The Quality of American Life: Perceptions, Evaluations, and Satisfactions*. New York: Ussell Sage Foundation.
- DIENER, E., (1994). Assessing subjective well-being: progress and opportunities. *Social Indicators Research*, 31, 103–157.
- DIENER, E., SANDVIK, E., PAVOT, W., GALLAGHER, D., (1991). Response artifacts in the measurement of subjective well-being. *Social Indicators Research*, 24, 35–56.
- DIENER, E., FUJITA, F., (1995). Resources, personal strivings, and subjective well-being: anomothetic and idiographic analysis. *Journal of Personality and Social Psychology*, 68, 926–935.
- DEATON, A., (2008). Income, health, and well-being around the world: evidence from the Gallup World Poll. *The Journal of Economic Perspectives*, 22, 53–72.
- GRAHAM, C., FELTON, A., (2006). Inequality and happiness: insights from Latin America. *Journal of Economic Inequality*, 4, 107–122.
- HAGERTY, M., (2000). Social comparisons of income in one's community: evidence from national surveys of income and happiness. *Journal of Personality and Social Psychology*, 78, 764–771.
- HUANG, L., XING, Z., (2005). An initial research on the wellbeing index applied to citizens in China. *Chinese Journal of Behavioral Medical Science*, 5, 464–465.
- KASSER, T., RYAN, R. M., (1994). A dark side of the American dream: correlates of financial success as a central life aspiration. *Journal of Personality and Social Psychology*, 65, 410–422.
- KUBISZEWSKI, I., COSTANZA, R., FRANCO, C., LAWN, P., TALBERTH, J., JACKSON, T., AYLNER, C., (2013). Beyond GDP: Measuring and Achieving Global Genuine Progress. *Ecological Economics*, 93, 66.
- KUSAGO, T., (2007). Rethinking of economic growth and life satisfaction on Post-Wwii Japan – A fresh approach. *Social Indicators Research*, 81, 79–102.

- MAGGINO, F., (2006). Perception and evaluation of the quality of life in Florence, Italy. *Community Quality of Life Indicators, Best Cases II*, 75–126.
- MAGGINO, F., NUVOLATI, G., (2011). Introduction. *Social Indicators Research*, 102, 1.
- MAGGINO, F., SCHIFINI D'ANDREA, S., (2003). Different scales for different survey methods: validation in measuring the quality of university life. *Advances in Quality of life Theory and Research*, 20, 233–256.
- MAGGINO, F., NUVOLATI, G., (2012). Quality of life in Italy (Research and Reflections). *Social Indicators Research Series*, 48, 7–103.
- MAX-NEEF, M., (1995). Economic growth and quality of life: a threshold hypothesis. *Ecological Economics*, 15, 115–118.
- MORAWETZ, D., ATIA, E., BIN-NUN, G., FELOUS, L., GARIPLERDEN, Y., HARRIS, E., SOUSTIEL, S., TOMBROS, G., ZARFATY, Y., (1977). Income distribution and self-rated happiness: some empirical evidence. *The Economic Journal*, 87, 511–522.
- NOLL, H. H., (2002). Towards a European system of social indicators: theoretical framework and system architecture. *Social Indicators Research*, 58, 51–52.
- OECD, (2006). Alternative Measures of well-being. *Economic Policy Reforms: Going for Growth*, 129–142.
- OSHIO, T., KOBAYASH, M., (2011). Area-level income inequality and happiness: evidence from Japan. *Journal of Happiness Studies*, 12, 633–649.
- PAVOT, W., DIENER, E., COLVIN, C. R., SANDVIK, E., (1991). Further validation of the satisfaction with Life scale: Evidence for the cross-method convergence of well-being Measures. *Journal of Personality Assessment*, 57, 149–161.
- ROJAS, M., (2011). The measurement of economic performance and social progress report and quality of life: moving forward. *Social Indicators Research*, 102, 169–180.
- SEIDLTIZ, L., DIENER, E., (1993). Memory for positive versus negative life events: theories for the differences between happy and unhappy persons. *Journal of Personality and Social Psychology*, 64, 654–664.
- SEN, A., K. (1993). *Capability and Well-Being, in the Quality of Life*, ed. Martha Nussbaum and Amartya Sen. Oxford: Clarendon Press.
- SUH, E., DIENER, E., FUJITA, F., (1996). Events and subjective well-being: only recent events matter. *Journal of Personality and Social Psychology*, 70, 1091–1102.

- WANG, P., (2011). The impact of income inequality on subjective well-being: evidence from Chinese General Social Survey data. *Chinese Journal of Population Science*, 3, 93–112.
- VEENHOVEN, R., (1984). *Conditions of happiness*. Dordrecht: D. Reidel Publishing.
- VEENHOVEN, R., (1995). The cross-national pattern of happiness: Test of predictions implied in three theories of happiness. *Social Indicators Research*, 34, 33–68.
- ZAPF, W., (1987). German social report. *Social Indicators Research*, 19, 5–171.
- XING, Z., (2009). Development of the revised subjective well-being scale for Chinese citizens. *Statistics in Transition*, 10, 301–316.
- XING, Z., (2011). A study of the relationship between income and subjective well-being in China. *Sociological Studies*, 151, 196–219.
- XING, Z., (2013). Bhutanese gross national happiness (GNH) measurement and enlightenment. *Dong Yue Tribune*, 34, 64–70.
- XING, Z., (2013). The Relationship between age and subjective well-being: evidence from five capital cities in mainland China. *Social Indicators Research*, 117, 743–756.
- XING, Z., ZHAO, Y., (2011). Canadian quality of life indicators systems oriented by public policy. *Shandong Social Sciences*, 185, 30–35.

APPENDIX**Subjective Well-being Scale for Chinese Citizens (SWBS-cc)**

The following items are related to something you have encountered in life or your attitudes towards life. Please read each question carefully and give an answer to it as soon as possible according to your intuition. Six options (from “strongly disagree” to “strongly agree”) are provided.

1. The society is providing us with more and more opportunities.
2. My wisdom grows with my age, and this makes me stronger and more capable.
3. Most of my life goals keep me feel refreshed, instead of making me depressed.
4. I often feel I am just being alive, not living a life.
5. I don't know the meaning of my life.
6. I often feel there must be something wrong with some of my physical organs.
7. I feel contented with my life when I compare myself against the others around.
8. I am satisfied with my family income.
9. I am often annoyed by trifling matters.
10. I am a lot worried about my own health.
11. I often find it very difficult for me to make friends with someone else.
12. I like myself.
13. I think most people have more friends than I do.
14. I really enjoy being with my family.
15. I am not as lucky as the people around.
16. I have great confidence in the development of the society.
17. I feel I did not get what I deserve, when comparing myself against the other people around.
18. It takes me a long time to get over unhappy experiences.
19. I am happy to find that I'm becoming more and more mature.
20. Sometimes I find it is very hard to communicate with other family members.

The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 461–487

OPERATIONALIZATION AND ESTIMATION OF BALANCED DEVELOPMENT INDEX FOR POLAND 1999-2016

Andrzej K. Koźmiński¹, Adam Noga², Katarzyna Piotrowska³,
Krzysztof Zagórski⁴

ABSTRACT

Stemming from assumption that Gross Domestic Product is an index oversimplifying economic development and not reflecting socio-economic development, the paper presents conceptualization, operationalization and estimation of Balanced Development Index (BDI), concerning both economic and social development in Poland. Actual values of this index as well as its four composite components (middle-level indexes) are presented for 1999-2013. A statistical model allowing estimation of BDI values as well as short-term forecasts is proposed alongside with the concept of balanced development. Application of this model for 1999-2016 is presented.

Key words: socio-economic development, index, economic indicators, social indicators, balanced development, evaluations, predictions.

1. Background

Gross Domestic Product has been criticized since the great crisis of 1930s as not reflecting the nature and the trends of socio-economic development and oversimplifying the complexity of economic development itself, even if no social aspects or consequences of it are under consideration (Kuznets 1934, 5-6). However, GDP or a very similar GNP (Gross National Product) have been almost exclusively used and interpreted as general measures of development of nations despite of all this criticism. The serious work towards improving or replacing them started only around 50 years ago. After some decline of interest in such

¹ President, Kozminski University, Warsaw, Poland. E-mail: kozmin@kozminski.edu.pl.

² Department of Economics, Kozminski University, Warsaw, Poland.
E-mail: anoga@kozminski.edu.pl.

³ Department of Statistical Methods and Informatics, Kozminski University, Warsaw, Poland.
E-mail: kpiotrowska@kozminski.edu.pl.

⁴ Empirical Sociology Center, Kozminski University, Warsaw, Poland.
E-mail: zagorski@kozminski.edu.pl.

endeavours, quite many indexes of development conceived according to “beyond GDP” paradigm (Constanza et al. 2009; Fleurbaey, Blanchet, 2013; Thiry, 2014) have been proposed, applied and published. The most known are Gross National Happiness (Pennock, Ura 2011, Ura et al. 2012), Better Life Index (OECD 2013), Human Development Index (UNDP 2013) as well as work by a group of economists and other social scientists commissioned by French President (Stiglitz et al. 2013). In Poland, the proposal to construct composite index describing both economic and social aspects of development was formulated by G. Kolodko (2008). Some work in this respect was also done in Polish Central Statistical Office (GUS 2011). The latter one, similar to the German proposal (Giesselmann et al. 2013), aimed at providing a comprehensive list of detailed indicators contrary to our project, the aim of which is to construct more general and complex indexes⁵.

Several of these projects concern only social indicators, whilst some others combine social and economic ones. Usually, however, even if they do not completely neglect economic aspects of development, they underemphasize them and overemphasize social aspects. Our approach is in line with the ‘beyond GDP’ paradigm understood not as “apart from GDP” but assigning equal weight to economic and social aspects of development.

The preliminary results of our work, including preliminary version have been already presented (Kozminski et.al., 2014). The project is still in progress. Our current aim is to substantially improve the initially presented index, to analyse and to predict changes in Polish economy and society. Here we will present the composition of the index, the idea concerning a balanced development, the trend of socio-economic development as measured by the index, a statistical model estimating values of the index and a comparison of obtained and estimated values as well as a short-term prediction.

2. Constructing the index

A primary and unique advantage of the Balanced Development Index is its composite (synthetic) character, enabling it to simultaneously measure:

- the achieved level of development, covering a wealth of both economic and social aspects;
- the system’s functional balance assessed by the degree of convergence measured as standard deviation of the four basic dimensions (realms) of the development.

⁵ We use the convention here according to which „index” is understood as a complex, composite measure, while „indicator” is understood as a simple, detailed one. Of course, this is only one of several possible conventions.

Our index concerns four socio-demographic dimensions, two economic and two social ones. Each of these is described by a set of indicators drawn from publicly available statistics and from public opinion research. Detailed indicators are then aggregated into four middle-level indexes, which – in turn – are aggregated into a general index (BDI).

One of the shortcomings of the previously proposed indexes is the neglect of the functioning of national economy in its international context. We consider that to be especially important in the present age of globalization. Our previous analysis, mentioned above, has suggested that the changes in international economic relations and in the global economic environment come before those taking place domestically. Thus, the first group of detailed indicators concerns various aspects of the functioning of Polish economy in its international, globalizing context. We call this group “external economic”. The second group of indicators characterizes various aspects of “internal economic” condition. Here, we adhere to the idea that GDP, as an important indicator, is needed but not sufficient to characterize the state of economy. We have also previously discovered that changes in public opinions about expected future conditions come ahead of subsequent evaluations and actual objective social change (Kozłowski et al. 2014). Therefore, social indicators have been grouped into the categories of “social expectations” (public hopes and fears concerning various aspects of economic, political and social life) and “current social situation” (subjective evaluations and objective state of socio-economic conditions).

The initial pre-selection of simple indicators, which together formed the initial four middle-level indexes, was largely intuitive (Kozłowski et al., 2014). It was, therefore, necessary to analytically and statistically verify the original list. The statistical verification was done first through a correlation analysis (not reported here for the lack of space), then the “Cronbach’s α ” was applied to measure the internal consistency of the four synthetic indexes and served as a final criterion for including particular indicators in them. The final list of the indicators is as follows:

A. External Economic Indicators

($\alpha = .761$)

- size of foreign direct investment (FDI) in Poland,
- WIG 20 (Warsaw Stock Exchange index),
- volume of import,
- volume of export,
- Euro/PLN exchange rate (inverted),
- spread (difference) between the Polish interest rate on 10-year bonds and German bonds (inverted)
- spread between the Polish interest rate on 10-year bonds and US bonds USA (inverted)

B. Internal Economic Indicators

($\alpha = .843$)

- production of electricity,
- number of dwellings completed,
- increase in real wages,
- Gross Domestic Product,
- size of consumption,
- size of accumulation,
- gross business profitability,
- public debt as GDP% (inverted),
- inflation rate (inverted),
- unemployment (inverted).

C. Social expectations (predictions)

($\alpha = .965$)

The ratio of positive (“will improve”) to negative (“will deteriorate”) answers to the questions whether in a year's time the situation will improve or deteriorate with respect to:

- anticipated changes in political situation in Poland,
- anticipated changes in the economic situation in Poland,
- anticipated changes in the overall situation in Poland,
- anticipated changes in the workplace,
- “Will your family live better or worse in a year's time?”,
- “Are you afraid or not that you may lose your job?”.

D. Current social situation

($\alpha = .958$)

- attitude to the government (ratio of the number of supporters to opponents),
- assessment of the political situation (ratio of those with 'positive' against 'negative' opinions),
- assessment of the economic situation (ratio of those with 'positive' against 'negative' opinions),
- how the family presently lives (ratio of those answering 'positively' against 'negatively'),
- assessment of situation in the workplace (ratio of those with 'positive' against 'negative' opinions),
- business confidence index,
- people below extreme poverty line (inverted),
- state budget expenditure on social welfare (GDP%) (inverted),
- state budget expenditure on health (GDP%),
- youth (15-24) unemployment (inverted),
- access to Internet (households%),

- birth rate,
- infant deaths (reversed),
- ratio of the population of pre-productive age compared to those of post-productive age,
- people aged 18-59 living in households in which no one is employed (%) (inverted)
- number of tertiary education graduates (public and private combined)
- number of scientific-research employees,
- young people not continuing education (%) (inverted),
- number of homicides (inverted),
- number of thefts (inverted),
- number of Polish students who went abroad on Erasmus grant,
- number of foreign students in tertiary education institutions.

The standardised value was calculated for each of these indicators in the period 1999-2013 (later data were not available). This was done according to a very simple statistical formula:

$$z = \frac{x_t - M}{SD}$$

Where: x_t is the value of an indicator in a year t , M is the average level of an indicator for the period 1999-2013 and SD is the indicator's standard deviation.

High values of Cronbach's α indicated that despite their great number the indicators can be collapsed into composite scales. Thus, non-weighted mean values of the standardised indicators were calculated for each year for the four different domains. They have constituted middle-level indexes. The correlations between them are presented in Table 1.

Table 1. Correlations between mid-level components of the general BDI index, 1999-2013 (Pearson's r above diagonal; Spearman's ρ below diagonal; Cronbach's $\alpha = .870$)

	External economic	Internal economic	Public expectations	Current social situation
External economic	x	.807**	.627**	.688**
Internal economic	.689**	X	.545*	.943**
Public expectations	.682**	.382	x	.516*
Current social situation	.650*	.854**	.575*	x

*) Significant at .05 level. **) Significant at .01 level.

Only one out of twelve correlation coefficients, namely Spearman ρ between the internal economic situation and public expectations, falls below .5 and is statistically insignificant due to a very small number of observations, even

though that correlation (.382) is of the level usually considered quite high in social sciences. Moreover, its equivalent among Pearson's correlations (.545) is statistically significant. The relatively low correlation, as compared to others (though still not negligible), between the current economic situation and public expectations indicates that the latter do not perfectly reflect economic situation of the nation. However, the expectations are strongly related to the external economic situation and to current social conditions. High correlations as well as high Cronbach's α (.870) of the four middle-level indexes allow computing the general BDI as their unweighted mean.

3. The general description of the index and its trend

As already stated, we have not applied different weights in calculating our indexes, neither to detailed variables contributing to middle-level indexes, nor to middle-level indexes constituting to the general one. High correlations between them did not allow for determining the weights by applying statistical criteria, such as often used factor analysis, the results of which are very unstable when colinearity occurs. Moreover, in substantive terms, high correlations between components suggest their more or less equal contribution to the higher-level index. On the other hand, we restrained ourselves from subjectively and intuitively assigning different weights, especially to middle-level indexes which we consider to be equally important. The changes in four middle-level indexes and their means constituting the BDI are presented in Figure 1.

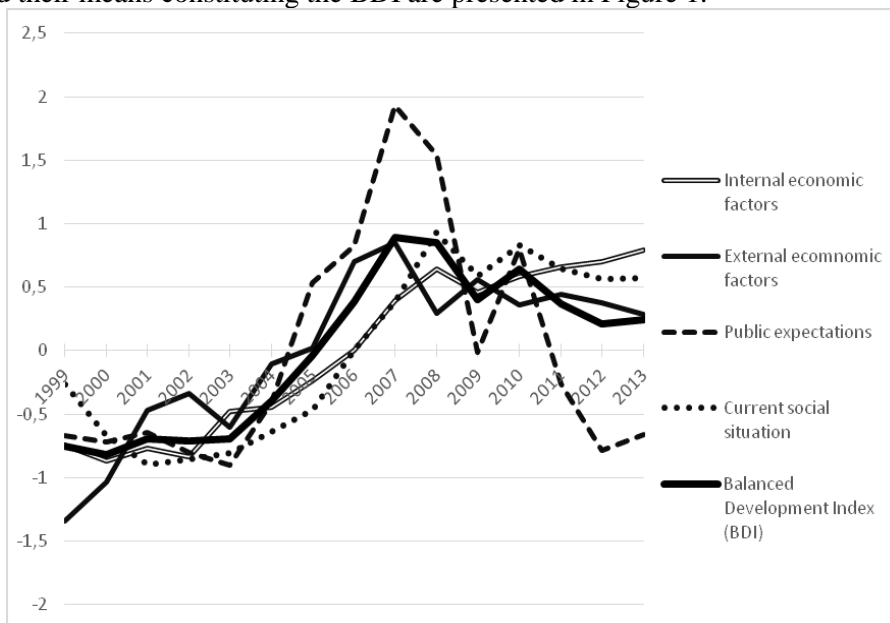


Figure 1. The four standardized components of Balanced Development Index (BDI), Poland 1999-2013

It is clear that the changes occurring in Poland take place in a similar manner for all four dimensions of development but with different amplitude. Thus, changes in the general index, calculated as an average, illustrate well the course of socio-economic development in Poland depicted by middle-level indexes. An important point of clarification is needed here. Economists and politicians emphasise the steady although slow GDP growth in Poland during and after the world financial crisis (see Fig. 2). However, the overall socio-economic index shows a slight downward trend since 2007, with some fluctuations. This downward trend is caused mostly by social trends, especially by worsening public expectations, and – to a much lesser but substantial extent – by external economic circumstances (functioning in international economic environment), while current economic situation measured by a middle-level composite index improves slightly since 2009 after a temporary decline.

The data presented in Figure 1 indicate that the variability of middle-level index of public expectations, which can be interpreted in terms of optimism and pessimism concerning the future, is much greater than the variability of the other three synthetic components of the index. This is the reason of already pointed out relatively low (though still not low in absolute terms) correlation between public expectations and economic situation.

The comparison of changes in GDP and BDI is especially important. This may be not statistically sound, since the first contributes to the second one, but such a comparison gives the idea about differences between both indexes as measures of development. As shown in Figure 2, the difference between standardized values of GDP and BDI has been recently increasing.

Two slowdowns (albeit not recessions) in economic development measured by GDP occurred twice in Poland during the investigated period: a weaker one (2000-2003) and a harsher one (2008-2009), though no full recovery is apparent until 2013.

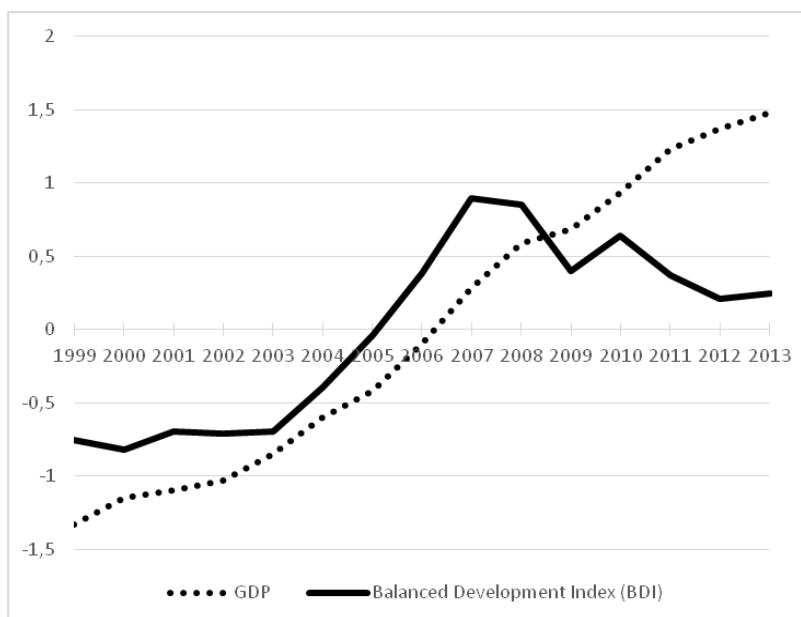


Figure 2. Changes in GDP and BDI, Poland 1999-2013 (standardised)

At the same times, the composite BDI index shows the first slowdown stronger than GDP and then the decline not shown by GDP whatsoever. In the first case, BDI suggests the stagnation while GDP was still growing, though at a smaller rate (by 3.9% in 2003). In the second case we see apparent BDI decline (with minor fluctuations only) despite the fact that GDP grew by 5.1% in 2008, 3.9% in 2010 and 4.3% in 2011. While the decline of socio-economic index has stopped in 2013, the gap between GDP and BDI remains quite big. The ability to show this discrepancy is in our opinion a very valuable feature of our synthetic index, which indicates the state of both the society and the economy much better than GDP does it.

4. Balanced development

Our idea of balanced development is based on an economic concept of equilibrium and rests upon the assumption that a certain minimum level of balance is needed to steer the system. According to the theory of Liapunov (1992), imbalances occur when individual elements excessively diverge from one another. However, some level of imbalance is needed for a change in the system (its development included). Disruption of the equilibrium is reflected in the growth of standard deviations of measured growth components. The perfect equilibrium is usually associated with a stagnation, though “dynamic equilibrium”

is also possible, defined as a parallel growth in different realms not changing the proportions between them. Our assumption is that some imbalances may pull the general index both up and down and have a positive as well as a negative effect. Imbalance is needed to knock the system out of stagnation but when it grows too much and over a long period of time, then this threatens the controllability of the system, its susceptibility to managerial impulses and stops its further development. We do not hypothesize about the optimum level of imbalance, but simply include its measure into our prediction model in order to see when the developmental trends does change. We are interested in the balance (understood as equilibrium, i.e. mutual adjustment and agreement measured as co-variation) between four domains of development embraced by BDI. Standard deviation between four middle-level indexes is used as a measure of balance.

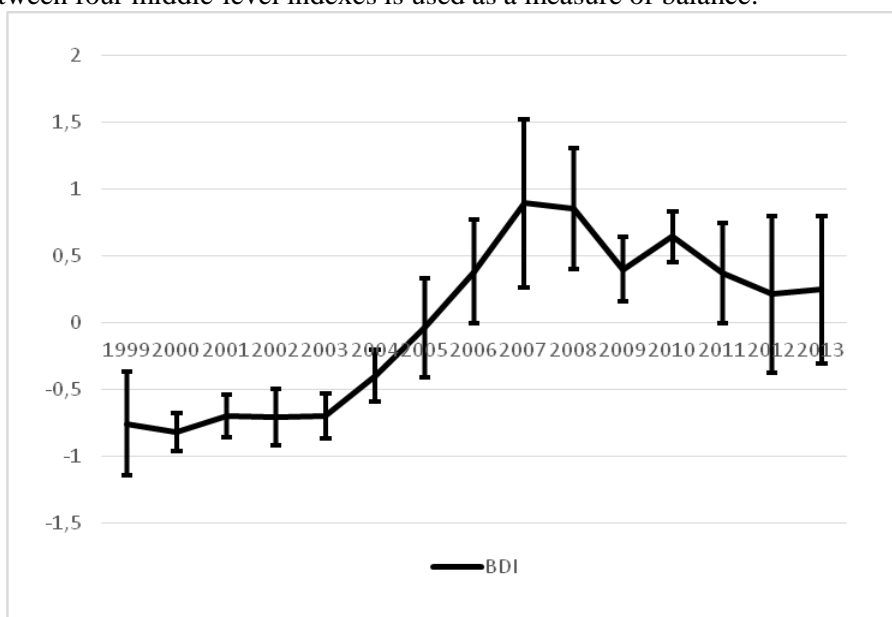


Figure 3. Changes in BDI and standard deviations of its four components, Poland 1999-2013

In accordance with our hypothesis, the changes from growth to decline, or vice versa, are taking place in the time of increase in the standard deviation and thus in conditions of imbalance. The period of a relative stagnation, or at least a slowdown in growth, was characterised by a high balance at the beginning of the 21st century. The beginning of economic crisis of 2007-09 was marked by a rising imbalance. The standard deviation exceeded .55 in 2007. This indicates a very high imbalance, and so the possibility of worsening the previously improving general socio-economic situation in the following years or - at best - halting further development. We still should see whether a similar rapid increase

in imbalance in 2012 will lead to the next reversal in the trend and in overall improvement of the situation in the forthcoming years. A slight improvement in 2013 may support such hypothesis. We can attempt to give an answer to this question based on our statistical model, presented below in the section devoted to forecasting.

5. Forecasts

A disadvantage with most of econometric forecasting methods is that they primarily predict “average” phenomena, free from discontinuities and other changes that essentially alter the course of studied processes. Fortunately, such rapid changes occurred during the period we examined, so we can try to analyse them and to predict further development taking them into account. We have attempted to design a model that allows not only for the description of current trends, but also for forecasting its change by anticipating the points where an excessive drift of all or some of the four middle-level indicators occurs. Although the beginning and the acceleration of growth demand a certain loss of socio-economic balance, it can result in a loss of control over the system and reverse the trend, if this loss of balance is too long and far-reaching. The danger of a decline may occur in conditions of imbalance growing after a period of development. There is the other side of the coin in this respect, since some imbalances occurring after a period of socio-economic decline may result in future development. Thus, we believe that the imbalance may incur both negative and positive changes in developmental trends. However, attempting prophecy or a simple extrapolation of known trends is a risky business. This is a reason of not attempting long-term predictions, which would be impossible anyway because of lack of statistical basis.

Our predictive model assumes that BDI value in year t depends on three factors:

- BDI value in the preceding year ($t-1$),
- change in BDI value between years ($t-2$) and ($t-1$),
- change in standard deviation of four BDI components in years ($t-2$) and ($t-1$).

As we do not expect large changes in BDI value from year to year, the relationship between BDI_t and BDI_{t-1} is assumed to be linear. The other two variables (both deltas) are, in turn, expected to have curvilinear, namely “S-shaped” (with two bends) relationship with BDI (which is represented by cubic polynomials in the model). As can be seen in Equation 2, the signs of highest order terms in both polynomials (b_4 and b_7) are negative, which indicates a curve that is first concave upward and then concave downward as the value of predictor increases (Cohen et al., 2003). This means that – with all other factors held constant – the increase in BDI delta as well as the delta of BDI’s standard deviation in previous year results in drop in the BDI value. Then the trend

reverses bigger deltas to co-occur with higher values of BDI. However, this is also up to a point (the second bend of the curve).

Our proposed model is as follows:

In its general form:

$$\text{Eq. 1 } BDI = b_0 + b_1 BDI_{t-1} + b_2 \Delta BDI + b_3 (\Delta BDI)^2 + b_4 (\Delta BDI)^3 + b_5 \Delta SD + b_6 (\Delta SD)^2 + b_7 (\Delta SD)^3$$

Specific values of the equation's coefficients for the period under analyses are:

$$\text{Eq. 2 } BDI' = -0,082 + 0,686 BDI_{t-1} + 0,229 \Delta BDI + 2,59 (\Delta BDI)^2 - 1,102 (\Delta BDI)^3 + 0,962 \Delta SD - 1,636 (\Delta SD)^2 - 20,855 (\Delta SD)^3$$

where :

BDI' - BDI value in the year t (predicted),

BDI_{t-1} - value of BDI in the previous year (t-1)

ΔBDI - $BDI_{t-1} - BDI_{t-2}$ (Difference between the BDI value during the previous year and two years earlier),

ΔSD - $SD_{t-1} - SD_{t-2}$ (Change in the standard deviation of the index's components between the previous year and two years ago),

SD_{t-1} - (Standard deviation of BDI components in the previous year) (t-1).

The basic premise here is to assume a curved shape of the relationship between the pace of development and the change of its degree of balance. We assume an "S" shape of basic relations between the two. The total consistency of imbalance is not conducive for development. Development requires some growth in imbalance in one or a few areas. However, too large imbalance in conditions of too quick growth causes an "overheating of the system", a lack of control and "fragility" and may lead to stagnation or decline. On the other hand, we assume that the imbalance occurring in conditions of prolonged stagnation or decline prevent excessive freezing and causes the trend to reverse to developmental one, so the cycle reverse.

Determining how one or the other (overheating and freezing) causes a trend to reverse is a matter of empirical analysis. Previous experience allows us to formulate an assumption that the desired level of balance, measured through the standard deviation of the four BDI components, is about .4. Until now, an imbalance at this level has always been associated with BDI rise during the next year. Deviations far above (such as .67 in 2007) or below (e.g. .17 in 2001) this level tend to be followed by a worsening or stagnation of socio-economic conditions. Our hypothesis is that high imbalance following the period of stagnation or decline should result in development, while high imbalance following period of growth should result in stagnation or decline. All in all, BDI constitutes a tool-box (Noll 2011) rather than a single universal tool, since it

allows for analyses of its changes as well as changes in its four composite components, their interrelations (including incongruence) and their influence on the generalised socio-economic development.

In order to forecast BDI changes, it was necessary to predict the standard deviation of its four components. For this purpose, we used the formula:

$$\text{Eq.3 } SD' = b_0 + b_1 S_{t-1} + b_2 \Delta S + b_2 (\Delta S)^2 + b_2 (\Delta S)^3$$

$$\text{Eq.4 } SD' = 0,145 + 0,817 S_{t-1} + 1,167 \Delta S - 2,614 (\Delta S)^2 - 21,147 (\Delta S)^3$$

(Notation as above).

The soundness of the proposed model is proven by a comparison of actual and estimated trends in the past. Our model reproduces the actual (empirically established) shape of the BDI curve very accurately. The Pearson's r correlation between the actual values and those estimated by the model is as high as .961. (Spearman's ρ : .918) Our estimates accurately indicate the points where the trend was changing, with an annual delay only observed in 2001 and 2003. From 2005 to 2012, actual and estimated curves completely overlap, even during the time of frequent fluctuations between 2007 and 2010. This has allowed us to predict the eventual change in the developmental trend for a short time period at least. The model allows us to estimate of IBD three years ahead. In conditions of extreme uncertainty caused by world-wide economic turbulences, we have not been tempted to make long-term predictions. The results presented in Figure 4 compare the observed and the estimated BDI curves up to 2013 and the estimates for the 2014-2016 period. We predict the sharp decline in the BDI value will come to a halt and perhaps the value will even slightly improve in 2014, for which no actual data existed in the time of the analysis.

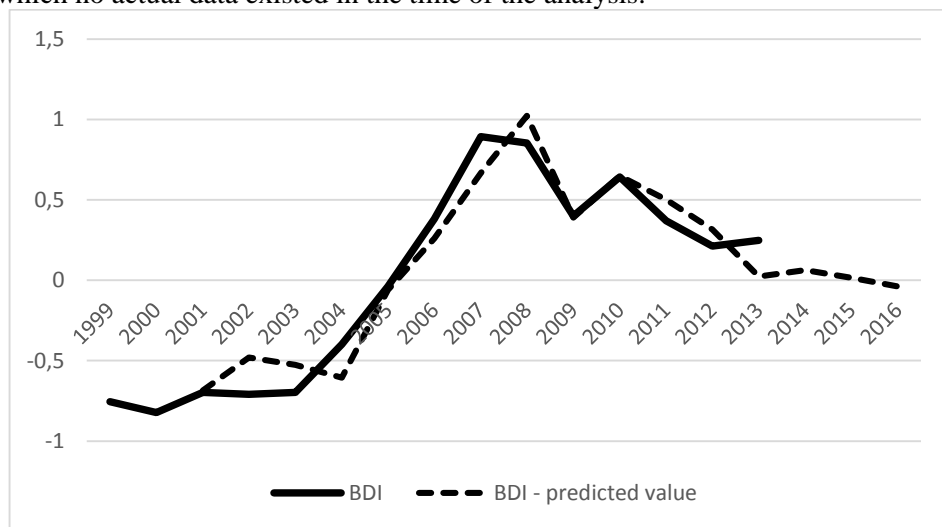


Figure 4. The actual and predicted course of socio-economic development, measured by BDI, 1999-2016

6. Conclusions

Although the time-scale during which we observed the dynamics of our indicators is short, it still gives a valuable picture of Polish socio-economic development. It covers tumultuous changes occurred during the 1999-2013 period, including the 2007 outbreak of the most serious financial and economic world crisis since 1929. Thus, the index underwent a severe stress test. All statistical models estimating and predicting complex processes may instigate various arguments. The ultimate proof for them is the consistency with empirically observed trends. Our model seems to pass this test, especially that the trend changes its direction quite often. The estimated BDI values correspond to the actual ones, which makes the short-term prediction reliable. In substantive terms, BDI suggests that socio-economic changes are less positive than those suggested by GDP growth itself, and that public expectations vary much more than the actual situation measured jointly by objective and subjective indicators.

REFERENCES

- BAURER, R. A. (ed.), (1966). *Social Indicators*. Cambridge: MIT Press.
- BES, (2014). *Il benessere equo esastenbile in Italia*. Roma: Cnel-Istat.
- CBOS, (2013). *Samopoczucie Polaków w roku 2013*. („Feelings of the Poles about Themselves in 2013”). *Komunikat z badań BS/181/2013*.
- COHEN, J., COHEN, P., WEST, S. G., AIKEN, L. S., (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd Edition Mahwah, New Jersey: L. Erlbaum Associates.
- CONSTANZA, R., HART, M., POSHER, S., TOLBERTH, J., (2009). *Beyond GDP: The Need for New Measures of Progress*, *The Pardee Papers*, 4.
- CROZIER, M., FRIEDBERG, E., (1982). *Człowiek i system. Ograniczenia działania zespołowego*. (Man and the System. Collective Action Limitations) Warsaw: PWN.
- DIENER, E., WIRTZ, D., TOW, W., KIM-PRIETO, C., DONG-WON, C., OISHI, S., BISWAS-DIENER, R., (2010). *New Well-Being Measures: Short Scales to Assess Flourishing and Positive and Negative Feelings*, *Social Indicators Research*, 97 (143–156).
- ESPINOZA, P. Z., SOMARRIBA ARECHVALA, N., (2013). *An assessment of social welfare in Spain: territorial analysis using a synthetic welfare indicator*. *Social Indicators Research*, 111: 1–23.
- FLEURBARY, M., BLANCHET, D., (2013). *Beyond GDP – Measuring Welfare and Assessing Sustainability*, Oxford: Oxford University Press.
- GIESSELMANN, M., HILMER, R., SIEGEL, N. A., WAGNER, G. G., (2013). *Measuring Well-Being: W3 Indicators to Complement GDP*, *DIW Economic Bulletin*, 5: 10–19.
- GUARDIOLA, J., PICAZO-TADEO, A., ROJAS, M., (2014). *Economic Crisis and Well-Being in Europe: Introduction*. *Social Indicators Research*. Published online first. DOI 10.1007/s11205-D14-0594-x.
- GUS, (2011). *Wskaźniki zrównoważonego rozwoju Polski*. (Indicators of Balanced Development of Poland). Katowice: GUS.
- HAGEN, L. M., (2008). *Public opinion and the economy*, in: W. Donsbach, M.W. Traugott (ed.), *The SAGE Handbook of Public Opinion Research*, London: SAGE Publications (526–533).

- HONE, L., JARDEN, A., SCHOFIELD, G., (2014). Psychometric Proprieties of the Flourishing Scale in New Zealand, *Social Indicators Research*, 119/2 (1031–1045).
- HUPPERT, F. A., SO, T. C., (2013). Flourishing Across Europe: Application of a New Conceptual Framework for Defining Well-Being, *Social Indicators Research*, 110: 837–861.
- JAROSZ, M., (2004). *Samobójstwa – Ucieczka przegranych. (Suicides – Losers' Escape)*. Warsaw: PWN.
- KEYNES, J. M., (2013). *The Collected Writings of John Maynard Keynes*, t. IV, Cambridge University Press, Cambridge UK.
- KAUFMANN, D., KRAAY, A., MASTRUZZI, M., (2010). *The Worldwide Governance Indicators: Methodology and Analytical Issues*, Working Paper No. 5430, World Bank Policy Research: <http://ssrn.com/abstract=1682130>.
- KOŁODKO, G., (2008). *Wędrujący świat (Wandering World)*. Warsaw: Prószyński i ska.
- KOŹMIŃSKI, A. K., ZAWIŚLAK, A. M., (1979). *Pewność i gra. (Certainty and Game)*. Warsaw: PWE.
- KOŹMIŃSKI, A. K., OBLÓJ, K., (1989). *Zarys teorii równowagi organizacyjnej. (Outline of the Organizational Stability Theory)* Warsaw: PWE.
- KOŹMIŃSKI, A. K., (1995). Mikro i makro równowaga funkcjonalna w warunkach transformacji. (Micro and Macro Functional Equilibrium during Transformation) *Ekonomista*, 1-2: 157–164.
- KOŹMIŃSKI, A. K., (2014). *Czas pokera. (Poker Time)*. Warsaw: Wolters Kluwer.
- KOŹMIŃSKI, A. K., (2014a). *Przedsiębiorcze państwo. (State Enterprise)* *Ekonomista*, 2: 321–336.
- KOŹMIŃSKI, A. K., NOGA, A., PIOTROWSKA, K., ZAGÓRSKI K., (2014). *Kozminski University Index of Balanced Socio-Economic Development (IBESD) for Poland*. In: Kolodko, G. (ed.) *Management and Economic Policy for Development*. New York: Nova Science Publishers (161–182).
- KOŹMINSKI, A. K., ZAGÓRSKI, K., (2015). *Spółeczeństwo w czasach kryzysu i niepewności: Założenia, pojęcia, metody (Society in Time of Crisis and Uncertainty)* in: Zagorski, K. (ed.) *Postawy ekonomiczne w czasach niepewności. (Economic Attitude in Time of Uncertainty)*. Warsaw: Scholar.
- KUZNETS, S., (1934). *National Income 1929-32 – A Report to the U.S. Senate, 73rd Congress, 2nd Session*. Washington: US Government Printing Office.

- LLEWELLYN CONSULTING, (2013). UK Infrastructure. The Challenges for Investors and Policymakers. White Paper:
https://www.pensioncorporation.com/sites/default/files/news/pic_uk_infrastructure_investment_white_paper_july_2013.pdf.
- LYAPUNOV, A. M., (1992). *The General Problem of the Stability of Motion*. London: Taylor & Francis.
- MADONIA, G., CRACOLICCI, M. F., CUFFARO, M., (2013). Exploring wider wellbeing in the EU 15 countries. An empirical application of the Stiglitz Report. *Social Indicators Research*, 111: 117–140.
- MARCH, J. G., (1981). Footnotes to Organizational Change. *Administrative Science Quarterly*, 26 (4): 563–577.
- MCGRANAHAN, D. V., RICHARD-PROUST, C., SOVANI, N. V., SUBRAMANIAN, M., (1970). *Content and Measurement of Socio-Economic Development – An Empirical Inquiry*. Geneva: United Nations Research Institute for Social Development Report Nr. 70.10.
- MOSER, C. A., (1973). *Social Indicators – Systems, Methods and Problems, The Review of Income and Wealth*, 19(2): 133–141.
- NEF The New Economics Foundation, (2009). *The Happy Planet Index*: www.dowssx417g17s.cloudfront.net/nefoundation/defaultpage/-/files/happy-planet-index-report.pdf.
- NOELLE-NEUMANN, E., (1984). *The spiral of silence: Public opinion – our social skin*, Chicago: The University of Chicago Press.
- NOLL, H. H., (2011). "The Stiglitz-Sen-Fitousi-Report : Old Wine in New Skins? Views from a Social Indicators Perspective". *Social Indicators Research*, 102 : 111–116.
- OECD The Organisation for Economic Co-operation and Development, (2013). *Better Life Index*. Downloaded from:
www.sttats.oecd.org/Index.aspx?DataSetCode=BLI#.
- PENNOCK, M., URA, K., (2011). Gross national happiness as a framework for health impact assessment. *Environmental Impact Assessment Review*, 31: 61–65 doi:10.1016/j.eiar.2010.04.003.
- REIG-MARTINEZ, E., (2013). Social and economic wellbeing in Europe and the Mediterranean Basin: Building an Enlarged Human Development Indicator. *Social Indicators Research*, 111: 527–547.
- REISS, A., (1970). Assessing the current crime wave. in: B.N. McLennan (red.) *Crime in Urban Society* New York: Donellen: 23–42.

- RODRİK, D., SUBRAMIANIAN, A., TREBBI, F., (2004). Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development, *Journal of Economic Growth*, 9: 131–165.
- SCHILLER, R. J., (2000). *Irrational Exuberance*. New York: Princeton University Press.
- SCHOPELMANN, P., GOOSENS, Y., MAKIPAA, A. (ed.), (2010). *Towards Sustainable Development – Alternative to GDP for Measuring Progress*, Wuppertal: Wuppertal Institute.
- SOMARRIBA, ARECHAVALA, N., ZARZOSA, ESPINA, P., PENA, TRAPERO, P., (2014). The Economic Crisis and its Effects on the Quality of Life in Europe. *Social Indicators Research*. Published online first. DOI 10.1007/s11205-D14-0595-9.
- STIGLITZ, J. E., SEN, A., FITOUSSI, J. P., (2013). Błąd pomiaru. Dlaczego PKB nie wystarcza. (Measurement Error – Why GDP is not Enough). Warsaw: PTE.
- STONE, R., (1966). Input-output and demographic accounting: A tool for educational planning. *Minerva* 4 (3): 365–380.
- STONE, R., BROWN, A., (1962). *A Computable Model of Economic Growth*. London: Chapman and Hall.
- TALEB, N., (2007). *The Black Swan: The Impact of the Highly Improbable* New York: Random House.
- TALEB, N., (2013). *Antykruchość – O rzeczach, którym służą wstrząsy, (Anti-fragility – About Things for Which the Shocks are Good)*. Warsaw: Kurhaus.
- THIRY, G., (2014). *Beyond GDP: Conceptual Grounds of Qualification. The Case of Index of Economic Well-Being (IEWB)*. *Social Indicators Research*. DOI 10.1007/s 1205-014-0650-6 (published online first).
- THOM, R., (1991). *Parabole i katastrofy. Rozmowy o matematyce, nauce i filozofii z Giulio Giorello i Simoną Marini, (Paraboles and Catastrophs. The Conversations with Giulio Giorello and Simone Marini about Mathematics, Science and Philosophy)* Warsaw: PIW.
- UN United Nations, (1975). *Towards a System of Social and Demographic Statistics, Studies in Methods, Series F, No.18*.
- UNDP United Nations Development Programme, (2013). *Human Development Index: www.hdr.undp.org/en/statistics/hdi*.
- URA, K., ALKIRE, S., ZANGMO, T., WANGDI, K., (2012). *An Extensive Analysis of GNH Index, Centre for Bhutan Studies: <http://www.grossnationalhappiness.com/wpcontent/uploads/2012/10/An%20Extensive%20Analysis%20of%20GNH%20Index.pdf>*.

WEINTRAUB, E. R., (1974). *General Equilibrium Theory* London and Basingstoke: Macmillan.

ZAGÓRSKI, K., (1972). Zintegrowany system statystyki społeczno-demograficznej (Integrated System of Socio-demographic Statistics), *Wiadomości Statystyczne*, 8: 9–11.

ZAGÓRSKI, K., (1972a). Dwa typy kompleksowej analizy procesów społecznych (Two Types of Complex Analyses of Social Processes), in a reader: *Ludność (Population)*, Warsaw: GUS.

ZAGÓRSKI, K., (1975). Socio-demographic Statistics in the System of Central Socio-economic Planning, *Bulletin of the International Statistical Institute*, 46: 511–522.

ZAGÓRSKI, K., (2011). Income and Happiness in Time of Post-Communist Modernization, *Social Indicators Research* 104, (331–349).

The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 479–482

ABOUT THE AUTHORS

Allin Paul is a Visiting Professor in the Department of Mathematics at Imperial College London. His research interests are the measurement of national well-being and the use of such measures in politics, policy, business and everyday life. He also lectures on official statistics. Paul spent his professional career as a statistician, researcher and policy analyst in several UK government departments and agencies, including the Equal Opportunities Commission and the Department for Culture, Media and Sport. He was latterly the director of the Measuring National Well-being programme at the Office for National Statistics.

Fattore Marco is an Assistant Professor in the Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Italy. Degree in Theoretical Physics, Ph. D. in Statistics. His main research field currently pertains to application of Partial Order Theory to the social sciences and, more generally, to the statistical treatment of multidimensional ordinal data. Particularly his interest lies (i) in the development of multi-criteria evaluation methodologies for multidimensional poverty, well-being, quality-of-life and similar socio-economic issues and (ii) in the measurement of inequality and polarization, on multidimensional systems of ordinal data. As an applied statistician, he also works in the field of Data Science and Big Data Analysis, primarily focusing on supervised and unsupervised classification. His teaching activity focuses on multivariate statistical analysis, applied to social and business sciences.

Koźmiński Andrzej K. is the founder, the first Rector and now the President of Kozminski University, the President of the Board of the International Business School in Warsaw, the corresponding member of the Polish Academy of Sciences and the member of the Presidium of Organization and Management Sciences Committee of the Polish Academy of Sciences. He has lectured at numerous renowned universities abroad. In 1981-1987, Prof. Koźmiński was the Dean of the Management Faculty of the University of Warsaw. He was also a Deputy President of the Central and the East European Management Development Association, a member of the Board of the European Foundation for Management Development based in Brussels and a member of the International Committee of the American Association of Collegiate Schools of Business. His research interests include theory of organization and management, international management, management in post-socialist countries, economic attitudes and socio-economic indicators. Prof. Koźmiński has published about 400 scientific papers and 47 books on organization, management and sociology in Poland and abroad.

Mackie Christopher is a study director for the Committee on National Statistics. He was a study director and editor for *Measuring Happiness, Suffering, and Other Dimensions of Experience* (NAS, 2013).

Maggino Filomena is a Professor of Social Statistics and Multivariate Statistical Analysis at the University of Florence. Her research interests embrace: (i) data production (in particular, subjective data assessment), (ii) data analysis (in particular, multivariate and dimensional analysis, scaling models and construction of composite and synthetic indicators), and (iii) data presentation and dissemination (with particular reference to defining a model aimed at assessing the quality of communication in statistics). The main field of application is quality-of-life and well-being measurement and analysis. Currently she serves as: Editor-in-Chief of Social Indicators Research journal (Springer), President and co-founder of the Italian Association for Quality-of-Life Studies (AIQUAV). Past-president of the International Society for Quality-of-Life Studies (ISQOLS). She is a member of several international associations, many editorial boards of scientific journals and scientific committees. She has organized and/or chaired sessions at numerous international conferences. Frequently invited all over the world to give lectures about building indicators and measuring well-being.

Noga Adam is a specialist in business cycles and enterprises, competition, accounting theories and macroeconomics. Currently he lectures at Kozminski University in Warsaw. In the past, he served as an advisor to the President of Poland, was a Vice-Rector of Warsaw School of Economics, a Chairman of the Supervisory Board of Polish Information and Foreign Investment Agency and lectured at the University of Warsaw as well as at various universities in France, Belgium and Canada. He also served several business organizations in various managerial capacities. In 1994, Prof. Noga received the Prize of Bank Handlowy for his book "Domination and effective competition". In 2009, he received an award for the best book on the enterprise "Theories of the firms", in which he presented an original "theory of confirm". In 2013, in the publication "The state as a leverage of household revenues," he presented an original theory of a "constant state" in economy.

Piotrowska Katarzyna works at Kozminski University, where she is a faculty member in the Department of Quantitative Methods and Information Technology. She has graduated from Kozminski University with MA diploma in Management, and from The Maria Grzegorzewska Academy of Special Pedagogy, where she has studied in the Department of Pedagogical Sciences. Being currently a student of Doctoral Studies in the Institute for Social Studies of the Warsaw University, she carries research in the field of cognitive aspects of entrepreneurial decision taking. Her main scientific interest areas include psychology of decision taking, perception of risk, statistics and indicators of economic and social well-being.

Qu Xiaxia is a Teaching Fellow at Shandong University of Finance and Economics; Doctor of Laws, with doctoral dissertation topic *Quality of Life Research in the Process of Harmonious Society Construction*.

Smith Conal is the head of the Well-being and Household Conditions section in the OECD Statistics Directorate. He was a lead author of the *OECD Guidelines on the Measurement of Subjective Well-being* (OECD, 2013).

Smith Dylan is a Social Psychologist and Associate Professor of Preventive Medicine at Stony Brook University. He is also affiliated with the Program in Public Health and the Department of Psychology at Stony Brook. Areas of research interest: quality of life in the context of illness and disability; subjective well-being measurement.

Tinkler Lucy has worked on subjective well-being within the UK Office for National Statistics (ONS) since the start of this programme of work in 2010. She has worked on developing the measures of subjective well-being, their introduction into ONS surveys, data analysis and dissemination of results. She is currently the Head of Measuring Personal (Subjective) Well-being.

Xing Zhanjun is a Professorial Research Fellow in the School of Political Science and Public Administration and the director of Humanities and Social Science Department in Shandong University. He is also the director of Centre for Quality of Life and Public Policy Research, and the director of Research Institute in Shandong University funded by National Bureau of Statistics of China. His research focuses on the Quality of Life and Public Policy, Subjective Well-being, and Personnel Evaluation. He is the author of three academic publications, and over 60 scholarly papers in academic journals. He has developed the Subjective Well-being Scale for Chinese Citizens (SWBS-cc), and has presided over a large-scale happiness survey research in six Chinese capital cities. His representative work is *Measuring Happiness: Research on Subjective Well-being* (in Chinese).

Zagórski Krzysztof teaches sociology and serves as a director of Empirical Sociology Center at Kozminski University, Warsaw. Previously, he was a director of Public Opinion Research Center (CBOS), and worked at the University of Warsaw, Central Statistical Office, Polish Academy of Sciences, Research School of Social Sciences of the Australian National University, and Institute of Applied Economic and Social Research of the Melbourne University. He has received his academic degrees in sociology at the University of Warsaw and Polish Academy of Sciences. He has authored and co-authored about 150 publications (books,

chapters and articles) devoted mainly to social stratification and mobility, public opinion and attitudes, social statistics and social indicators, economic and political sociology as well as living conditions and satisfactions.