Essays on the Determinants and Measurement of Subjective Well-Being

Martin Berlin

Academic dissertation for the Degree of Doctor of Philosophy in Economics at Stockholm University to be publicly defended on Friday 18 August 2017 at 10.00 in hörsal 11, hus F, Universitetsvägen 10 F.

Abstract

This thesis consists of four self-contained essays in economics, all concerned with different aspects of subjective well-being. The abstracts of the four studies are as follows.

Beyond Income: The Importance for Life Satisfaction of Having Access to a Cash Margin. We study how life satisfaction among adult Swedes is influenced by having access to a cash margin, i.e. a moderate amount of money that could be acquired on short notice either through own savings, by loan from family or friends, or by other means. We find that cash margin is a strong and robust predictor of life satisfaction, also when controlling for individual fixed effects and socio-economic conditions, including income.

Decomposing Variation in Daily Feelings: The Role of Time Use and Individual Characteristics. I explore the potential of using time-use data for understanding variation in affective well-being. Using the Princeton Affect and Time Survey, I decompose variation in daily affect into explained and unexplained within- and between person variation. Time use is found to mostly account for within-variation. Hence, its explanatory power is largely additive to that of individual characteristics. The explanatory power of time use is small, however. Activities only account for 1–7% of the total variation and this is not increased much by adding contextual variables.

The Association Between Life Satisfaction and Affective Well-Being. We estimate the correlation between life satisfaction and affect — two conceptually distinct dimensions of subjective well-being. We propose a simple model that distinguishes between a stable and a transitory component of affect, and which also accounts for measurement error in self-reports of both variables, including current-mood bias effects on life satisfaction judgments. The model is estimated using momentarily measured well-being data, from an experience sampling survey that we conducted on a population sample of Swedes aged 18–50 ($n=252$). Our main estimates of the correlation between life satisfaction and long-run affective well-being range between 0.78 and 0.91, indicating a stronger convergence between these variables than many previous studies that do not account for measurement issues.

Do OLS and Ordinal Happiness Regressions Yield Different Results? A Quantitative Assessment. Self-reported subjective well-being scores are often viewed as ordinal variables, but the conventional wisdom has it that OLS and ordered regression models (e.g. ordered probit) produce similar results when applied to such data. This claim has rarely been assessed formally, however, in particular with respect to quantifying the differences. I shed light on this issue by comparing the results from OLS and different ordered regression models, in terms of both statistical and economic significance, and across data sets with different response scales for measuring life satisfaction. The results are mixed. The differences between OLS, probit and logit estimates are typically small when the response scale has few categories, but larger, though not huge, when an 11-point scale is used. Moreover, when the error term is assumed to follow a skewed distribution, larger discrepancies are found throughout. I find a similar pattern in simulations, in which I assess how different methods perform with respect to the true parameters of interest, rather than to each other.

Keywords: *subjective well-being, happiness, life satisfaction, affect, income, cash margin, time use, measurement error, ordinal response models, cardinality.*

Stockholm 2017
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-142560

ISBN 978-91-7649-858-3
ISBN 978-91-7649-859-0
ISSN 0283-8222

Stockholm University

Department of Economics

Stockholm University, 106 91 Stockholm

ESSAYS ON THE DETERMINANTS AND MEASUREMENT OF SUBJECTIVE WELL-BEING

Martin Berlin

Stockholm
University

Essays on the Determinants and Measurement of Subjective Well-Being

Martin Berlin

ISBN print 978-91-7649-858-3
ISBN PDF 978-91-7649-859-0
ISSN 0283-8222

Cover photo: Bartolomeo Perrotta
Back cover photo: Eva Dalin

Printed in Sweden by Universitetsservice US-AB, Stockholm 2017
Distributor: Swedish Institute for Social Research

To Mema, Nils and Frans, to
whom I wish happiness.

Table of Contents

Abstract

This thesis consists of four self-contained essays in economics, all concerned with different aspects of subjective well-being. The abstracts of the four studies are as follows.

Beyond Income: The Importance for Life Satisfaction of Having Access to a Cash Margin. We study how life satisfaction among adult Swedes is influenced by having access to a cash margin, i.e. a moderate amount of money that could be acquired on short notice either through own savings, by loan from family or friends, or by other means. We find that cash margin is a strong and robust predictor of life satisfaction, also when controlling for individual fixed effects and socio-economic conditions, including income.

Decomposing Variation in Daily Feelings: The Role of Time Use and Individual Characteristics. I explore the potential of using time-use data for understanding variation in affective well-being. Using the Princeton Affect and Time Survey, I decompose variation in daily affect into explained and unexplained within- and between-person variation. Time use is found to mostly account for within-variation. Hence, its explanatory power is largely additive to that of individual characteristics. The explanatory power of time use is small, however. Activities only account for 1–7% of the total variation and this is not increased much by adding contextual variables.

The Association Between Life Satisfaction and Affective Well-Being. We estimate the correlation between life satisfaction and affect—two conceptually distinct dimensions of subjective well-being. We propose a simple model that distinguishes between a stable and a transitory component of affect, and which also accounts for measurement error in self-reports of both variables, including current-mood bias effects on life satisfaction judgments. The model is estimated using momentarily measured well-being data, from an experience sampling survey that we conducted on a population sample of Swedes aged 18–50 ($n = 252$). Our main estimates of the correlation between life satisfaction and long-run affective well-being range between 0.78 and 0.91, indicating a stronger convergence between these variables than many previous studies that do not account for measurement issues.

Do OLS and Ordinal Happiness Regressions Yield Different Results? A Quantitative Assessment. Self-reported subjective well-being scores are often viewed as ordinal variables, but the conventional wisdom has it that OLS and ordered regression models (e.g. ordered probit) produce similar results when applied to such data. This claim has rarely been assessed formally, however, in particular with respect to quantifying the differences. I shed light on this issue by comparing the results from OLS and different ordered regression models, in terms of both statistical and economic significance, and across data sets with different response scales for measuring life satisfaction. The results are mixed. The differences between OLS, probit and logit estimates are typically small when the response scale has few categories, but larger, though not huge, when an 11-point scale is used. Moreover, when the error term is assumed to follow a skewed distribution, larger discrepancies are found throughout. I find a similar pattern in simulations, in which I assess how different methods perform with respect to the true parameters of interest, rather than to each other.

Sammanfattning

Denna avhandling består av fyra fristående uppsatser i nationalekonomi, alla på temat subjektivt välbefinnande. Nedan följer sammanfattningar av de fyra delstudierna.

Studie 1: Vi undersöker sambandet mellan livstillfredsställelse och tillgång till en kontantmarginal för ett svenskt befolkningsurval. Vi finner att en kontantmarginal, en summa pengar som kan uppbådas med kort varsel antingen genom eget sparande eller till exempel lån från familj eller vänner, är en stark och robust prediktor av livstillfredsställelse. Det gäller även när vi kontrollerar för individ-fixa effekter och en rad socioekonomiska variabler, inklusive inkomst.

Studie 2: Jag undersöker användbarheten av tidsanvändningsdata för att förstå variation i affektivt välbefinnande. Jag använder Princeton Affect and Time Survey och dekomponerar daglig affekt i förklarad och oförklarad intra- och interindividuell variation. Jag finner att tidsanvändning till största delen förklarar intraindividuell variation. Dess förklaringsvärde är således i stort sett additivt till förklaringsvärdet av individkarakteristika. Tidsanvändningens totala förklaringsvärde är dock litet. Aktiviteter förklarar 1–7% av variationen i affekt, vilket inte ändras nämnvärt när även kontextuella variabler beaktas.

Studie 3: Vi skattar korrelationen mellan livstillfredsställelse och affekt, två konceptuellt distinkta dimensioner av subjektivt välbefinnande. Vi skisserar en enkel modell som gör åtskillnad mellan en stabil och en transitorisk komponent i affektivt välbefinnande. Modellen tar också tar hänsyn till mätfel i självskattningar av båda variabler, inklusive humöreffekter på självskattad livstillfredsställelse. Vi skattar modellen med momentana välbefinnandedata från en mobiltelefonbaserad enkät som vi genomförde på ett befolkningsurval av svenskar i åldern 18–50 ($n = 252$). Våra huvudestimat av korrelationen mellan livstillfredsställelse och långsiktig affekt ligger i intervallet 0.78–0.91. Det är en starkare sambandsstyrka än vad som funnits i många tidigare studier som inte tar hänsyn till mätproblematik.

Studie 4: Självskattat subjektivt välbefinnande betraktas ofta som en ordinal variabel. Det är, emellertid, en vanlig uppfattning att OLS och ordinala regressionsmodeller (t.ex. ordered probit) ger liknande resultat för sådana data. Detta påstående har dock knappt undersökts formellt, i synnerhet med avseende på att kvantifiera skillnaderna. Syftet med denna studie är att belysa denna fråga. Jag jämför resultat från OLS med resultat från olika ordinala regressionsmodeller, i termer av både statistisk och ekonomisk signifikans, och i olika datamängder som skiljer sig med avseende på vilken svarsskala som används för att mäta livstillfredsställelse. Resultaten är inte entydiga. Skillnaderna mellan skattningar från OLS, probit och logit är vanligtvis små när svarsskalan har få kategorier, men större, om än inte väldigt stora, när en 11-gradig svarsskala används. Vidare finner jag genomgående större skillnader när feltermen antas vara skevt fördelad. Jag finner ett liknande mönster i simuleringar, i vilka jag undersöker hur väl olika metoder skattar de sanna parametrarna av intresse, snarare än hur samstämmiga dessa metoder är inbördes.

Acknowledgments

First and foremost, I want to greatly thank my supervisor Markus Jäntti for his continuous support, expertise and patience. From day one, Markus encouraged me to pursue my interest in subjective well-being, despite the topic not being in the mainstream and with barely anyone else working on it in Stockholm at the time. It has been a true learning experience for me to discuss my research with Markus, as well as working as his teaching assistant in econometrics. Markus is a true gentleman, and a role model in academia as well as outside of it.

I am indebted to the great academic and social environment at SOFI, owing of course to all the people here, but I want to thank the following people in particular. Anders Björklund, for his continuous encouragement during my time at SOFI. Anders has always been generous with his time, whether providing thoughtful comments on my drafts or discussing other matters. Matthew Lindqvist, whose door has always been open for interesting discussions about research. Torsten Santavirta and I have had an endless stream of conversations about everything but work, really, and it has been a pleasure to get to know you as well as your family during these years. I have also enjoyed many a conversation with Kristian Koerselman, who took my research forward from time to time, whether by insightful comments or assistance with technical issues ranging from server support to typesetting. I am also grateful for excellent administrative support by Isabelle Andersson, Viviana Milbers, Maria Mårtensson, Elma Sose and Simon Stenborg.

I also want to thank the Department of Economics for the opportunity of pursuing a PhD, as well as for offering numerous valuable teaching opportunities. I am particularly grateful to Mahmood Arai and Michael Lundholm for introducing me to the world of econometrics with R. I also want to thank Ingela Arvidsson and Anne Jensen for excellent administrative support throughout my PhD time.

I want to thank Filip Fors at Umeå University, with whom I have had many insightful discussions about subjective well-being. They increased my understanding of the topic and motivated me to probe further into it. I also want to thank Göran Landgren at Umeå University for invaluable help with programming the survey that I carried out together with Filip, and which Study 3 in this thesis is based on.

The PhD experience was enriched by many fellow students across institutions and cohorts. In particular, I have enjoyed the company of the following people. Eric Sjöberg, my co-author Niklas Kaunitz, my co-RA's and co-eaters Pamela Campa and Erik Prawitz, and Manja Gärtner.

One of the students, whom I met during my first year, while we were both at the IIES, had a particular impact—my wife Mema. You have been the most patient of all during this journey, and an endless source of happiness to me. Thank you for love and support, I love you.

It has also been a pleasure to get to know my parents-in-law Rocco and Vita, who have provided invaluable babysitting and household assistance that enabled me to finish this thesis. Finally, I want thank my parents Lars and Agneta for having faith in me and supporting me during my PhD studies, as well as before.

* * *

I am grateful for financial support from Forte, for part of my time as a PhD-student at SOFI.

Introduction

This thesis is about subjective well-being (SWB). As evident from the name, SWB is about how well off people are in a subjective sense, i.e. as perceived from their own perspective. SWB is often called happiness in everyday language. Though not wrong *per se*, the word happiness sometimes tends to obscure two other key features of SWB (Diener, 1984). First, that SWB is concerned not only with happy (or unhappy) mind states, but with the whole range of variation from unhappiness to happiness. Second, that SWB encompasses two different dimensions: evaluative and affective well-being. The former is called life satisfaction when it refers to an evaluation of one's overall situation. It is, essentially, how you think your life is going. Affective well-being, or simply affect, on the other hand, is about the emotions and moods that are experienced momentarily as you live your life. We can think of affect either in terms of specific positive and negative feelings, such as joy, sadness and stress, or in terms of how these combine to form a sense of overall affective balance within a given time frame.

The study of SWB is an interdisciplinary field with origins as far back as the 1920's (Angner, 2011), but it has been growing particularly fast during the past ten to twenty years or so. This thesis is at the intersection of economics and psychology, with occasional references to work by sociologists and philosophers. It thus reflects the interdisciplinary nature of SWB research. Yet, it is written from the perspective of an economist, as described below.

1 Happiness Economics

One possible (though somewhat clunky) definition of economics is: the analysis of how the distribution of individual welfare outcomes in a given population is affected by the allocation of some set of scarce resources. To the extent that it is warranted to talk about "happiness economics" as a distinct sub-field, it can be defined in terms of two ideas, to be added to the definition proposed above. The first idea is to explicitly conceptualize welfare in terms of SWB, as compared to preferences (or their utility-representation) employed in standard neo-classical economics. The second idea is to use self-report measures of SWB as a means of studying such welfare outcomes empirically, rather than deducing welfare indirectly by means of observed behaviour, i.e. revealed preferences.

The rationale for equating welfare and SWB is rather self-evident—it is an outcome that we can think of as an end goal in itself, rather than as a means for something else. In this respect, SWB is fundamentally different from other, presumably welfare-relevant outcomes, such as income. In line with this view, economists such as Layard (2005) have proposed that happiness should be the main policy goal. Other advocates of the happiness perspective in economics, e.g. Frey and Stutzer (2002) and Van Praag and Ferrer-i Carbonell (2004), have to a larger extent motivated SWB in neo-classical terms, as an empirical measure of (cardinal) utility. Yet other economists studying SWB, like Benjamin et al. (2014), maintain the neo-classical framework and view aspects of SWB as (important) arguments among others in the utility function. Regardless of the exact interpretation, there is a growing recognition within economics as well as in other fields, that SWB is an important outcome worth studying. There is also a growing policy interest, as manifested e.g. in the Sarkozy report by Stiglitz et al. (2009) about alternative welfare measures, and subsequent national and international policy intitiatives.

1.1 Model

Let y^* denote a cardinal SWB variable, representing either life satisfaction or affect, and referring either to an individual's overall well-being or to a specific time frame within an individual. We are interested in the determinants of y^*, which we can think of in terms of the linear model

$$y^* = \boldsymbol{x'}\boldsymbol{\beta} + \epsilon, \tag{1}$$

where $\boldsymbol{x}$ is a vector of variables assumed to influence SWB, i.e. the "scarce resources" part, with corresponding coefficients $\boldsymbol{\beta}$. The error term ϵ captures additional variation in y^*, not accounted for by $\boldsymbol{x}$. The main interest is in estimating the elements of $\boldsymbol{\beta}$, which represent the well-being weights of $\boldsymbol{x}$. When such estimates are combined with information about the costs of changing $\boldsymbol{x}$, it is possible to assess which allocations of $\boldsymbol{x}$ are more cost-effective than others, in terms of generating individual SWB (or population-level SWB, given some social welfare function for aggregating y^*).

The vector $\boldsymbol{x}$ could encompass any resources, goods or exposure to specific policies, but is of particular relevance in contexts in which welfare-maximizing outcomes cannot be expected to come about by means of well-functioning markets. This can be the case due to externalities or irrational behaviour, or simply because there is no market, e.g. as is the case for government-provided health care and education in many countries.

Most of the economics literature on SWB revolves around estimating $\boldsymbol{\beta}$ from variations of Equation (1) by means of "happiness regressions". Ideally, such estimates should be causal, so as to be informative of the well-being consequences of policies involving changes in $\boldsymbol{x}$. Depending on the context, even estimates of $\boldsymbol{\beta}$ that are not strictly causal may be more informative than having no information at all, however, at least as a first step.

1.2 Measures

Conceptually, SWB is about self-perceived well-being. In addition, *measures* of SWB are nearly always self-reported, as it is hard to come up with other reliable ways of eliciting self-perceived mental states. Reported well-being, denoted y, may in turn differ from true (or latent) well-being, y^*, e.g. if people are only able to approximately report their well-being or if responses are not truthful. We could think of this problem in terms of a reporting function (Oswald, 2008), denoted $r()$, which maps y^* to y, i.e.

$$y = r(y^*). \tag{2}$$

For example, even though y^* is assumed to be cardinal, y might not be, if $r()$ is an ordinal mapping. The function $r()$ may also include classical or non-classical measurement error. Although we can make some plausible assumptions about $r()$, such that it is increasing in y^*, we cannot infer the shape of $r()$, due to the fact that only y is observed. Taking the idea of a reporting function seriously thus adds a considerable layer of complexity to the problem of estimating $\boldsymbol{\beta}$ from Equation (1).

2 Outline

As suggested by the title of this thesis, I address two different themes relating to the framework just presented: the determinants and measurement of SWB. The first two papers are concerned with determinants, i.e. $\boldsymbol{\beta}$ and $\boldsymbol{x}$ from Equation (1), whereas the last two papers are primarily concerned with measurement issues relating to y^* and ϵ in Equation (1), and to the reporting function $r()$ in Equation (2).

In **Study 1**, *Beyond Income: The Importance for Life Satisfaction of Having Access to a Cash Margin* (with Niklas Kaunitz), we address the relationship between economic conditions and life satisfaction—a topic that has received particular attention by economists in previous research. We find that a direct measure of cash margin, i.e. whether one could come up with a moderate amount of money on short notice, through owns savings, borrowing or by other means, is a strong and robust predictor of life satisfaction in a population sample of Swedes. This is true also when controlling for individual fixed effects and socio-economic conditions, including income. Since it shows not to matter whether cash margin comes from own savings or with help from family members, this measure captures something beyond wealth.

In **Study 2**, *Decomposing Variation in Daily Feelings: The Role of Time Use and Individual Characteristics*, I explore the usefulness of time use variables for explaining variation in affect in a US population sample. Affect is measured (retrospectively) for three different occasions of the previous day, for each respondent. This allows me to decompose the overall variation into explained and unexplained within- and between-person variation. I find that activities, and the context in which they take place, capture variation in affect that is distinct from the variation captured by individual socio-economic characteristics, as well as life satisfaction. Although this suggests that there is value added to the time use approach for understanding SWB, I also find that time use only accounts for a small share of the total variation in affect.

The outcome in **Study 1** is life satisfaction, whereas it is affective well-being in **Study 2**. This raises the question of whether it is preferable to study either outcome over the other, and how they relate to each other. I try to answer this question in **Study 3**, *The Association Between Life Satisfaction and Affective Well-Being* (with Filip Fors). This study is thus concerned with the left-hand side of Equation (1), i.e. what particular aspect of SWB to choose for y^*, but also with the issue of measurement error in $r()$.

We propose a simple model that distinguishes between a stable and a transitory component of affect, and which also accounts for measurement error in self-reports of both variables, including current-mood bias effects on life satisfaction judgments. We estimate the model using momentarily measured well-being data, from an experience sampling survey that we conducted on a population sample of Swedes aged 18–50. We find a strong correlation between life satisfaction and long-run affective well-being, both in absolute terms and relative to many previous studies that do not account for measurement issues.

Study 4, *Do OLS and Ordinal Happiness Regressions Yield Different Results? A Quantitative Assessment*, can also be motivated in terms of a discrepancy between the other studies, namely whether reported SWB, y, should be treated as ordinal as in **Study 1**, or as cardinal as in **Study 2** and **Study 3**. The conventional wisdom has it that OLS (assuming cardinality) and ordinal regression models produce similar results when applied to SWB data. This claim has rarely been assessed formally, however, in particular with respect to quantifying the differences. I examine this issue by comparing the results from OLS and different ordered regression models (e.g. ordered probit), in terms of both statistical and economic significance, and across data sets with different response scales for measuring life satisfaction. I also use simulations, in which I assess how OLS and ordered regressions perform with respect to the true parameters of interest, rather than to each other. My results do not overturn the conventional wisdom, but they paint a more nuanced picture.

This study is thus concerned with ordinality of the reporting function $r()$, the distribution of ϵ, and how different assumptions with regard to these matter for the estimates of $\boldsymbol{\beta}$.

I conclude this introduction with a note to the reader. Although I have tried to make the terminology and notation in this thesis somewhat consistent, all four papers were written to be self-contained. Hence, some amount of discrepancy, as well as redundancy, is to be expected. Moreover, the papers appear in the chronological order that they were written. As a result, prior papers may not incorporate insights reflected in subsequent ones.

References

Angner, E. (2011). The evolution of eupathics: The historical roots of subjective measures of wellbeing. *International Journal of Wellbeing 1*(1), 4–41.

Benjamin, D. J., O. Heffetz, M. S. Kimball, and A. Rees-Jones (2014). Can marginal rates of substitution be inferred from happiness data? Evidence from residency choices. *The American Economic Review 104*(11), 3498–3528.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin 95*(3), 542–575.

Frey, B. S. and A. Stutzer (2002). What can economists learn from happiness research? *Journal of Economic Literature 40*(2), 402–435.

Layard, R. (2005). *Happiness: Lessons from a new science.* New York: Penguin.

Oswald, A. J. (2008). On the curvature of the reporting function from objective reality to subjective feelings. *Economics Letters 100*(3), 369–372.

Stiglitz, J. E., A. Sen, and J.-P. Fitoussi (2009). Report by the commission on the measurement of economic performance and social progress.

Van Praag, B. M. and A. Ferrer-i Carbonell (2004). *Happiness quantified: A satisfaction calculus approach.* Oxford: Oxford University Press.

Study 1

Beyond Income: The Importance for Life Satisfaction of Having Access to a Cash Margin*

Annual income twenty pounds, annual expenditure nineteen nineteen and six, result happiness. Annual income twenty pounds, annual expenditure twenty pounds ought and six, result misery.

—Charles Dickens, *David Copperfield*

1 Introduction

Social scientists are to an increasing extent treating subjective well-being (SWB) measures—especially evaluations of overall life satisfaction and happiness—as useful welfare measures.[1] There is now a fast-growing literature on the determinants of SWB, and especially on the relationship between life satisfaction and economic conditions.

---

*This paper is co-authored with Niklas Kaunitz and has been published in the Journal of Happiness Studies, 2015, 16(6), pp 1557–1573. The final publication is available at Springer via `http://dx.doi.org/10.1007/s10902-014-9575-7`. This version differs slightly from the published one with respect to formatting. This paper has benefited greatly from suggestions for improvements by Markus Jäntti and Anders Björklund. The authors also wish to express their gratitude towards Johan Egebark, Louise Johannesson, Maria Perrotta Berlin, and seminar participants at SOFI and at the Department of Economics at Stockholm University for valuable comments.

[1] See e.g. Frey and Stutzer (2002) and Di Tella and MacCulloch (2006) on the use of SWB data in economics, and Diener et al. (1999) for a general survey of the psychology literature.

It is by now well-established, across different types of samples and estimation settings, that the within-country association between life satisfaction and household income is positive (Argyle, 2003, Clark et al., 2008, and Diener and Biswas-Diener, 2002, review this literature). As expected from economic theory, the relationship is concave, and it is typically estimated using the logarithm of income (see e.g. Layard et al., 2008, and Stevenson and Wolfers, 2008). Controlling for individual fixed effects has been found to reduce the impact of income—although a positive association clearly remains—whereas the life satisfaction–income association is relatively robust to whether life satisfaction is modelled as an ordinal or a cardinal variable (Ferrer-i-Carbonell and Frijters, 2004).

Despite the robust statistical association between life satisfaction and income, the impact of income must be considered small, both in absolute terms and relative to other determinants of life satisfaction, and the explanatory power of income is also fairly small.[2] This is somewhat puzzling, at least in the light of economists' attention to income in many other settings.

However, typical "happiness regressions", in which life satisfaction is regressed on household or personal income contemporaneous with the well-being response, may give a misleading answer to the broader question of whether money buys happiness. A person with low income may, for example, not worry much about money if he or she also has low expenses, high wealth, or is able to borrow money easily. To the extent that consumption determines well-being, contemporaneous income is only a noisy proxy that could be expected to lead to downward-biased estimates.

Although the potential problem of using income to represent material circumstances has been recognised (see e.g. Diener and Biswas-Diener, 2002), there are only a few studies that investigate how SWB is related to economic conditions defined more broadly than contemporaneous income. An early study by Mullis (1992) shows that well-being is better predicted by a composite measure including both a proxy for permanent income, based on earnings averaged over several years, and a measure of annuitised net worth, scaled by household size. In a similar vein, Headey and Wooden (2004), Headey et al. (2008) and D'Ambrosio et al. (2009) find that substantially more variation in life satisfaction can be accounted for when adding wealth to the analysis. The two

[2] We are not the first ones to make this interpretation, and it is discussed e.g. by Headey et al. (2008) and Christoph (2010).

latter studies also show that income averaged over several years is more relevant in terms of both magnitude and explanatory power. Moreover, Headey et al. (2008) also find that consumption expenditures is at least as important as income. Inspired by sociologial poverty research, Christoph (2010) finds that a deprivation index—a checklist of amenities that the household is lacking—is a better predictor of life satisfaction than income.

In this paper, we add to this literature by investigating how life satisfaction is influenced by yet another variable related to, but distinct from, income: having access to a *cash margin*. Specifically, we use a decade-long panel sample of the Swedish Level of Living Survey, in which respondents were asked whether they could come up with a moderate sum of money within a week—either through their own savings or by some other means, e.g. borrowing from family or friends. Lack of such a margin can be interpreted as a more direct measure of economic distress than having a low income, but at the same time capturing something distinct from wealth. To the best of our knowledge, this variable has not been considered as a determinant of life satisfaction before.

The rest of this paper is organised as follows: we describe our data and our method in Sections 2 and 3, whereafter we present the results in Section 4. We discuss our results and conclude in Section 5.

2 Data

2.1 Data Sources and Sample

Our main data source is the Swedish Level of Living Survey (LNU): a socioeconomic panel survey designed to be representative of the Swedish population aged 18–75 (Jonsson and Mills, 2001). Interviews were conducted by the Swedish statistical agency, Statistics Sweden, either face-to-face in the respondent's home or by telephone.

The LNU is unusual because of the long time span inbetween survey waves—we use the two waves from 1991 and 2000—and hence our panel models capture long-term intra-individual variation. We include all individuals in the 1991 wave that were re-interviewed in 2000, except those living with their parents and those with any item non-response on the variables used in our analysis.[3]

[3]We drop 327 individuals living with their parents in either 1991 and 2000. These are mostly youths that move out from their parents' home between 1991 and 2000. The motivation for this sample restriction is that income comparisons between this group, mainly

Our balanced estimation sample consists of $N = 6,406$ observations and $n = 3,203$ individuals.

Our second data source is income register data matched to each respondent in the LNU, as well as to his or her partner (as identified by the survey).

2.2 Variables

We use two different satisfaction measures as outcome variables. *Satisfaction with life circumstances* (henceforth SLC) is based on the following question:

> We have now been through a lot of questions about your living conditions in different areas. How do you yourself view your own conditions? By and large, do you think that your situation is: very good, rather good, neither good nor bad, rather bad, or very bad?

This question is located at the very end of the survey, within a block of judgments and opinions, and at this point the respondent has been interviewed about his or her circumstances across several domains, such as family situation, health, education and occupation. It is thus plausible that the question captures satisfaction across all these domains. The second outcome measure, *satisfaction with daily life* (henceforth SDL), is located shortly before the SLC question within a block of questions of a more psychological character:[4]

> Do you usually feel that your daily life is a source of personal satisfaction? (Yes, most often / yes, sometimes / no)

Although we choose distinct labels for these two satisfaction measures, we believe that both are comparable to the life satisfaction measures found in other surveys (e.g. "All things considered, how satisfied are you with your life as a whole these days?", in the World Values Survey). Given the phrasings and the survey context, it is likely that SLC is somewhat more sensitive to external aspects of life—including economic conditions—whereas SDL should tap more into internal aspects of well-being. By considering both outcomes we can to some extent assess this, which is interesting in its own right.[5]

---

supported by their parents' income, and others are hard to interpret. The attrition rate (of those eligible for re-interview in 2000) between the two waves is 22.0%.

[4]This measure has been used by Andersson (2008) who studies the effects of self-employment on well-being, and by Gerdtham and Johannesson (2001) who study the correlates of well-being with a focus on health.

[5]Both of our satisfaction measures can be considered mostly cognitive and evaluative in nature, in comparison to more specific measures of positive and negative affect which are

As income measure we use the individual's income in the case that he or she lacks a spouse, and when there is a spouse, the simple average of the spouses incomes (i.e. income per capita among spouses). The income variable, which is based on register data, includes labour and capital incomes net of taxes, as well as important transfers such as child allowance and social welfare benefits. However, our results do not hinge on using this particular measure of household income (see Appendix B.2 for details).

The main contribution of this paper is to complement the income variable with a more immediate measure of economic conditions: whether one has access to a cash margin. The variable is based on the following question in the LNU:

> If a situation suddenly arose where you had to come up with 10,000 kr, could you manage it? (yes / no)

The figure amounts to ca. \$1,170 in 2011 prices, and was adjusted to 12,000 kr in the 2000 survey, keeping it roughly constant in real terms. These amounts correspond to slightly less than the monthly median income in our sample (the median was 11,728 kr in 1991 and 12,865 kr in 2000). 8.6% of the respondents in our sample reply "no" to this question in 1991, and the shares who lack a cash margin are, from the lowest income quartile to the highest: 13.1%, 9.6%, 8.1% and 3.5%. Lack of a cash margin is thus not rare, nor is it solely a low-income phenomenon. The responses are distributed similarly in 2000, but with somewhat fewer people lacking a cash margin (8.1%), which might be due to the older sample at this point, as older people are more likely to have a cash margin.

We use a follow-up question asked to those who have a cash margin, in order to distinguish between those who have own savings and those who are able to borrow from a close family member, from a relative or friend, from the bank, or acquire money by some other means. There are thus six different cases, including those who lack a cash margin.[6]

Finally, we employ a set of standard control variables that are likely to

also encompassed in the concept of SWB (Diener et al., 1999). Different measures that can broadly be classified as evaluative can still vary along an evaluative-affective continuum, however, and more evaluative measures have been found to correlate more strongly with material circumstances (Diener et al., 2010).

[6] Respondents are only able to choose one alternative for this question. Although this is not clear from the survey documentation, we interpret the responses as hierarchical, in the sense that responses higher up in the order of response categories (as we list them) are chosen first when possible. This interpretation implies, for example, that those who have a cash margin through a bank loan are not able to borrow from family members or friends.

correlate with both well-being and economic conditions: age (5 categories), sex, health (index based on 44 symptoms), marital and parental status (5 categories), education level (3 categories) and employment status (6 categories). The control variables are described further in Appendix A.

2.3 Descriptive Statistics

The sample distribution of SLC in 1991 and 2000 is shown as a transition matrix in Table 1. The overall satisfaction distribution is quite stable between 1991 and 2000, with few people reporting low satisfaction, which is typical for life satisfaction data. Still, there is a fair amount of transitions. Though the SDL measure has fewer response categories, the dynamics are similar.

Table 1: Satisfaction with life circumstances 2000 conditional on 1991 (%)

	Satisfaction 2000					
Satisf. 1991	1.	2.	3.	4.	5.	All 1991
1. Very bad	7.7	23.1	23.1	46.2	0.0	0.4
2. Rather bad	9.5	11.9	14.3	47.6	16.7	1.3
3. Neither	0.0	3.4	21.6	60.2	14.8	2.7
4. Rather good	0.5	1.7	3.2	64.9	29.7	53.0
5. Very good	0.1	0.3	0.8	33.1	65.7	42.6
All 2000	0.5	1.4	2.9	50.9	44.3	

$n = 3,203$. Rows 1–5 show the satisfaction distribution in 2000 (in %) conditional on the satisfaction in 1991.

The first column of Table 3 shows descriptive statistics for the most important variables, for all observations in our pooled 1991–2000 sample. To highlight the raw patterns in the data, columns 2–4 show means by the levels of SDL. Those reporting higher daily satisfaction are more likely to be cohabiting, have some higher education and somewhat less health problems. Moreover, satisfied individuals are more likely to be working part-time or being self-employed, and are less likely to be unemployed. These differences are broadly in line with previous research.[7] The patterns with respect to SLC (not shown) are similar, which is indicative of the validity of our two outcome measures.

[7]For example, Clark and Oswald (1994) document lower well-being among unemployed in Britain, and Stutzer and Frey (2006) find married persons to be happier.

Table 2: Satisfaction with daily life 2000 conditional on 1991 (%)

	Satisfaction 2000			
Satisfaction 1991	1.	2.	3.	All 1991
1. No	27.5	41.0	31.5	5.6
2. Yes, sometimes	6.7	45.4	48.0	33.3
3. Yes, most often	3.2	26.5	70.3	61.2
All 2000	5.7	33.6	60.7	

$n = 3,203$. Rows 1–3 show the satisfaction distribution in 2000 (in %) conditional on the satisfaction in 1991.

The two rightmost columns in Table 3 show descriptive statistics by cash margin. It can be seen that individuals who lack a cash margin share, on average, several characteristics with those having a low satisfaction: i.e. they have less education, worse health, are less likely to be married or cohabiting, more likely to be single parents and more likely to be unemployed. There are also differences between these two groups, however. Those without margins are more likely to work part-time rather than full-time, and are are also younger, whereas age is only weakly correlated with SLC and SDL.

3 Method

The observed values of our outcome variables have an ordinal interpretation—we only know the order of the outcomes and not their magnitudes. Hence, we treat each observed life satisfaction outcome variable y as an ordinal representation of an interval-scale latent variable y^*. The number of observed categories of y is 5 for SLC, and 3 for SDL. Our model for the latent variable, for individual i and period t, is

$$y^*_{it} = \alpha_i + \boldsymbol{cm}'_{it}\boldsymbol{\beta}_{cm} + \beta_{inc}\log(inc_{it}) + \boldsymbol{z}'_{it}\boldsymbol{\gamma} + \varepsilon_{it}, \quad (1)$$

where α_i is an individual-fixed effect, $\boldsymbol{z}_{it}$ is a vector of control variables (with categorical variables included as sets of dummy variables), and ε_{it} is an error term. The vector $\boldsymbol{cm}$ consists of five dummy variables corresponding to the six different cases of cash-margin described in the previous section (using cash margin through own savings as the reference category), and the corresponding

Table 3: Descriptive statistics, means (sd)

		Satisfied w. daily life			Cash margin	
	All	No	Sometimes	Yes	No	Yes
SLC = Very good	0.43	0.15	0.28	0.55	0.19	0.46
SLC = Rather good	0.52	0.61	0.66	0.43	0.67	0.51
SLC < Rather good	0.05	0.24	0.06	0.02	0.15	0.04
Income	13,274	11,908	12,779	13,672	10,970	13,484
	(8,710)	(4,568)	(6,565)	(9,924)	(3,383)	(9,012)
No cash margin	0.08	0.20	0.11	0.06		
Has cash margin						
Own savings	0.72	0.57	0.66	0.76		0.78
Loan from family	0.04	0.06	0.04	0.04		0.05
Loan from friend	0.07	0.09	0.08	0.06		0.08
Bank loan	0.07	0.08	0.09	0.07		0.08
Other	0.01	0.01	0.01	0.01		0.01
Female	0.52	0.50	0.49	0.53	0.68	0.50
Age	45.94	44.88	45.35	46.35	41.43	46.35
	(13.30)	(14.20)	(13.07)	(13.32)	(13.12)	(13.24)
High school	0.42	0.43	0.45	0.41	0.46	0.42
Higher education	0.26	0.19	0.23	0.29	0.16	0.27
Symptom index	6.44	10.27	7.08	5.74	9.50	6.16
	(5.59)	(7.54)	(5.92)	(4.97)	(7.47)	(5.30)
Cohabiting	0.77	0.55	0.74	0.81	0.57	0.79
Cohab. parent	0.40	0.27	0.40	0.42	0.40	0.40
Single parent	0.04	0.08	0.04	0.04	0.14	0.03
Full time	0.54	0.48	0.57	0.54	0.44	0.55
Part time	0.16	0.11	0.16	0.17	0.21	0.16
Self-employed	0.08	0.05	0.06	0.09	0.02	0.08
Unemployed	0.03	0.08	0.04	0.02	0.10	0.02

1991 and 2000 pooled sample. $N = 6,406$, $n = 3,203$. See Appendix A for an explanation of the control variables.

coefficient vector $\boldsymbol{\beta}_{cm}$ is thus what we are mainly interested in. We are also interested in comparing the estimate of $\boldsymbol{\beta}_{cm}$ with that of β_{inc}, the coefficient on log household income.

Assuming that the error terms ε_{it} are independent and follow a logistic distribution, Equation (1) can be estimated with ordered logit regression. The robustness of this specification is examined in Appendix B; this includes testing the restrictiveness of the logarithm form for income as well as using OLS instead of ordered logit for estimating the parameter values.

The purpose of including fixed effects α_i is to control for time-constant unobserved individual characteristics that correlate with both life satisfaction and the independent variables of interest. Failing to control for fixed effects has been shown to produce biased estimates of various determinants of life satisfaction, including income (Ferrer-i-Carbonell and Frijters, 2004). The same could be true for cash margin if, for instance, stable personality traits influence both well-being and the likelihood of having access to a cash margin.

However, it is well-known that implementing fixed effects in the ordered logit model is not straight-forward (see e.g. Wooldridge, 2010, Section 15.8). Our approach is to use the *BUC estimator* proposed by Baetschmann et al. (2015). The BUC estimator, which is an extension of the fixed-effects binary logit model (Chamberlain, 1984), utilises all possible dichotomisations of the dependent variable, but discards observations where the dependent variable is constant over time.[8] This method allows for arbitrary correlation between the individual effects and the explanatory variables, as in the linear fixed-effects model.

To facilitate interpretation of the estimates we have scaled coefficients and standard errors by the standard deviation of the latent dependent variable. Hence, the reported coefficients measure the impact on life satisfaction in terms of standard deviations, associated with a unit change in the explanatory variable. For measuring goodness of fit we use $\overline{R^2} = \mathrm{var}(\hat{y}^*)/\mathrm{var}(y^*)$.[9] Similarly

[8]Given K discrete outcomes, the BUC estimator creates $K-1$ new observations from each original one by transforming the dependent variable to one of the $K-1$ possible dichotomisations, thus giving rise to a $K-1$ times larger data set. Since the expanded data set has a binary outcome variable we can use the Chamberlain fixed-effects model. Finally, to account for the dilution of observations, standard errors are made cluster-robust with respect to individuals. Hence the acronym *BUC*, "Blow-Up and Cluster". The BUC estimator is similar to the methods of Das and van Soest (1999) and Ferrer-i-Carbonell and Frijters (2004), but is argued to be more robust for small samples (Baetschmann et al., 2015).

[9]Letting $\boldsymbol{x}$ denote the vector of all covariates, with corresponding coefficients $\boldsymbol{\beta}$, the latent variable variance follows from Equation (1): $\mathrm{var}(y^*) = \boldsymbol{\beta}'\mathrm{var}(\boldsymbol{x})\boldsymbol{\beta} + \mathrm{var}(\varepsilon)$, where $\mathrm{var}(\varepsilon) = \frac{\pi^2}{3}$ is imposed in the logit model and $\mathrm{var}(\hat{y}^*) = \hat{\boldsymbol{\beta}}'\mathrm{var}(\boldsymbol{x})\hat{\boldsymbol{\beta}}$ is estimated from the sample. $\overline{R^2} = \mathrm{var}(\hat{y}^*)/\mathrm{var}(y^*)$ is now straightforwardly obtained from the above expressions. The normalisation procedure as well as the idea for measuring goodness of fit was first suggested by McKelvey and Zavoina (1975).

to R^2 for linear models, $\overline{R^2}$ is the share of dependent (latent) variable variance explained by the covariates. Note that for the fixed-effects regressions, $\overline{R^2}$ measures the share of explained variance *after* the fixed effects have been eliminated.

4 Results

The results for satisfaction with life circumstances and satisfaction with daily life are presented in Tables 4 and 5, respectively. Recall that the reference cash margin category is that of having own savings (the most common case). Hence, all cash margin estimates should be interpreted as the well-being difference relative to this group.[10] For comparison we present pooled cross-section regressions without fixed effects (columns 1 and 2), as well as panel regressions including fixed effects (columns 3 and 4). Moreover, to disentangle the separate influences of economic factors and the control variables, we present regressions including cash margin and income only, without the control variables (columns 1 and 3). Our preferred specification includes both fixed effects and the full set of controls (column 4). The full regression output, showing the coefficient estimates also for the control variables, is reported in Appendix C.

The lack of a cash margin, relative to having own savings, has a strong negative association with both life satisfaction measures, regardless of specification. This relationship is most pronounced for SLC, ranging from -0.76 in the cross-section to -0.52 when including fixed effects and control variables. But cash margin has a distinct impact on satisfaction with daily life as well, at between -0.50 and -0.31 standard deviations of SDL. For both satisfaction measures, additional controls and individual fixed effects reduce the impact, but not dramatically so. The impact of not having a cash margin is greater than the impact of cohabitation or marriage (see Appendix C), which is typically found to be one of the most important correlates of life satisfaction in the literature. Comparing with the impact of income in the cross-section, it takes more than a five-fold increase in income to balance the decrease in SLC associated with the lack of a cash margin—in 2000 this corresponded to going

[10] An alternative specification is to contrast lack of cash margin with a single category of having a cash margin, thus adding all cases under "Has cash margin" in Tables 4 and 5 to the reference category. This would slightly reduce the coefficient for lack of cash margin, to $(-0.76, -0.5, -0.54, -0.41)$ for SLC and $(-0.48, -0.29, -0.27, -0.23)$ for SDL.

Table 4: Results, satisfaction with life circumstances (ordered logit regressions)

	Pooled cross-section		Fixed effects	
	(1)	(2)	(3)	(4)
No cash margin	−0.76***	−0.58***	−0.65***	−0.52***
	(0.06)	(0.06)	(0.12)	(0.12)
Has cash margin				
Own savings (ref. category)	—	—	—	—
Loan from family	−0.05	−0.11	−0.07	−0.08
	(0.07)	(0.07)	(0.13)	(0.12)
Loan from friend	−0.28***	−0.23***	−0.24**	−0.18*
	(0.06)	(0.06)	(0.11)	(0.10)
Bank loan	−0.43***	−0.33***	−0.33***	−0.34***
	(0.05)	(0.05)	(0.11)	(0.11)
Other	−0.20	−0.11	0.10	0.11
	(0.13)	(0.14)	(0.25)	(0.22)
Income (log)	0.35***	0.35***	0.05	0.11
	(0.04)	(0.05)	(0.09)	(0.10)
CONTROLS	NO	YES	NO	YES
N	6,406	6,406	1,316	1,316
$\overline{R}^2$	0.08	0.17	0.05	0.15

Significant at *** 1%, ** 5%, * 10%. Coefficients from ordered logit regressions, standardised by latent variable standard deviation. Standard errors cluster-robust w.r.t. individuals. Controls include health, family conditions, employment status, level of education, age group, and sex (cross-section only). All cross-section specifications include year-fixed effects.

from the 10th to the 95th income percentile.[11] For SDL, it takes a twenty-fold increase in income to compensate for the lack of a cash margin. As reported in Section 2.2, while lack of a cash margin is more prevalent in low-income groups, it is not limited to the latter. Consequently, these results do not represent a low-income effect but capture something distinct from income level.

It turns out that it matters not only if, but also *how* an individual has access to a cash margin. Borrowing from close family rather than having access to own savings is not associated with significantly lower levels of satisfaction (although the point estimates suggest a small negative cost). In contrast, those

[11] The equivalent income change is computed as $e^{|-0.58/0.35|} = 5.2$.

Table 5: Results, satisfaction with daily life (ordered logit regressions)

	Pooled cross-section		Fixed effects	
	(1)	(2)	(3)	(4)
No cash margin	−0.50***	−0.33***	−0.34***	−0.31***
	(0.05)	(0.06)	(0.11)	(0.11)
Has cash margin				
Own savings (ref. category)	—	—	—	—
Loan from family	−0.08	−0.07	−0.16	−0.16
	(0.07)	(0.07)	(0.11)	(0.11)
Loan from friend	−0.25***	−0.15***	−0.24**	−0.20**
	(0.05)	(0.06)	(0.10)	(0.10)
Bank loan	−0.23***	−0.14***	−0.09	−0.13
	(0.05)	(0.05)	(0.10)	(0.10)
Other	−0.13	−0.08	−0.12	−0.12
	(0.15)	(0.15)	(0.22)	(0.22)
Income (log)	0.16***	0.11**	−0.11	−0.08
	(0.04)	(0.04)	(0.09)	*(0.10)*
Controls	No	Yes	No	Yes
N	6,406	6,406	1,413	1,413
$\overline{R^2}$	0.03	0.09	0.02	0.07

Significant at *** 1%, ** 5%, * 10%. Coefficients from ordered logit regressions, standardised by latent variable standard deviation. Standard errors cluster-robust w.r.t. individuals. Controls include health, family conditions, employment status, level of education, age group, and sex (cross-section only). All cross-section specifications include year-fixed effects.

who need to take a bank loan, or borrow from a friend, experience a sizeable drop in well-being in comparison to those with own economic resources. These results suggest that a sense of economic security is what matters, rather than wealth *per se*—the more uncertain the means to accessing a cash margin are, the stronger the adverse effect on well-being. Own savings and lack of a cash margin can thus be seen as two ends of a continuum.[12]

[12]Social relations have been found to be important for SWB (see e.g. Powdthavee, 2008), and it is possible that our cash margin variable to some extent reflects this. The fact that those who lack a cash margin have a lower life satisfaction than those who have it by means of a bank loan suggests the importance of something beyond social relations alone, however. See also footnote 6.

The impact of income is less robust. While statistically significant in the cross-section regressions, the association disappears when controlling for individual-fixed effects, for both outcome measures. For satisfaction with life circumstances in the pooled cross-section, a doubling of income is associated with a 0.24 standard-deviation increase in SLC ($0.35 \cdot \log 2$). The corresponding figure for SDL is a mere 0.08 standard-deviation increase.[13]

For both outcome measures the share of explained variation is generally small, dropping further when adding fixed effects. This is consistent with previous literature stressing the importance of individual-fixed effects as determinants of subjective well-being. SLC seems to be somewhat easier to explain by observed factors, as indicated by higher $\overline{R^2}$-values. If the regressions are run without income, $\overline{R^2}$ is virtually unaffected for SDL, while for SLC in the cross-section, $\overline{R^2}$ is somewhat lower at 0.06 (instead of 0.08). Regressing life satisfaction on income and year-fixed effects alone, we get $\overline{R^2}$-values in the range 0.00–0.03 (highest for SLC).

In general, economic factors have a stronger impact on SLC than on SDL. Even so, the impact of cash margin on both life satisfaction measures is impressive. While including control variables generally reduces this impact, its relative importance and statistical significance remain robust.

5 Discussion

We have shown that there is a strong association between life satisfaction and having access to a cash margin. This is true whether we consider different individuals in the cross-section or the same individuals over time. The results are robust to controlling for other socio-economic factors, including different measures of, and specifications of income. The positive impact on life satisfaction is largest when one has own savings or is able to borrow from one's family. There is a substantial satisfaction cost of having to borrow from friends or the bank, however, perhaps due to a larger social cost and more insecurity.

Our interpretation of these results is that having a sense of economic security, in a broader sense than wealth or income, is important for subjective

[13] As a comparison, Sacks et al. (2010) find coefficients in the range 0.22–0.28 for the log-income impact on standard deviations of life satisfaction. One reason for the lack of income effects in the long run may be that we use a very long panel, with nine years between the two waves. This is in line with the literature on adaptation of well-being, stating that many factors correlated with well-being have a diminishing impact over time (Clark, Diener, Georgellis, and Lucas, 2008).

well-being. In the case of wealth, we could not test this explicitly, but is something we infer from the fact that being able to borrow from family appears, from a well-being perspective, to be as good as having own savings. The fact that the cash margin variable remains important even when controlling for income is somewhat puzzling, however, and warrants further discussion.

First, lack of a cash margin might capture low income *relative* to the individual's own consumption standard. The individual's consumption standard is presumably slow-moving, and is a function both of own past consumption and that of some reference group.[14] A mismatch between income and consumption standard might for example arise if one fails to adjust one's consumption when faced with a negative income shock. A permanent mismatch would also be possible if the standard rises in line with income increases in the reference group.[15]

Second, lack of a cash margin might reflect self-control problems, which may lead to failure to save, or "overspending". This interpretation is in line with the fact that people also report lack of a cash margin towards the top of the income distribution. However, a sizeable impact remains when accounting for stable personality traits through the inclusion of fixed effects. Hence, to the extent that the results can be attributed to self-control problems, these are time-varying rather than fixed.

The above explanations are perhaps not mutually exclusive, and may also vary in relevance depending on what part of the income distribution we consider. Some of the characteristics of those reporting lack of a cash margin (see Section 2.3), e.g. marital and employment status, do indeed indicate that this group has less economic resources. At the macro level, the share of people without a cash margin has also been shown to co-vary with the share of people defined as absolutely poor and the share of people seeking social assistance (Jonsson et al., 2010). The fact that those without a cash margin are younger is consistent with these explanations: it is conceivable that younger people, to a greater extent, have their consumption standards misaligned with their income levels, but it is also plausible that they are less accustomed to long-term

[14] Such a framework is discussed in Clark, Frijters, and Shields (2008). Robert Frank has also written extensively about the importance of reference groups and "positional goods", see e.g. Frank (1985).

[15] A related explanation is that the lack of a cash margin could capture differences in living expenses across regions. To test this hypothesis, we have run regressions where we include regional dummies and their interaction with income, but with no substantial results: the income interactions are generally not significant and lack of cash margin remains equally strong. Hence, we reject this explanation.

economic responsibility.

Regardless of the exact interpretation of the cash margin variable, its strong association with life satisfaction supports the view that, from a well-being perspective, measures of the household's economic situation should be defined more broadly than income. For future research, it would be interesting to combine measures of long-term income, wealth and consumption (as in Headey et al., 2008), with a measure of cash margin.

References

Andersson, P. (2008). Happiness and health: Well-being among the self-employed. *Journal of Socio-Economics 37*(1), 213–236.

Argyle, M. (2003). Causes and correlates of happiness. In D. Kahneman, E. Diener, and N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*, Chapter 18, pp. 353–372. New York: Russell Sage Foundation.

Baetschmann, G., K. E. Staub, and R. Winkelmann (2015). Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 178*(3), 685–703.

Chamberlain, G. (1984). Panel data. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of econometrics*, Volume 2, Chapter 22, pp. 1247–1318. Amsterdam: North-Holland.

Christoph, B. (2010). The relation between life satisfaction and the material situation: A re-evaluation using alternative measures. *Social Indicators Research 98*(3), 475–499.

Clark, A., P. Frijters, and M. Shields (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature 46*(1), 95–144.

Clark, A. and A. Oswald (1994). Unhappiness and unemployment. *The Economic Journal 104*(424), 648–659.

Clark, A. E., E. Diener, Y. Georgellis, and R. E. Lucas (2008). Lags and leads in life satisfaction: A test of the baseline hypothesis. *The Economic Journal 118*(529), F222–F243.

D'Ambrosio, C., J. Frick, and M. Jäntti (2009). Satisfaction with life and economic well-being: Evidence from Germany. *Schmollers Jahrbuch 129*(2), 283–295.

Das, M. and A. van Soest (1999). A panel data model for subjective information on household income growth. *Journal of Economic Behavior & Organization 40*(4), 409–426.

Di Tella, R. and R. MacCulloch (2006). Some uses of happiness data in economics. *Journal of Economic Perspectives 20*(1), 25–46.

Diener, E. and R. Biswas-Diener (2002). Will money increase subjective well-being? *Social Indicators Research 57*(2), 119–169.

Diener, E., D. Kahneman, W. Tov, and R. Arora (2010). Income's association with judgments of life versus feelings. In E. Diener, J. Helliwell, and D. Kahneman (Eds.), *International differences in well-being*, Chapter 1, pp. 3–15. Oxford: Oxford University Press.

Diener, E., E. Suh, R. Lucas, and H. Smith (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin 125*(2), 276–302.

Ferrer-i-Carbonell, A. and P. Frijters (2004). How important is methodology for the estimates of the determinants of happiness? *Economic Journal 114*(497), 641–659.

Frank, R. (1985). *Choosing the right pond: Human behavior and the quest for status*. Oxford: Oxford University Press.

Frey, B. and A. Stutzer (2002). What can economists learn from happiness research? *Journal of Economic Literature 40*(2), 402–435.

Gerdtham, U. and M. Johannesson (2001). The relationship between happiness, health, and socio-economic factors: Results based on swedish microdata. *Journal of Socio-Economics 30*(6), 553–557.

Headey, B., R. Muffels, and M. Wooden (2008). Money does not buy happiness: Or does it? A reassessment based on the combined effects of wealth, income and consumption. *Social Indicators Research 87*(1), 65–82.

Headey, B. and M. Wooden (2004). The effects of wealth and income on subjective well-being and ill-being. *Economic Record 80*, S24–S33.

Jonsson, J. O. and C. Mills (2001). The Swedish Level-of-Living-Surveys: A general overview and description of the event history data. In J. Jonsson and C. Mills (Eds.), *Cradle to grave: Life-course change in modern Sweden*, pp. 228–242. Durham: Sociology Press.

Jonsson, J. O., C. Mood, and E. Bihagen (2010). Fattigdomens förändring, utbredning och dynamik. In Socialstyrelsen (Ed.), *Social rapport 2010*, Chapter 3, pp. 90–126. Västerås: Edita Västra Aros.

Layard, R., G. Mayraz, and S. Nickell (2008). The marginal utility of income. *Journal of Public Economics 92*(8–9), 1846–1857.

McKelvey, R. D. and W. Zavoina (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology 4*(1), 103–120.

Mullis, R. (1992). Measures of economic well-being as predictors of psychological well-being. *Social Indicators Research 26*(2), 119–135.

Powdthavee, N. (2008). Putting a price tag on friends, relatives, and neighbours: Using surveys of life satisfaction to value social relationships. *Journal of Socio-Economics 37*(4), 1459–1480.

Sacks, D., B. Stevenson, and J. Wolfers (2010). Subjective well-being, income, economic development and growth. NBER Working Paper 16441, National Bureau of Economic Research.

Stevenson, B. and J. Wolfers (2008). Economic growth and subjective well-being: Reassessing the Easterlin paradox. *Brookings Papers on Economic Activity 1*, 1–87.

Stutzer, A. and B. Frey (2006). Does marriage make people happy, or do happy people get married? *Journal of Socio-Economics 35*(2), 326–347.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2 ed.). Cambridge, MA: MIT Press.

A Control Variables Definitions

Income: average of spouses' combined monthly disposable income, i.e. net of taxes and transfers, based on tax register data contemporaneous with the survey year. (Equals individual income when the respondent is neither married nor cohabiting.)

Age group, 5 categories: 18–26, 27–36, 37–50, 51–64 and 65–75. Age is approximated by subtracting birth year from the survey year.

Highest completed education, 3 categories: *basic school* (education level is junior high school or lower, includes no schooling), *high school* (highest completed education level is high school, gymnasium, or a short vocational training), *higher education* (has completed some higher level education, i.e. a university diploma or a longer vocational education).

Symptom index: numeric variable based on the summation of 44 separate symptom scores that take the values 0, 1 or 2, if the respondent has no, mild, or severe symptoms, respectively. Hence, an index score of zero indicates perfect health.

Marital status, 3 categories: *not married*, *cohabiting* (including married), *divorced or widowed*.

Children in household, 3 categories: indicates whether there are any children currently living in the respondent's household, regardless of how many: *no children*, *cohabiting parent* (children living in household together with married or cohabiting respondent), *single parent* (children living in household, respondent is neither married nor cohabiting).

Employment status, 5 categories: respondents in LNU may hold multiple employment statuses, e.g. working full-time while searching for a job. We define mutually exclusive employment indicators in the following lexicographic order, meant to capture main activity: *full-time*, *part-time*, *self-employed* (works either in a firm partly or fully owned by him or herself, or in a free profession), *unemployed* (is currently searching for a job), *retired* (at least 65 years old and receiving pension) and *other*. The last category includes all respondents not falling into any other category, e.g. students.

B Sensitivity Analysis

B.1 Functional Form of Income

In line with most of the previous literature on life satisfaction we have imposed a logarithmic functional form upon the relationship between income and well-being. To assess the plausibility of this assumption, we have also estimated the model

$$y^*_{it} = \alpha + \boldsymbol{cm}'_{it}\boldsymbol{\beta}_{cm} + \sum_{q \in Q} \beta_q I_q(inc_{it}) + \boldsymbol{z}'_{it}\boldsymbol{\gamma} + \varepsilon_{it}, \tag{2}$$

where Q represents a partitioning of $[0, 100]$ into quantiles and $I_q(\cdot)$ is the indicator function for quantile q. That is, we here estimate the mean impact of income on life satisfaction separately for each quantile.

Our results indicate that SLC does indeed display an approximately logarithmic association with income, while for SDL the relationship is somewhat weaker. The case for choosing a logarithmic functional form is strengthened further by the fact that $\overline{R^2}$-values for the cross-section regressions are virtually unaffected by changing to the less restrictive income specification, for both outcome measures.[16]

Most importantly, the cash margin estimates are unaffected by changing the income specification, for satisfaction with both circumstances and daily life. The results from the flexible income specification are, thus, in accordance with our previous results.

B.2 Choice of Income Measure

Above we argued that erroneous functional form of income in the regressions cannot explain the large estimates of the impact of cash margin. However, in theory it is possible that the income measure itself is flawed, and that this biases the estimates of cash margin.

We have tested the results using three different income measures: individual disposable income, per-spouse household disposable income (our preferred income measure) and equivalised household disposable income. The latter is defined as total household disposable income, excluding earnings by children, divided by the square root of the number of family members in the household.

[16]These results are consistent with Layard et al. (2008) who examine the functional relationship between happiness and income in a number of countries. They find that across most specifications, the relationship is logarithmic, or somewhat more concave.

The idea behind this measure is to include the economic cost of children, but also capture household economies of scale.[17]

While there are some differences in how these measures interact with well-being (although they tend to converge as control variables are included), the cash margin estimates are largely unaffected by the choice of income measure.[18] Thus, while choosing the correct income measure is important for estimating the impact of income on life satisfaction, the cash margin estimates are robust in this regard.

B.3 Consistency of the BUC estimator

The BUC estimator first expands the data set by using all possible dichotomisations of the dependent variable, thus deriving a binary dependent variable. At the next stage, the Chamberlain fixed-effects model for binary logit eliminates all observations where the, now binary, outcome variable is constant across all time periods. Taken together, this implies that the BUC estimator weights each original observation $(y_{it}, \boldsymbol{x}_{it})$ proportionally to the span that y_{it} covers within that individual. Individuals with constant y_{it} over time will, consequently, not influence the coefficient estimates $\hat{\boldsymbol{\beta}}$ at all.

Under the assumptions of the BUC estimator, this sample reweighting does not threaten consistency. However, if, for example, $(\boldsymbol{\beta} \mid y, i)$ should vary across outcomes or individuals, the BUC estimator would give the mean impact of the covariates for the weighted subsample, rather than for the original representative sample.[19] In such a case, it would be difficult to compare the cross-section

[17]For single-person households without children all three measures coincide, for single person households with children the first two measures coincide, and for two-spouse households with two children the household measures coincide. As regards scaling, the per-spouse household measure can be considered a filtered version of the individual measure, reducing within-couple variation (noise) but maintaining the order of magnitude. The equivalised income, on the other hand, is anchored at singles without children and couples with two children, i.e. it coincides with the per-spouse measure at these points. At other points, however, equivalised income is scaled differently: for example, singles with children have their income adjusted downwards while couples without children have their income adjusted upwards.

[18]When it comes to explanatory power, $\overline{R^2}$ is consistently higher with both household measures than when using individual income. Per-spouse household income is the most robust measure; for the two other measures, including control variables tends to push estimates towards the per-spouse estimate.

[19]This could happen if, e.g., unhappy individuals would tend to experience higher fluctuations in well-being than happy individuals. In fact, from examining the correlation between $\bar{y}_i$ and $|\Delta y_i|$ it turns out that such an interrelationship does exist in our data. The correlation is stronger for satisfaction with daily life (-0.39) than for satisfaction with life circumstances (-0.28). This correlation could, of course, be the result either of larger coefficient magnitudes or higher error variance for these individuals (or both).

estimates with those from the fixed-effects specifications.

As an informal test we have run all cross-section regressions also for the reduced sample used in the fixed-effects estimations (i.e. dropping individuals with constant satisfaction levels over time, separately for each outcome measure). For the SLC specifications the sample reduction leads to the cash margin coefficients being mildly attenuated, but in no case significantly different from earlier. For SDL the attenuation is somewhat stronger, at most a decrease in magnitude of around a third. The attenuation is generally a bigger problem for the income estimates than for the cash margin estimates. Consequently, the sample reduction associated with the BUC estimator is not a likely source of bias for our results on cash margin.

We have also estimated all models with OLS rather than ordered logit. In addition to simply using equidistant integer values for the outcome variables (1–5 for SLC, 1–3 for SDL), we have also estimated all specifications with the outcome variables transformed to binary outcomes in a linear probability model (with the cutoff between the top category and those below, in both cases). In neither case is the overall picture different from that described below: the variables of interest have similar relative magnitudes and statistical significance. These results can be obtained from the authors upon request.

C Detailed Regression Results

Table C.1: Results, all covariates (ordered logit regressions)

	Satisfaction with living conditions				Satisfaction with daily life			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
No cash margin	−0.76***	−0.58***	−0.65***	−0.52***	−0.50***	−0.33***	−0.34***	−0.31***
	(0.06)	(0.06)	(0.12)	(0.12)	(0.05)	(0.06)	(0.11)	(0.11)
CM: Own savings (ref. category)	—	—	—	—	—	—	—	—
CM: Loan, family	−0.05	−0.11	−0.07	−0.08	−0.08	−0.07	−0.16	−0.16
	(0.07)	(0.07)	(0.13)	(0.12)	(0.07)	(0.07)	(0.11)	(0.11)
CM: Loan, friend	−0.28***	−0.23***	−0.24**	−0.18*	−0.25***	−0.15***	−0.24**	−0.20**
	(0.06)	(0.06)	(0.11)	(0.10)	(0.05)	(0.06)	(0.10)	(0.10)
CM: Bank loan	−0.43***	−0.33***	−0.33***	−0.34***	−0.23***	−0.14***	−0.09	−0.13
	(0.05)	(0.05)	(0.11)	(0.11)	(0.05)	(0.05)	(0.10)	(0.10)
CM: Other	−0.20	−0.11	0.10	0.11	−0.13	−0.08	−0.12	−0.12
	(0.13)	(0.14)	(0.25)	(0.22)	(0.15)	(0.15)	(0.22)	(0.22)
Income (log)	0.35***	0.35***	0.05	0.11	0.16***	0.11**	−0.11	−0.08
	(0.04)	(0.05)	(0.09)	(0.10)	(0.04)	(0.04)	(0.09)	(0.10)
Year 2000	−0.04*	−0.00			−0.05**	−0.06**		
	(0.02)	(0.03)			(0.02)	(0.03)		
Age ≤ 36		−0.16***		0.06		−0.04		0.04
		(0.06)		(0.11)		(0.06)		(0.11)
Age ≤ 50		−0.28***		0.04		−0.07		−0.06
		(0.07)		(0.18)		(0.07)		(0.17)
Age ≤ 64		−0.28***		0.11		0.01		−0.12
		(0.07)		(0.23)		(0.07)		(0.23)
Age ≤ 75		−0.18		0.14		0.21		−0.17
		(0.15)		(0.38)		(0.16)		(0.36)
Female		0.30***				0.18***		

Continued on next page

Table C.1 continued

		(0.03)				(0.03)		
Educ. HS		0.02				0.04		
		(0.04)				(0.04)		
Educ. above HS		0.12***				0.16***		
		(0.04)				(0.05)		
Symptom index		−0.03***		−0.03***		−0.03***		−0.03***
		(0.00)		(0.01)		(0.00)		(0.01)
Cohabiting		0.39***		0.44***		0.26***		0.20*
		(0.05)		(0.10)		(0.05)		(0.10)
Divorced/widow		−0.13*		−0.15		0.03		0.01
		(0.07)		(0.15)		(0.07)		(0.14)
Cohab. parent		−0.09**		−0.11		0.03		−0.10
		(0.04)		(0.07)		(0.04)		(0.07)
Single parent		−0.20**		−0.26		0.08		−0.14
		(0.08)		(0.16)		(0.08)		(0.13)
Part-time worker		0.06		0.13		0.03		0.00
		(0.04)		(0.08)		(0.04)		(0.08)
Self-employed		0.20***		0.17		0.21***		0.29**
		(0.06)		(0.12)		(0.06)		(0.13)
Unemployed		−0.14*		−0.27*		−0.18**		−0.03
		(0.08)		(0.14)		(0.08)		(0.12)
Retired		0.04		0.18		−0.02		0.13
		(0.14)		(0.28)		(0.16)		(0.26)
Other occupation		0.03		0.07		0.03		0.07
		(0.05)		(0.09)		(0.05)		(0.09)
Fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
N	6,406	6,406	1,316	1,316	6,406	6,406	1,413	1,413
$\overline{R^2}$	0.08	0.17	0.05	0.15	0.03	0.09	0.02	0.07

Significant at *** 1%, ** 5%, * 10%. Coefficients from ordered logit regressions, standardised by latent variable standard deviation. Standard errors cluster-robust w.r.t. individuals.

Study 2

Decomposing Variation in Daily Feelings: The Role of Time Use and Individual Characteristics

1 Introduction

Subjective well-being (SWB) encompasses both specific positive and negative feelings, affect, and cognitive evaluations of one's overall situation, life satisfaction. Happiness can thus be characterized in terms of prevalence of positive feelings, absence of negative feelings, and a sense of overall life satisfaction (Diener, 1994). It is an ongoing discussion to what extent these components of SWB represent the same underlying construct (see Busseri and Sadava, 2011), but it is typically advocated that they should be measured and analysed separately (see e.g. Diener et al., 1999).

Yet, the economics literature on SWB has mostly relied on single-item survey measures of life satisfaction. Presumably, this is both because such data are relatively easy to collect for large representative samples, and because of the possible interpretation of life satisfaction as a proxy for overall experienced utility.

There are compelling reasons not to focus only on the cognitive aspect of SWB, however. First, to the extent that well-being is truly multidimensional, affective and evaluative measures may not always give the same result, as suggested e.g. by Knabe et al. (2010) in a study of German unemployed and by

Kahneman et al. (2010) in a study of French and American women. Second, affect data may be preferable from a measurement perspective if people's life satisfaction judgments are subject to cognitive biases (Kahneman, 1999, Kahneman and Krueger, 2006) and sensitive to context effects (Schwarz and Strack, 1999 and Connolly, 2013). Third, as affect varies from moment to moment, in contrast to life satisfaction, affect measurement can be combined with "ecological" data about the individual's behaviour and environment, thus making it possible to study the proximate determinants of happiness.

A main candidate among such proximate determinants is *time use*, i.e. the activities that people do, with whom and where they do it, etc. Motivated by these ideas, Krueger et al. (2009) have proposed a promising research program on National Time Accounting (NTA, see also Kahneman et al., 2004b and Krueger, 2007).[1] The basic theoretical framework underlying NTA is that individual well-being, during a given time horizon, can be regarded as the sum of the well-being generated by each activity, weighted by the time spent in each activity.

The measurement of combined time use and well-being data proposed for NTA is based on the Day Reconstruction Method (DRM), developed by the same authors (Kahneman et al., 2004a). The DRM (described further in Section 2.1) is, in short, a survey method in which the respondent reports how he or she spent the previous day in terms of distinct episodes, of which some are rated with respect to affective well-being.[2] As part of the NTA program, DRM data were collected for a representative sample of US adults in the Princeton Affect and Time Survey (PATS). Subsequently, a similar well-being module has been added to the American Time Use Survey (ATUS), carried out by the Bureau of Labor Statistics.

The motivation for this paper is to explore the potential of using such combined time-use and affect data for measuring and understanding SWB. Rather than investigating the impact of specific aspects of time use, I focus on the simple question of how much total explanatory power time use variables have with respect to affect. Arguably, such a bird's-eye perspective is relevant for assessing the usefulness of decomposing well-being differences over time or

[1] See Juster and Stafford (1985) for an earlier approach to time use and well-being.

[2] The Experience Sampling Method (Larson and Csikszentmihalyi, 1983) and the Ecological Momentary Assessment method (Stone and Shiffman, 1994) are often considered to be the gold standard for collecting such data in real-time, but have been relatively expensive and cumbersome to implement. The recent development of smart-phone based survey methods is changing this, however.

across countries in terms of differences in time use, as proposed in Kahneman et al. (2004b), or for extrapolating national well-being based on historical time-use data, as in Krueger (2007). To the best of my knowledge, the only previous analysis along these lines is by White and Dolan (2009). However, their focus is on augmenting the DRM with measures of meaningfulness, for which time-use variables are found to be better predictors, compared to affect.

I base my analysis on the PATS data set, which includes three measurements of affect for each individual (within a day). This allows me to decompose the total variation in affect into within-and between-person variation, explained or unexplained by time use, respectively. In case of the between-variation, I also contrast the explanatory power of time use with that of individual socio-economic characteristics as well as life satisfaction—or put more concretely, are some people happier because of who they *are*, or is it because of what they *do*?

My main findings are as follows. A sizeable share of the overall variation in affect—up to two thirds—can be attributed to variation within individuals. In principle, there is thus plenty of room for explanatory factors other than demographic characteristics or fixed personality traits—the latter often being pointed out as important determinants of SWB. However, I find that time use only accounts for a small part of the variation in affect. Activities account for 1–7% of the variation, depending on the specific affect outcome, and adding contextual variables does not increase this figure much. Moreover, time use mostly explains within-variation. As a consequence, the explanatory power of time use is additive to that of individual characteristics and life satisfaction. Finally, I find that time use is a relatively better predictor of positive affect, whereas individual characteristics are better predictors of negative affect.

The rest of this paper is structured as follows. I describe the PATS data in Section 2 and the method in Section 3. In Section 4, I present the results, which are discussed further in Section 5.

2 Data

2.1 The Princeton Affect and Time Survey

The PATS was designed by Krueger et al. (2009), based on the ATUS and DRM. The survey was conducted in 2006 through telephone interviews administered by the Gallup Organization. The sample was selected using a random-digit dial technique covering all households in continental U.S. with a residential telephone line, and one person of age 16 or older was selected from each sampled household. The final sample consists of 3959 persons, reflecting a 37% response rate.[3]

Respondents were assigned an interview-day randomly, across all days of the week. During the interview, the respondent was asked to divide the previous day into distinct *episodes*, defined as non-overlapping time intervals associated with one main activity. Information about where the activity took place and who else was present was also collected. After the whole day was described, three episodes from the non-sleeping part of the day were sampled without replacement, with probabilities proportional to the episodes' duration.[4] The respondents were then asked to what extent they experienced the following six feelings during the sampled episodes: happy, interested, tired, stressed, sad and pain. For each episode, all of these feelings were assessed on a scale from 0 to 6, where 0 means "not at all" and 6 means that the feeling was "very strong". 18 affect scores were thus collected for every individual, partial non-response aside.

Demographic and socio-economic information was also collected, including education level, employment status and household income. In particular, respondents were asked the following question about general life satisfaction: "Taking all things together, how satisfied are you with your life these days? Are you: very satisfied, satisfied, not satisfied, or not at all satisfied?" For more details on the PATS, see Krueger et al. (2009) and Krueger (2007).

The sample used in the subsequent analysis is obtained by removing observations with partial non-response in affect scores and relevant covariates.[5] Individuals with fewer than three episodes are also removed, yielding a balanced sample of $N = 11,469$ observations and $n = 3,823$ individuals.

[3] I use the supplied weights to make sure that the weighted data conform to the distribution of the Current Population Survey, with respect to observable demographic characteristics. The weights, which also account for the sampling structure, are used throughout this paper (unless noted otherwise).

[4] Three 15-minute intervals were randomly selected, and if two intervals were contained within the same block of time associated with one activity (i.e. an episode), one of these was dropped and a new interval was selected. Although the affect ratings refer to these 15-minute intervals, the intervals will be referred to as episodes in this paper for simplicity.

[5] In order to preserve the sample size as far as possible, missing values in independent category variables are assigned to separate missing/other categories, included as separate indicator variables in the analysis.

2.2 Affect Outcomes

It is not obvious how one should account for the multi- dimensionality of affective well-being. A natural starting point is to look at measures that summarize information across different affect dimensions relating to the same episode. Two such summary measures are the *u-index* and *net affect*, and these will be the focus of the analysis.

The u-index, proposed by Krueger et al. (2009), measures time spent in an unpleasant state. More precisely, the episode-level u-index is equal to one whenever the maximum score among the three negative feelings stress, sadness and pain exceeds the score of happiness, and zero otherwise. Since the unit of observation in this paper is an episode, the u-index is thus simply an indicator variable rather than a proper index, i.e. it is not aggregated within individuals or activities. An advantage of the u-index is that it is ordinal at the level of feelings, meaning that it is robust to individual differences in the use of response scales that apply to all feelings within the same individual (Krueger et al., 2009). An interpersonal comparison of the u-index between persons A and B is thus not invalidated if A only uses the lower end of the response scale for all feelings, whereas B only uses the upper end.

Table 1: Affect score response distribution

	Intensity (% of responses)								
	0	1	2	3	4	5	6	Mean	Std dev
Happy	5.3	2.8	6.9	17.1	18.9	23.3	25.7	4.1	1.7
Interested	7.1	3.8	8.2	15.3	18.0	20.6	27.0	4.0	1.8
Tired	25.0	8.8	12.7	15.3	14.7	13.0	10.6	2.7	2.1
Stressed	47.9	12.4	12.8	9.5	6.7	5.4	5.3	1.5	1.9
Sad	75.2	8.5	5.6	4.0	2.6	1.8	2.4	0.7	1.4
Pain	72.4	5.7	5.4	5.1	4.4	3.4	3.6	0.9	1.7
U-index								0.2	0.4
Net affect								3.1	2.4

$N = 11,469$, $n = 3,823$. An intensity of 0 means that the feeling was not experienced at all and 6 means that the feeling was very strong.

Although the u-index uses information on several affect dimensions, a drawback is that it discards information about the intensity of feelings. Net affect, on the other hand, is defined as the happiness score minus the average of stress, sadness and pain, and thus reflects the full range of variation in intensity for these emotions.

Moreover, happiness, sadness, stress and pain are also analysed separately. The outcomes interestedness and tiredness are more ambiguous with respect to the dimension positive–negative valence, i.e. if the experience is good or bad (see Russell, 1980), and will thus not be analysed here. For completeness, they are still reported along with descriptive statistics for the other outcomes, however.

The distribution of episode-level affect scores, along with their means and standard deviations, are shown in Table 1. As can be seen from the table, the typical episode is a pleasant one. The mean of happiness is 4.1, with half of the episodes given a rating of 5 or 6. Sadness and pain were not experienced at all for about three-fourths of all episodes, and a little less than half are not stressful at all. The share of unpleasant episodes, as summarized by the u-index, is 19%. Overall, there is a fair amount of variation in affect ratings, as well as in the u-index and net affect.

Table 2: Episode-level affect score correlations

	Happy	Interested	Tired	Stressed	Sad	Pain	U-index
Interested	0.47						
Tired	−0.10	−0.13					
Stressed	−0.28	−0.10	0.34				
Sad	−0.26	−0.05	0.19	0.46			
Pain	−0.12	−0.05	0.26	0.28	0.34		
U-index	−0.54	−0.20	0.22	0.56	0.45	0.38	
Net affect	0.86	0.38	−0.26	−0.62	−0.58	−0.46	−0.71

$N = 11,469$, $n = 3,823$.

To get a sense of the overlap between different affect outcomes it is also instructive to look at pairwise correlations within episodes, shown in Table 2. These correlations reflect both the affective content of the episode and stable differences in individual well-being. The correlation between the two summary measures, the u-index and net affect, is $r = -0.71$, which is quite strong, but not so strong that either could be considered redundant *ex ante*. The correlation between happiness and any negative emotion is rather weak, e.g. $r = -0.26$ for happy and sad, which is a motivation for studying them separately. Net affect is most strongly correlated with happiness, with $r = 0.86$, whereas the correlations between the u-index and the underlying feelings are rather similar in terms of absolute magnitudes (from $r = 0.38$ for pain up to $r = 0.56$ for stress).

2.3 Time Use Variables

Activities in the PATS are classified as in the ATUS, at three different levels of detail. The episode observations in the sample considered (i.e. those for which affect was also sampled) fall in 240 different categories at the most detailed 6-digit level. A trade-off between a rich characterization of activities and a tractable empirical model must thus be made. The approach taken here is to use the intermediate 4-digit level of aggregation as a benchmark, whereafter a reasonable number of categories is obtained by merging categories with few observations, and categories that could be expected to be affectively similar. In a few cases with a large number of observations, e.g. television watching, I use the more detailed 6-digit classification as a separate category. The resulting classification has 24 categories.[6] The sample distribution of activities according to this classification, ranked by the mean of net affect, is shown in Table A.1 in Appendix A. Work, TV, meals and travel account for around half of the sampled activities. Sports, socializing and meals are examples of activities ranking high in net affect.

Other aspects of time use are accounted for by a set of variables that I refer to as context variables. These are indicators of whether the activity is undertaken at home or somewhere else, with whom (7 levels), in what month, on what day of the week, and at what time of the day (6 levels).[7]

2.4 Individual-Level Variables

The individual-level variables used capture both demographic characteristics and socio-economic circumstances: sex, age, marital status (7 levels), education (7 levels), employment status (5 levels), household income (10 levels) and life satisfaction (4 levels). I refer to this set of variables as characteristics. I use the original answer categories in the PATS survey, except for education and

[6] Although this classification is to some extent arbitrary, one can argue that this choice of activity classification corresponds to a level of detail that is reasonable for other empirical applications. A more detailed classification of activities runs into the problem of very small cell sizes, but would of course fit the data better in a mechanical sense. See also Krueger (2007) for an approach based on cluster analysis.

[7] As for activities, these variables were derived from more detailed classifications. The number of categories was reduced in order to avoid small cells due to overlap with activities.

employment, for which I use fewer categories. All variables, except for age and its square, are used as sets of indicator variables in the estimations.

3 Method

I use a simple mixed-effects framework to estimate what amount of the total variation in affect that can be attributed to observables, on one hand, and unexplained within- and between-variation, on the other hand. I estimate the following model:

$$y_{it} = \alpha + \boldsymbol{w}_{it}'\boldsymbol{\delta} + \boldsymbol{z}_i'\boldsymbol{\gamma} + b_i + \epsilon_{it}, \tag{1}$$

where y_{it} is an affect score for episode $t = 1, 2, 3$, for individual i, $\boldsymbol{w}_{it}$ is a vector of episode-level time-use variables (with coefficients $\boldsymbol{\delta}$), $\boldsymbol{z}_i$ is a vector of individual-level variables (with coefficients $\boldsymbol{\gamma}$), b_i is an individual random effect, and ϵ_{it} is an episode-level error term. The random effects b_i and the errors ϵ_{it} are assumed to be normally distributed and independent, with variances σ_b^2 and σ_ϵ^2, respectively. Following the independence assumption on b_i and ϵ_{it}, we have that $\mathrm{var}\,(y|\boldsymbol{w}, \boldsymbol{z}) = \sigma_b^2 + \sigma_\epsilon^2$. The variance components of the random effects, σ_b^2 and σ_ϵ^2, are estimated from Equation (1) by maximum likelihood,[8] whereafter the share $\widehat{\sigma}_b^2 / \left(\widehat{\sigma}_b^2 + \widehat{\sigma}_\epsilon^2\right)$ is computed to indicate the share of unexplained variation that can be attributed to individual characteristics, i.e. between-variation.

The share of variation accounted for by the covariates is estimated by means of a pseudo coefficient of determination, based on the log-likelihoods of the estimated model and that of a null model.[9] The measure is defined as

$$\overline{R^2} = 1 - \exp\left[-\frac{2}{n}\left(\ell_{ur} - \ell_0\right)\right], \tag{2}$$

where ℓ_{ur} and ℓ_0 denote the log-likelihoods of the unrestricted model and the null model, respectively. The null model used here includes an intercept and the random effects terms b_i and ϵ_{it} only, i.e. it has the same covariance structure as the unrestricted model.

I use a non-parametric clustered bootstrap approach to obtain confidence intervals for σ_b^2, σ_ϵ^2 (and associated variance shares) and $\overline{R^2}$. These statistics are

[8] See e.g. Pinheiro and Bates (2000).

[9] This measure was proposed by Cox and Snell (1989) and Magee (1990). Its properties and interpretation, which are described by Nagelkerke (1991), are similar to that of the usual R^2 measure.

computed for each of $1,000$ data sets resampled at the level of the individual. Thereafter a 95% confidence interval around the original estimate is computed using the distances between the median and the 2.5th and 97.5th percentiles in the bootstrap distributions.

Although the random-effects framework is useful for decomposing the variation in affect into distinct components, the independence assumption regarding the individual effects b_i is not very realistic. In particular, individuals' affective well-being (or response behaviour) could vary systematically with their time use or their characteristics, which would lead to biased estimates of the coefficient vectors $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$. However, this is of somewhat less concern here, as the focus is on a descriptive characterization of the variation in affect, rather than on effect sizes. Moreover, we will see that individuals' general life satisfaction, which is included in some specifications, in fact can be interpreted as a pseudo fixed effect that controls for stable individual well-being. Foreshadowing the results, it turns out that the random-effects assumption appears not to be crucial with respect to time use variables, since the variation attributable to time use remains virtually unchanged when life satisfaction is controlled for. Demographic and socio-economic characteristics are more sensitive to the inclusion of life satisfaction, however.

4 Results

The results for the summary measures u-index and net affect are presented first, followed by separate results for the emotions happiness, sadness, stress and pain, that the summary measures are based on.

4.1 U-Index

To get an overview of the structure of variation in the u-index, Equation (1) is first estimated with the random terms and an intercept only (the null model). As can be seen from column (1) in the upper panel of Table 3, the variance components are estimated with reasonable precision, and variation within individuals accounts for slightly more than two-thirds of the total variation in the u-index ($100 - 30.7 = 69.3\%$).[10] In principle, there is thus much scope for time use variables, which vary both within and between individuals, to account

[10]Standardized affect outcomes are used in all estimations, but the sum of the estimated variance components do not sum to exactly one, presumably due to weighting.

for variation in the u-index. The remaining between-variation in time spent in an unpleasant state could potentially be either due to individual differences in time use or characteristics. Individuals may for example allocate time between leisure and work, or different types of leisure or work, depending on their income. Alternatively, it is possible that some people enjoy any given activity more, e.g. due to better health or a more cheerful personality. It is important to keep in mind that all episode observations for each individual refer to the same day. The between-variation component thus captures both long-run variation between individuals and between-day variation within individuals.

In the second column, the importance of activities is assessed by adding a set of 23 activity indicator variables. Together, these account for 2.8% of the variation in the u-index, as seen from the $\overline{R^2}$-value.[11] In absolute terms, the explanatory power of activities thus appears to be small, which is somewhat discouraging with respect to the overall usefulness of a time-use approach to well-being. Moreover, it can be seen from the relatively larger decline in unexplained within-variance that time use mostly explains within-variation.

To allow for a richer characterization of time use, a set of context variables that measure where, with whom and when the episode took place are added in column (3). The incremental explanatory power of these variables is small, however, only 0.4 percentage points. This reflects the fact that activities and context are strongly correlated. For example, one tends to associate with one's colleagues when working, meaning that these variables to a large extent convey the same information.[12]

Columns (4)–(6) of Table 3 show the same models as columns (1)–(3), but with added individual-level controls for age, sex, marital status, education, employment and household income. By themselves (column 4), these variables account for 1.1% of the variation in the u-index and, by definition, the unexplained between-variation decreases. Among the individual-level variables (results not shown), socio-economic factors (education, employment status and household income) alone account for 0.8% of the variation and are thus more important than demographic variables (age, sex and marital status), which by themselves account for 0.4%. Among the former, income is especially interesting, as it may capture quality aspects of time use not reflected in the

[11] Note that the unexplained variance shares are computed with the covariates partialled out, so that they sum to 100% irrespective of the value of $\overline{R^2}$.

[12] Adding activities and context variables in the reverse order shows that activity alone is a better predictor than context, as context by itself only accounts for 1.3% of the variation in affect. The corresponding figure for net affect is 3.4%.

Table 3: Results: u-index

	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{\sigma}_b^2$	0.30	0.29	0.29	0.28	0.28	0.27
	(0.27, 0.33)	(0.26, 0.32)	(0.26, 0.32)	(0.25, 0.31)	(0.25, 0.31)	(0.24, 0.30)
$\widehat{\sigma}_\epsilon^2$	0.67	0.65	0.65	0.67	0.65	0.65
	(0.64, 0.71)	(0.62, 0.69)	(0.62, 0.69)	(0.64, 0.71)	(0.62, 0.69)	(0.61, 0.68)
$\widehat{\sigma}_b^2$, % share	30.7	31.0	30.8	29.7	29.9	29.7
	(27.9, 33.4)	(27.9, 33.8)	(27.7, 33.5)	(26.9, 32.3)	(27.0, 32.6)	(26.8, 32.3)
$\overline{R^2}$, %		2.8	3.2	1.1	3.9	4.5
		(2.0, 3.6)	(2.5, 4.2)	(0.6, 1.6)	(3.2, 4.9)	(3.7, 5.5)
			Models including life satisfaction			
$\widehat{\sigma}_b^2$	0.26	0.25	0.25	0.25	0.25	0.24
	(0.23, 0.29)	(0.23, 0.28)	(0.22, 0.28)	(0.22, 0.28)	(0.22, 0.27)	(0.21, 0.27)
$\widehat{\sigma}_\epsilon^2$	0.67	0.65	0.65	0.67	0.65	0.64
	(0.63, 0.70)	(0.61, 0.68)	(0.61, 0.68)	(0.63, 0.70)	(0.61, 0.68)	(0.61, 0.68)
$\widehat{\sigma}_b^2$, % share	28.0	28.2	28.0	27.4	27.5	27.2
	(25.2, 30.5)	(25.3, 30.8)	(25.1, 30.5)	(24.7, 29.9)	(24.7, 30.1)	(24.5, 29.8)
$\overline{R^2}$, %	2.9	5.6	6.1	3.5	6.3	6.8
	(2.2, 3.6)	(4.7, 6.7)	(5.2, 7.2)	(2.8, 4.3)	(5.4, 7.5)	(5.9, 8.0)
Activities		✓	✓		✓	✓
Context			✓			✓
Characteristics				✓	✓	✓

$N = 11,469$, $n = 3,823$. $\widehat{\sigma}_b^2$, $\widehat{\sigma}_\epsilon^2$, denote between- and within-variance components. The between %-share, $100 \cdot \widehat{\sigma}_b^2/(\widehat{\sigma}_b^2 + \widehat{\sigma}_\epsilon^2)$, is also shown. Bootstrapped 95% c.i.'s in parentheses. Activities is a set of 23 indicator variables. Context includes indicators for where, with whom and when the episode took place. Characteristics include sex, age, age squared, education level, marital status, employment status and income category.

Table 4: Results: net affect

	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{\sigma}^2_b$	0.4	0.4	0.4	0.4	0.4	0.4
	(0.41, 0.47)	(0.40, 0.46)	(0.39, 0.46)	(0.39, 0.45)	(0.38, 0.44)	(0.37, 0.43)
$\widehat{\sigma}^2_\epsilon$	0.5	0.5	0.5	0.5	0.5	0.5
	(0.47, 0.52)	(0.44, 0.49)	(0.43, 0.48)	(0.47, 0.52)	(0.43, 0.48)	(0.43, 0.47)
$\widehat{\sigma}^2_b$, % share	47.4	48.5	48.4	46.3	47.2	47.1
	(44.8, 49.7)	(45.9, 51.0)	(46.0, 50.9)	(43.7, 48.6)	(44.7, 49.6)	(44.6, 49.4)
$\overline{R^2}$, %		5.7	7.3	1.3	7.1	8.7
		(4.9, 6.7)	(6.3, 8.3)	(0.9, 1.8)	(6.2, 8.1)	(7.6, 9.7)
			Models including life satisfaction			
$\widehat{\sigma}^2_b$	0.4	0.4	0.3	0.4	0.3	0.3
	(0.34, 0.39)	(0.33, 0.38)	(0.32, 0.37)	(0.33, 0.38)	(0.32, 0.37)	(0.31, 0.36)
$\widehat{\sigma}^2_\epsilon$	0.5	0.5	0.4	0.5	0.5	0.4
	(0.46, 0.51)	(0.43, 0.48)	(0.43, 0.47)	(0.46, 0.51)	(0.43, 0.48)	(0.43, 0.47)
$\widehat{\sigma}^2_b$, % share	42.8	43.7	43.6	42.4	43.2	43.0
	(40.5, 45.1)	(41.4, 46.1)	(41.3, 45.9)	(40.0, 44.5)	(40.7, 45.4)	(40.6, 45.4)
$\overline{R^2}$, %	4.7	10.3	11.7	5.2	10.9	12.4
	(4.0, 5.4)	(9.3, 11.4)	(10.7, 12.9)	(4.4, 6.0)	(9.8, 12.0)	(11.3, 13.5)
Activities		✓	✓		✓	✓
Context			✓			✓
Characteristics				✓	✓	✓

$N = 11,469$, $n = 3,823$. $\widehat{\sigma}^2_b$, $\widehat{\sigma}^2_\epsilon$, denote between- and within-variance components. The between %-share, $100 \cdot \widehat{\sigma}^2_b/(\widehat{\sigma}^2_b + \widehat{\sigma}^2_\epsilon)$, is also shown. Bootstrapped 95% c.i.'s in parentheses. Activities is a set of 23 indicator variables. Context includes indicators for where, with whom and when the episode took place. Characteristics include sex, age, age squared, education level, marital status, employment status and income category.

ATUS classifications and context variables. Higher income makes it possible to eat better food at any given meal, for example (Krueger et al., 2009). Income accounts for 0.4% of the total variation in the u-index, and thus half of what is accounted for by all socio-economic factors, but in absolute terms, its explanatory power must be considered low.[13]

The added explanatory power of time use variables (columns 5 and 6) is similar to before, which is expected given that they were found to mostly explain within-variation. This fact can also be interpreted as that there is not much selection into activities based on individual characteristics. Relatively speaking, time use accounts for almost three times as much variation as the individual-level variables, and taken together, they account for 4.5% of the variation in the u-index.

The lower panel of Table 3 has the same structure as the top panel, but with added controls for life satisfaction. As expected, higher life satisfaction is associated with a lower probability of experiencing negative affect. The estimated coefficients for "not satisfied", "satisfied" and "very satisfied", without any other controls, are -0.11, -0.30 and -0.39 (results not shown in table).[14] On average, persons with higher life satisfaction thus have a substantially lower probability of experiencing negative affect. The explanatory power of life satisfaction alone is moderate, however—the $\overline{R^2}$ is 2.9%, which is on par with the $\overline{R^2}$ for activities, but almost three times as much as for the other individual-level factors together.

The life satisfaction estimates should be interpreted carefully, however. First, one should keep in mind that the unit of observation is an episode within a day, rather than an individual. The estimated relationship is thus between life satisfaction and momentary affective well-being. Given that there is substantial within-individual variation in affect, one could be tempted to interpret the estimated association as a lower bound on the association between life satisfaction and individual affective well-being, appropriately aggregated over a longer period.[15]

[13]The income association is mostly driven by those with incomes less than $10,000 and is basically flat beyond incomes of $50,000. This is in line with the findings of Kahneman and Deaton (2010), although they find a satiation point at $75,000.

[14]Since the u-index is a binary variable, the coefficients should be interpreted in terms of the probability of experiencing a predominantly negative emotion, relative to the reference category "not at all satisfied".

[15]This interpretation is supported by the fact that the explanatory power of life satisfaction increases by 89% when the outcome is the the average daily u-index instead of the episode-level u-index (based on a comparison of regular R^2-values from OLS). The corresponding

However, the quantitative interpretation is further complicated by the possibility of correlated measurement errors, given that the respondent was asked about affect and life satisfaction at the same occasion. If the respondent was in a particularly good or bad mood at the time of the interview, and if this affected both the affect and life satisfaction reports in the same direction, then the correlation between the two measures would be inflated.

Although the magnitudes of the estimates relating to life satisfaction should be taken with a grain of salt, it was mentioned previously that life satisfaction can be interpreted as a pseudo fixed effect that controls for stable individual differences in well-being. It is therefore interesting to see how the explanatory power of time use and individual characteristics change when life satisfaction is controlled for. By contrasting columns (2) and (3) in the upper and lower panel, we see that the explanatory power of activities and context is almost completely additive to that of life satisfaction, which is again expected due to time use mostly explaining within-variation. Nevertheless, this suggests that there is no selection into more pleasant activities based on life satisfaction, consistent with the results for the other individual-level variables. That individuals who are more satisfied with their life also have fewer unpleasant daily experiences can hence not be explained in terms of how they spend their time, at least according to the current model framework. In other words, time use and life satisfaction capture distinct variation in affective well-being.

In contrast to the time use variables, the explanatory power of individual characteristics (columns 4–6, lower panel) is not additive to that of life satisfaction. Whereas these alone account for 1.1% of the variation in the u-index (column 4, upper panel), their marginal explanatory power when life satisfaction is controlled for is only 0.6% (columns 1 and 4, lower panel). A plausible interpretation is that demographic and socio-economic variables only influence emotional well-being insofar as they influence life satisfaction—a channel which is well-documented in the research on the determinants of life satisfaction.

4.2 Net Affect

The results for net affect, which is defined as the score of "happy" minus the average score of "sad", "stressed" and "pain", are presented in Table 4. For this outcome, the total variation is split almost evenly across between- and

increase for net affect is 53%.

within-variation.[16] Given that net affect uses the whole range of variation in the underlying emotions, it is difficult to compare it directly to the u-index. It is, for example, hard to say whether there is actually more systematic variation between individuals when the full range of well-being is considered, or whether this simply reflects individual differences in how the response scales are used.

From column (2) we see that activities account for 5.7% of the variation in net affect, i.e. twice as much as for the u-index. The added explanatory power of context (column 3) is 1.6%. Individual characteristics alone account for 1.3% of the total variation (column 4), and is additive to time use (columns 5 and 6), just as for the u-index.

Life satisfaction accounts for 4.7% of the total variation (column 1, lower panel), which is substantially more than for the u-index in absolute terms, but relative to the explanatory power of time use within the same outcome, the results are similar. The remaining results for net affect are also similar to those for the u-index: the explanatory power of time use is almost completely additive to that of life satisfaction (columns 2 and 3, upper and lower panel), though somewhat less so for context variables, whereas demographic and socio-economic variables are largely redundant once life satisfaction is controlled for. It is quite natural that the explanatory power in general is higher for net affect than for the u-index, given that the latter is constrained to be either zero or one. Besides this, the pattern of variation is strikingly similar across the two measures.[17]

4.3 Happiness, Sadness, Stress and Pain

The same models as for the u-index and net affect were run for the outcomes happiness, sadness, stress and pain. Table A.2 and Table A.3 in Appendix A correspond (in compact form) to the upper and lower panels of Table 3 and Table 4, respectively, i.e. without and with controls for life satisfaction. The

[16] This is consistent with the findings of White and Dolan (2009) based on a German sample, who find within- and between- components to account for 55% and 45%, respectively, of an affect balance measure in a model conditional on activities. The corresponding figures were 59% and 41% when duration-weighted responses were considered.

[17] A third summary measure was also derived as the principal component from a factor analysis of happiness, sadness, stress and pain, with corresponding factor loadings of 0.40, −0.57, −0.56 and −0.45. This variable was strongly correlated with net affect ($r = 0.90$) and, as expected, it also showed a similar pattern of variation when analysed as a separate outcome, although the model fit was slightly worse. For instance, activities alone accounted for 4.2% of the variation, characteristics for 2.2%, life satisfaction for 4.6%, and the full model accounted for 11.0% of the variation. 55.7% of the unconditional variation was between-variation.

pattern of variation for happiness is most similar to that of net affect, as one might have expected given the strong direct correlation between the two measures. The main difference is that individual characteristics and life satisfaction account for less variation in happiness. Time use accounts for 7.6% of the total variation, whereas demographic and socio-economic variables together only account for 0.5%, and life satisfaction accounts for 2.5%.

With respect to time use and individual characteristics, this pattern is reversed for the two negative outcomes sadness and pain. Time use only accounts for 1.6% and 2.0% of the variation in sadness and pain, respectively, whereas individual characteristics account for 2.1% and 4.0%. The explanatory power of time use is additive to that of characteristics, as was also the case for the u-index and net affect.

Moreover, individual characteristics have substantial independent explanatory power even when life satisfaction, which itself accounts for 2.9% of the variation in sadness and 1.9% in pain, is controlled for. Separate regressions (results not shown) reveal that employment status is the single most important individual-level variable with respect to pain, accounting for 2.1% of the variation. Not surprisingly, disabled persons experience more pain relative to those in employment, but to less extent this is also true for the retired and others not in employment. There is also a negative association between pain and income, mostly driven by the group with lowest income.

Finally, stress is the outcome which is best explained by time use variables, which account for 7.9% of the total variation. In comparison, individual characteristics and life satisfaction account for 1.3% and 2.1% of the variation in stress. With the exception of stress, time use thus does a better job of explaining variation in positive affect than in negative affect, both in absolute terms and relative to demographic and socio-economic characteristics within the same measure. A possible explanation for this result is that individuals are able to avoid many unpleasant activities, especially painful ones. Remaining feelings of sadness and pain might then instead be related to individuals' general circumstances or characteristics.

5 Discussion

I have shown how variation in affective well-being can be accounted for by individuals' time use and their characteristics, and how these results are affected when controlling for general life satisfaction. Time use is primarily found to capture within-variation in affect, and hence it is distinct from the variation captured by individual characteristics, including life satisfaction. The positive association between affect and life satisfaction can thus not be explained in terms of selection into more pleasant activities. Activities were generally found to be more important for affect than the context in which they took place. Time use was also found to be more important than individual characteristics for the u-index and for net affect, happiness and stress, whereas the opposite was true for sadness and pain. In absolute terms, the explanatory power of both time use and characteristics was found to be quite low, however, despite using a rich set of covariates.

Intuitively, how people spend their time should be important for their well-being, so how should these results be interpreted? The intuition that time use matters may still be correct, for a number of reasons. First, it is possible that important quality aspects of time use are largely unobserved. An interesting example is provided by Killingsworth and Gilbert (2010), who find that "mind wandering" is a better predictor of momentary well-being than activities. Besides raising the obvious question of what determines mind wandering, their results point more generally to the importance of internal, rather than external, aspects of time use. Second, it might be the case that the set of well-being outcomes considered here is too narrow. This point is highlighted by White and Dolan (2009), who find that time use accounts for less variation in affect measures similar to those in this paper ("pleasure"), than in alternative measures of meaningfulness ("reward").[18] Third, even if time use—in an absolute sense—is rather unimportant for well-being in the short run, it may at the same time be important in the long run. Many activities of course serve the purpose of investing in future well-being, either in the short or long run, e.g. washing dishes or studying. More generally, there is reason to question the sensibility of the independence assumption in the theoretical framework proposed by Juster et al. (1985), that Krueger et al. (2009) builds on. To give a concrete example, in line with the findings of Knabe et al. (2010), watching TV after a day of work or after a day of staying at home unemployed is probably not worth the same in terms of momentary well-being. Hence, future research may consider a less mechanical approach to time use than the accounting framework used in

[18] White and Dolan (2009) estimate time use models with pleasure and reward as separate outcomes, and report R^2-values of 0.07 and 0.23 (or 0.09 and 0.35 for duration-weighted outcomes), respectively.

this paper, perhaps by studying between-person differences in time use over a longer period.

That time use is a worse predictor of negative feelings (with the exception of stress) than for positive feelings, and a worse predictor of the u-index compared to net affect, is interesting in itself, and especially so given one of the motivations for the u-index. Quoting Kahneman and Krueger (2006): "we suspect that many policymakers are more comfortable with the idea of minimizing a specific concept of misery than maximizing a nebulous concept of happiness". Unfortunately, the results of this paper suggest that the time-accounting approach to well-being may be especially limited with respect to negative experiences. A likely explanation is that individuals to a large extent are able to avoid unpleasant activities, highlighting the general problem of deriving useful information from time-use data when individuals self-select into activities. For instance, most people should be able to avoid particular activities associated with physical pain. For the purpose of studying negative experiences, it may thus be more fruitful to focus on "states" and general life circumstances (e.g. health, economic conditions, social status), rather than day-to-day activities.

Finally, two major limitations of this study should be mentioned. First, although the affect data analysed here refer to specific episodes, they were collected in a retrospective survey during the following day rather than in the moment. It is thus possible that one would find different patterns if one instead used methods for measuring affect in the moment. A second limitation is the lack of repeated affect data for the same individuals over a longer period of time. As such, it is only possible to provide a partial characterization of the variation in affective well-being. For example, one would like to know how representative a randomly sampled day is, with respect to an individual's long-term well-being. It is also somewhat of an open question how well life satisfaction correlates with momentary well-being aggregated over a longer time period. Collecting such data should therefore be a priority for future research.

References

Busseri, M. and S. Sadava (2011). A review of the tripartite structure of subjective well-being: Implications for conceptualization, operationalization, analysis, and synthesis. *Personality and Social Psychology Review 15*(3), 290–314.

Connolly, M. (2013). Some like it mild and not too wet: The influence of weather on subjective well-being. *Journal of Happiness Studies 14*(2), 457–473.

Cox, D. and E. Snell (1989). *Analysis of binary data.* London: Chapman & Hall.

Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social Indicators Research 31*, 103–157.

Diener, E., E. Suh, R. Lucas, and H. Smith (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin 125*(2), 276–302.

Juster, F. and F. Stafford (Eds.) (1985). *Time, goods, and well-being.* Ann Arbor, MI: Institute for Social Research, University of Michigan.

Juster, F. T., P. N. Courant, and G. K. Dow (1985). A conceptual framework for the analysis of time allocation data. In F. Juster and F. Stafford (Eds.), *Time, goods, and well-being*, pp. 333–351. Ann Arbor, MI: Institute for Social Research, University of Michigan.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener, and N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*, Chapter 1, pp. 3–25. New York: Russell Sage Foundation.

Kahneman, D. and A. Deaton (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences 107*(38), 16489–16493.

Kahneman, D. and A. Krueger (2006). Developments in the measurement of subjective well-being. *The Journal of Economic Perspectives 20*(1), 3–24.

Kahneman, D., A. Krueger, D. Schkade, N. Schwarz, and A. Stone (2004a). A survey method for characterizing daily life experience: The day reconstruction method. *Science 306*(5702), 1776–1780.

Kahneman, D., A. B. Krueger, D. Schkade, N. Schwarz, and A. Stone (2004b). Toward national well-being accounts. *American Economic Review 94*(2), 429–434.

Kahneman, D., D. Schkade, C. Fischler, A. Krueger, and A. Krilla (2010). The structure of well-being in two cities: Life satisfaction and experienced happiness in Columbus, Ohio; and Rennes, France. In E. Diener, D. Kahneman, and J. Helliwell (Eds.), *International differences in well-being*, Chapter 2, pp. 16–33. New York: Oxford University Press.

Killingsworth, M. A. and D. T. Gilbert (2010). A wandering mind is an unhappy mind. *Science 330*(6006), 932–932.

Knabe, A., S. Rätzel, R. Schöb, and J. Weimann (2010). Dissatisfied with life but having a good day: Time-use and well-being of the unemployed. *The Economic Journal 120*(547), 867–889.

Krueger, A. (2007). Are we having more fun yet? Categorizing and evaluating changes in time allocation. *Brookings Papers on Economic Activity 2007*(2), 193–215.

Krueger, A., D. Kahneman, D. Schkade, N. Schwarz, and A. Stone (2009). National time accounting: The currency of life. In A. Krueger (Ed.), *Measuring the subjective well-being of nations: National accounts of time use and well-being*, pp. 9–86. University of Chicago Press.

Larson, R. and M. Csikszentmihalyi (1983). The experience sampling method. In H. T. Reis (Ed.), *New directions for methodology of social & behavioral science*, Volume 15, pp. 41–56. San Francisco: Jossey-Bass.

Magee, L. (1990). R2 measures based on wald and likelihood ratio joint significance tests. *The American Statistician 44*(3), 250–253.

Nagelkerke, N. (1991). A note on a general definition of the coefficient of determination. *Biometrika 78*(3), 691–692.

Pinheiro, J. and D. Bates (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer Verlag.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology 39*(6), 1161–1178.

Schwarz, N. and F. Strack (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. Diener, and N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*, Chapter 4, pp. 61–84. New York: Russell Sage Foundation.

Stone, A. and S. Shiffman (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine 16*, 199–202.

White, M. and P. Dolan (2009). Accounting for the richness of daily activities. *Psychological Science 20*(8), 1000–1008.

A Tables

Table A.1: Distribution and mean net affect of activities

	Mean affect	Frequency, %	Obs.[a]
Sports	4.3	2.8	289
Religious activities	4.2	1.3	149
Caring for pets	4.2	0.9	100
Socializing	3.8	4.4	519
Eating and drinking	3.8	10.3	1170
Other leisure	3.6	7.0	758
Reading	3.5	3.5	465
Caring for others	3.5	4.2	438
Volunteering	3.4	0.4	51
Garden work	3.4	2.4	312
Shopping	3.3	2.8	336
Phone calls	3.2	1.0	126
Traveling	3.0	10.2	1120
TV	3.0	15.3	1890
Food preparation	2.9	4.7	580
Personal care	2.8	2.4	277
Work	2.6	15.6	1681
Repairs and maintenance	2.5	1.2	127
Housework	2.5	4.6	524
Using services	2.4	0.6	73
Household management	2.3	1.9	229
Other	2.2	0.1	19
Education	2.2	1.9	142
Health care	1.8	0.7	94
Sum		100	11,469

[a] Number of observations are unweighted. Activities are ranked by mean net affect.

Table A.2: Results, specific affect measures

	(1)	(2)	(3)	(4)	(5)	(6)
HAPPY						
$\widehat{\sigma}_b^2$, % share	41.1	41.8	41.6	40.6	41.3	41.0
	(38.8, 43.5)	(39.5, 44.1)	(39.2, 43.8)	(38.4, 43.0)	(39.0, 43.6)	(38.7, 43.3)
$\overline{R^2}$, %		5.7	7.6	0.5	6.1	8.1
		(4.8, 6.6)	(6.6, 8.6)	(0.2, 0.7)	(5.3, 7.1)	(7.1, 9.2)
SAD						
$\widehat{\sigma}_b^2$, % share	49.0	49.2	49.2	47.0	47.2	47.1
	(45.7, 52.3)	(45.9, 52.5)	(45.8, 52.6)	(43.7, 50.4)	(44.0, 50.6)	(43.8, 50.5)
$\overline{R^2}$, %		1.0	1.6	2.1	3.1	3.7
		(0.6, 1.4)	(1.1, 2.1)	(1.6, 2.7)	(2.4, 3.8)	(2.9, 4.5)
STRESS						
$\widehat{\sigma}_b^2$, % share	45.2	45.9	45.8	43.8	44.7	44.6
	(42.9, 47.5)	(43.7, 48.3)	(43.5, 48.1)	(41.7, 46.2)	(42.4, 47.1)	(42.4, 47.0)
$\overline{R^2}$, %		7.2	7.9	1.3	8.4	9.1
		(6.2, 8.2)	(6.9, 9.0)	(1.0, 1.8)	(7.3, 9.5)	(8.0, 10.2)
PAIN						
$\widehat{\sigma}_b^2$, % share	66.3	66.9	66.9	63.0	63.6	63.7
	(63.7, 68.9)	(64.4, 69.4)	(64.4, 69.5)	(60.3, 65.7)	(61.1, 66.4)	(61.1, 66.3)
$\overline{R^2}$, %		1.8	2.0	4.0	5.7	5.9
		(1.2, 2.5)	(1.4, 2.8)	(3.2, 4.7)	(4.8, 6.6)	(4.9, 6.9)
Activities		✓	✓		✓	✓
Context			✓			✓
Characteristics				✓	✓	✓

$N = 11,469$, $n = 3,823$. $\widehat{\sigma}_b^2$ denotes unexplained between-variance component, presented as the %-share of total unexplained variation. Bootstrapped 95% c.i.'s in parentheses. Activities is a set of 23 activity indicators. Context includes indicators for where, with whom and when the episode took place. Characteristics include sex, age, age squared, education level, marital status, employment status and income category.

Table A.3: Results, specific affect measures, controlling for life satisfaction

	(7)	(8)	(9)	(10)	(11)	(12)
HAPPY						
$\widehat{\sigma}_b^2$, % share	38.6	39.2	39.0	38.2	38.8	38.5
	(36.2, 40.9)	(37.0, 41.4)	(36.7, 41.2)	(35.8, 40.5)	(36.4, 41.0)	(36.1, 40.7)
$\overline{R^2}$, %	2.5	8.1	9.9	2.9	8.5	10.4
	(2.1, 3.1)	(7.1, 9.2)	(9.0, 11.0)	(2.4, 3.5)	(7.5, 9.5)	(9.4, 11.5)
SAD						
$\widehat{\sigma}_b^2$, % share	46.0	46.3	46.3	44.8	45.0	45.0
	(42.7, 49.4)	(42.9, 49.5)	(42.9, 49.5)	(41.5, 48.1)	(41.7, 48.3)	(41.6, 48.3)
$\overline{R^2}$, %	2.9	3.8	4.4	4.2	5.2	5.7
	(2.2, 3.7)	(3.0, 4.6)	(3.6, 5.3)	(3.5, 5.1)	(4.4, 6.2)	(4.9, 6.8)
STRESS						
$\widehat{\sigma}_b^2$, % share	43.0	43.4	43.3	41.9	42.6	42.5
	(40.8, 45.4)	(41.1, 45.8)	(40.9, 45.6)	(39.6, 44.3)	(40.3, 44.9)	(40.2, 44.9)
$\overline{R^2}$, %	2.1	9.4	10.1	3.2	10.3	10.9
	(1.7, 2.7)	(8.3, 10.6)	(9.0, 11.2)	(2.7, 3.9)	(9.1, 11.4)	(9.7, 12.1)
PAIN						
$\widehat{\sigma}_b^2$, % share	64.9	65.5	65.5	62.1	62.8	62.8
	(62.3, 67.6)	(62.9, 68.0)	(62.9, 68.1)	(59.4, 64.9)	(60.2, 65.5)	(60.2, 65.4)
$\overline{R^2}$, %	1.9	3.7	3.9	5.0	6.8	7.0
	(1.4, 2.5)	(3.0, 4.5)	(3.1, 4.8)	(4.3, 5.9)	(5.8, 7.8)	(6.0, 8.0)
Activities		✓	✓		✓	✓
Context			✓			✓
Characteristics				✓	✓	✓

$N = 11,469$, $n = 3,823$. $\widehat{\sigma}_b^2$ denotes unexplained between-variance component, presented as the %-share of total unexplained variation. Bootstrapped 95% c.i.'s in parentheses. Activities is a set of 23 indicator variables. Context includes indicators for where, with whom and when the episode took place. Characteristics include sex, age, age squared, education level, marital status, employment status and income category.

Study 3

The Association Between Life Satisfaction and Affective Well-Being*

1 Introduction

A key question for both research and policy initiatives that assess individual welfare in terms of subjective well-being (SWB), is whether it primarily should be conceptualized and measured in terms of life satisfaction or affective well-being. According to the established definition of SWB, the former refers to a cognitive evaluation of one's life as a whole, and the latter to more specific experiences of positive and negative feelings (Diener, 1984).[1] A common view among psychologists is that these components are separate constructs that should be studied on their own terms (e.g. Diener et al., 1999), but there is no established theory of how the components relate to each other.[2]

The multi-dimensionality of SWB—or at least its implications—has received relatively little attention within the subfield of "happiness economics", however.[3] There are several possible reasons for this. First, it could perhaps

*This paper is joint work with Filip Fors. We want to thank Göran Landgren for invaluable help with programming the survey, which this paper is based on.

[1]Some definitions of SWB also include meaningfulness (or eudaimonia) as a separate component. This is not discussed further in this paper, but see e.g. Haybron (2016) for a discussion, and White and Dolan (2009) for an application.

[2]See Busseri and Sadava (2011) for an overview of different theoretical models, and Lucas (2016) for a discussion.

[3]In an early survey in the economics literature, Frey and Stutzer (2002) note the distinc-

be attributed to an intuitive understanding of happiness as one-dimensional. After all, it is hard to imagine that a person whose daily life is characterized by negative feelings, is also satisfied with his or her life as a whole, and vice versa. Second, economists have been inclined to interpret SWB data as an empirical proxy for utility, i.e. an index of preference fulfillment, which is one-dimensional by construction. A third reason is practical—there are few data sources for affect that are based on population samples, whereas several ones exist for life satisfaction, either in the form of national socio-economic surveys (e.g. the US General Social Survey and the German Socio-Economic Panel) or international surveys (e.g. the World Values Survey and the European Social Survey). Consequently, most applied research has studied the determinants of life satisfaction, but the results have often been framed generically in terms of happiness or SWB.[4]

One-dimensionality of SWB is in fact one of its major selling points, as it is necessary for unambiguous welfare comparisons across individuals (given interpersonal comparability) and across countries and over time (given a social welfare welfare function). A one-dimensional measure of SWB is particularly useful for public policy purposes, as it provides a clear framework for cost-benefit analysis, in which the marginal well-being increase of a given policy can be weighed against the marginal costs, to yield a welfare-maximizing outcome.[5] As such, (one-dimensional) SWB is not subject to the problem inherent to multi-dimensional approaches to welfare, of assigning welfare weights to different components, such as health and income.

If SWB were instead genuinely multi-dimensional—in the sense of life satisfaction and affective well-being being only weakly or or moderately associated with each other, as well as having different determinants—one would either

---

tion between life satisfaction and affect, but do not discuss any implications. A later survey by Di Tella and MacCulloch (2006) does not take note of the distinction at all. The issue has been highlighted e.g. by Adler (2013) and Clark (2016).

[4]Confusion about the terminology may partly stem from the fact that global SWB is elicited with questions phrased in terms of "happiness" in some surveys (as in the GSS). This phrasing is not wrong per se, but when such questions refer to one's life as a whole or life in general, in contrast to a narrow time frame, they should properly be categorized as measures of life satisfaction and not affect.

[5]More precisely, the marginal rate of substitution of two goods/outcomes can be equated with the ratio of their respective costs, or one can express the well-being impact of some good/outcome in terms of a SWB "money-metric". Needless to say, there are huge challenges for such analyses, e.g. when it comes to establishing causal effects on well-being. In any case, a SWB-based approach has some clear advantages compared to approaches based on revealed or stated preferences. These advantages have been discussed extensively elsewhere, see e.g. Dolan and Fujiwara (2016), for a recent discussion.

have to concede the advantages just described, or take an explicit normative stance of whether to emphasize evaluative or affective aspects of well-being.[6]

It thus appears central to establish the extent of empirical overlap between life satisfaction and affective well-being. Previous research is not very clear on this point, however. There is substantial variation in previous estimates, but several studies (discussed further in Section 2 below) have found only a moderate correlation between individual-level life satisfaction and affect. As noted above, there are relatively few studies on the determinants (or correlates) of affective well-being, but some of these point to differences with respect to the determinants of life satisfaction, e.g. Kahneman and Deaton (2010) and Kahneman et al. (2010), who find that income is more strongly associated with life satisfaction than it is with affective well-being, and Knabe et al. (2010), who find that unemployment is more strongly associated with life satisfaction than with affect.

However, we believe that it is premature to rule out the possibility that discrepancies between life satisfaction and affect, at least to a large extent, are driven by measurement issues, rather than genuine differences between the two variables. In particular, we identify four measurement issues in previous estimates of the association between life satisfaction and affect.

First, affect measures are likely to pick up short-term fluctuations in well-being, as they typically refer to a relatively narrow time frame (e.g. "right now", or the previous day or week), whereas this is not the case for life satisfaction measures, which implicitly refer to a more stable condition (e.g. satisfaction "these days"). Our view, which is central for this paper, is that an affect-based measure of individual welfare should capture the individual's stable, or long-run, affective well-being.[7] Short-run measures of affect are at best noisy

[6]Notably, Kahneman (1999) has taken an extreme stance on this matter, by arguing that life satisfaction is basically an unreliable and biased account of affect. Consequently, Kahneman has referred to evaluative and affective well-being in terms of remembered and experienced utility. Although he has subsequently nuanced this position, with respect to the normative relevance of evaluative well-being (Kahneman and Riis, 2005), he has continued to emphasize discrepancies between evaluative and affective well-being (e.g. Kahneman et al., 2010).

[7]This point has been made previously by Campbell et al. (1976), and highlighted by Eid and Diener (2004). One could argue that this is more of a conceptual point, rather than a measurement issue, but given that the point should be quite uncontroversial, we think that it is appropriate to frame the problem in terms of measurement. Note also that we do not argue that intra-individual variation in affect has no normative relevance, but rather that such variation is of second-order importance. It may e.g. be the case that among two time-profiles of well-being with the same average, one is preferrable to the other. A parallel can also be drawn to the distinction between current and lifetime (permanent) income (see e.g.

measures of long-run affect, and the use of such measures can thus be expected to cause attenuated correlation estimates, given that the association of interest is between life satisfaction and long-run affect.

Second, many studies do not take into account measurement error in self-reports of life satisfaction and affect, which causes correlations to be biased towards zero. Previous research has shown that reliability-adjusted correlation estimates, based on multiple items and/or repeated measurements, are markedly stronger than unadjusted estimates (Eid and Diener, 2004; Krueger and Schkade, 2008; Schimmack et al., 2002).

Third, affect measures are often based on retrospective questions, and are therefore prone to various forms of recall bias (Robinson and Clore, 2002). As it is hardly possible to remember exactly how one felt moment-to-moment in the past, respondents need to "guesstimate" how they felt. If this process is unbiased, but with a white-noise error, the resulting measure would have lower reliability compared to one for which such estimation is not necessary. The resulting correlation with life satisfaction would thus be downward biased, as noted in the previous point. Another possibility is that such a retrospective assessment is mixed up with a cognitive evaluation of the period in question, rather than a retrieval of past affective experiences. In this case, one would instead expect an upward bias in the correlation between such a measure and life satisfaction. The problem of recall bias is the main reason why momentary measurement—as in the Experience Sampling Method (ESM; Larson and Csikszentmihalyi, 1983) and in the Ecological Momentary Assessment (EMA; Stone and Shiffman, 1994)—is usually regarded as the gold standard for measuring affective well-being.

Note that this third point, together with the first point above, form a paradox, or trade-off—a valid measure of experienced affect should, as far as possible, be momentary, in order to reduce recall bias and avoid conflation with evaluative well-being. But momentary measures are by construction only snapshots of an individual's life, and can therefore not be reliable summary measures of individual well-being. In our view, the only way to solve this dilemma is by using repeated, momentary measurements.

Fourth, life satisfaction judgements have been shown to be biased by current mood and situational variables, such as the question order and the weather at the time of the survey (e.g. Connolly, 2013; Deaton, 2012; Schwarz et al., 1987;

Haider and Solon, 2006 regarding measurements issues in this context).

Schwarz and Strack, 1999; but see also Lucas, 2016 for a critical discussion). Hence, the correlation between life satisfaction and affect can be expected to be upward-biased if they are both measured on the same survey occasion.

The purpose of the present paper is to estimate the association between life satisfaction and long-run affective well-being, simultaneously taking into account all of these problems. Although the sign of the net bias from these issues is unclear, our overall hypothesis is that correlation estimates will be stronger when these issues are properly accounted for, compared to "naive" estimates for which this is not the case. We investigate this hypothesis within a simple measurement model framework, in which SWB is decomposed into a component capturing stable individual well-being, and another one capturing measurement error and temporary fluctuations. To estimate the model, we conducted a smart-phone based experience-sampling survey—on a population sample of Swedes aged 18–50 (n = 252)—in which respondents' momentary well-being was measured repeatedly during a seven-week period.

Previewing the results, our main correlation estimates range between 0.78 to 0.91, thus indicating a strong convergence between life satisfaction and affect. Our estimates are markedly stronger compared to estimates that do not account for measurement issues, both in previous literature and based on our own data.

The remainder of this paper is structured as follows. In the next section, we review previous literature about the association between life satisfaction and affect. In Section 3, we outline the model and how we estimate it, whereafter we describe the survey and the data in Section 4. Our main results, on the satisfaction–affect correlation, are presented in Section 5. In Section 6, we present additional results on how life satisfaction and affect differ in terms of their socio-economic correlates. We discuss our results further and conclude in Section 7.

2 Previous Literature

In this section, we review previous research on the association between life satisfaction and affective well-being. Some studies indicate that this association differs across countries, and in particular that it is stronger in individualistic cultures, compared to collectivistic ones (Suh et al., 1998; Schimmack et al., 2002; Kuppens et al., 2008). We limit our review to studies from individualistic/Western countries, however, consequent with that the current study is set in Sweden. We also limit our review to results on the direct correlation between life satisfaction and affect, rather than differences in their correlates, but we return to the latter question in Section 6.

Estimates of the correlation between life satisfaction and affect that are not adjusted for measurement error typically range between 0.3 and 0.6. Previous studies suffer to varying degrees from the four measurement problems identified in the introduction, and the variation in previous estimates can partly be understood in light of these. Using longer time frames for the affect questions, or asking "trait-like" questions about the frequency of affect experienced in general, tends to produce stronger correlations, compared to using a narrow time frame. Studies that explicitly account for measurement error also tend to obtain stronger correlations.

Part of the variation relates to other method differences, however, and in particular to the specific choice of SWB measures used. Life satisfaction usually correlates more strongly with positive affect than with negative affect. Aggregated measures, based on averaging several affect items (questions), tend to produce stronger correlations than single items. Such averaging is in fact an indirect way of dealing with measurement error. Importantly, measures of *hedonic balance* or *net affect*, that capture the balance between positive and negative affect—e.g. by subtracting the mean of a set of negative affect items from the mean of a set of positive affect items—typically generate stronger correlations with life satisfaction. This makes intuitive sense, to the extent that people's full spectrum of emotional well-being matters for their life satisfaction (and/or if life satisfaction influences both positive and negative affect). Measures of hedonic balance are not always used, however.

Many studies are based on convenience samples, e.g. university students, but it is not clear from our review if the satisfaction–affect association varies systematically across different types of samples. Finally, a non-negligible part of the heterogeneity in previous estimates can probably be attributed to random sample variation.

Starting with the work that does not account for measurement error, Lucas et al. (1996) study the association between various measures of life satisfaction and affect, elicited from three different student samples. Life satisfaction was measured with the five-item Satisfaction With Life Scale (SWLS; Diener et al., 1985, see further description in Section 4.4) and positive and negative affect, respectively, were measured with the ten-item PANAS scales (Watson et al., 1988). Their estimates of the correlation between satisfaction and positive affect range between 0.42 and 0.52, and those between satisfaction and negative affect range between −0.51 and −0.30. Affect was also measured with the forty-item Affect Balance Scale (Derogatis, 1975), producing somewhat stronger correlations (at most 0.65 and −0.58, for positive and negative affect, respectively).

Studies based on population surveys do not appear to yield systematically different results, per se, but such surveys typically include fewer items to measure SWB. Kööts-Ausmees et al. (2013) study a large sample of adults in 21 countries, using the European Social Survey. They report correlations between satisfaction and positive and negative affect equal to 0.49 and −0.51, respectively. Their satisfaction measure is based on two items, and the affect measures are based on four items each, referring to the frequency of emotions during the previous week. The British Office for National Statistics (2011) reports correlations between single-item measures of life satisfaction and positive and negative affect experienced yesterday (happy/anxious), equal to 0.55 and −0.26, respectively. Their estimates are based on a population sample of British adults. Wiest et al. (2011) report correlations between life satisfaction and affect measured with SWLS and PANAS (referring to the past months), respectively, based on data from a population sample of Germans aged 40–85. Their estimates for positive and negative affect are 0.30 and −0.29, respectively.

Several studies of affect within the economics literature are based on the Day Reconstruction Method (DRM), developed by Kahneman et al. (2004; see also Kahneman and Krueger, 2006). In DRM surveys, respondents are asked about time use and affect during the previous day, which is partitioned into distinct episodes defined by the main activity performed during that time interval (e.g. commuting to work, working, eating lunch). The respondents provide details about the context of each episode (e.g. where did the activity take place, and with whom), and they also rate episodes according to a set of emotions (e.g. to what extent they felt happy, sad or stressed). A central idea behind the DRM—besides linking time use and SWB variables—is to emulate experience sampling methods (ESM/EMA) for momentary affect measurement, but without imposing the costs and response burden associated with such methods.

A series of DRM studies have found rather weak correlations between life satisfaction and affect. Kahneman et al. (2004) study a sample of employed women in the US, and obtain a correlation of 0.38 between a single item life satisfaction measure (with four response categories) and net affect. Krueger and Schkade (2008) also study a sample of US women, and obtain a corresponding estimate of 0.31. They also find the correlation between satisfaction and the u-index, a measure of time spent in an unpleasant state, to be −0.26. Knabe et al. (2010) study a sample of unemployed Germans, and obtain a correlation of 0.32 between a single-item life satisfaction measure (on a numeric 0–10 scale) and net affect. Based on a pooled sample of French and US women, Kahneman et al. (2010) obtain an estimate of 0.36, between satisfaction measured with SWLS and an affect measure defined as the value of positive affect minus the maximum value of negative affect. The fact that the DRM studies yield weaker correlations is not surprising, given the short time frame of the affect measure, which refers to the previous day only.

In the study by Krueger and Schkade (2008), the DRM survey was administered twice to the same respondents, two weeks apart. The authors use test-retest correlations of life satisfaction and affect to adjust correlation estimates for measurement error. These estimates are substantially stronger, and equal to 0.50 for net affect and −0.48 for the u-index.

Psychologists typically handle measurement error by using survey instruments that consist of multiple items designed to measure the same latent construct. The inter-item correlation, which is a measure of reliability, can then be used to disattenuate the correlation between the variable measured with error and another variable (that could be handled likewise), using different type of latent variable models (or structural equation models). The disattenuated correlation can in this context be interpreted as pertaining to the latent variables of interest.

The study by Schimmack et al. (2002) accounts for measurement error using such methods. They estimate the correlation between SWLS and a hedonic balance measure (based on several items), in samples of US and German students. They obtain unadjusted correlations equal to 0.61 and 0.62, respectively, for the US and the German samples, and the corresponding measurement-adjusted correlations are equal to 0.68 and 0.76.

Luhmann et al. (2012) compare the importance of time frames of life satisfaction and affect questions, in a US population sample of persons older than 50 years. Life satisfaction and positive and negative affect were measured using two items each, for different time frames, and a latent variable model was used to account for measurement error. Their estimates of the correlation between satisfaction in general and positive and negative affect today, are 0.33 and −0.25, respectively. When satisfaction and affect were instead assessed using a common time frame referring to the past two months, the strength of these correlations is increased to 0.63 and −0.46, respectively.

Eid and Diener (2004) estimate a latent variable model that accounts for both measurement error and temporary deviations in life satisfaction and affect, using a student sample. Three measurements of life satisfaction (SWLS) and current mood (a multi-item measure) were made four weeks apart. They obtain a correlation of 0.74 between the stable components of general mood and stable life satisfaction, i.e. a doubly disattenuated correlation.

Of the studies reviewed here, Eid and Diener (2004) are the only ones to use a momentary measure of affect. Even so, this measure was elicited by means of a standard survey, and not ecologically, i.e. in the context of the respondent's everyday life. Neither do any of the above studies measure well-being at random times. As a consequence, it is not clear, even for the studies that use a shorter time frame, if the affect measure is even a representative snapshot of an individual's well-being. In fact, some study designs are explicitly non-random (e.g. Krueger and Schkade, 2008, in which respondents were surveyed on two Thursdays, two weeks apart).

A major contribution of the current study is to address these weaknesses, by using a genuine experience-sampling design in which affect is measured momentarily, ecologically, repeatedly, and during random points in time, over an extended period. Another contribution relative to previous studies is that our analysis explicitly accounts for the possibility of a current-mood bias of life satisfaction, since we avoid correlating satisfaction and affect measures from the same occasion.

3 Model

3.1 Model

In this section, we propose a simple measurement model, which formalizes the measurement issues outlined in the introduction. The equation

$$l_{it} = l_i^* + e_{it} \tag{1}$$

decomposes reported (observed) life satisfaction l_{it}, for individual i at time t, into an individual-stable latent component l_i^* and a temporary deviation e_{it}. Similarly,

$$a_{it} = a_i^* + u_{it}, \tag{2}$$

decomposes reported affect a_{it} into a stable component a_i^* and a deviation u_{it}. Conceptually, life satisfaction refers to a relatively stable condition, so e_{it} should be interpreted as a measurement error. The presence of such an error can be expected because of the arbitrariness of the response scale of self-reported satisfaction. While we would expect most people to be able to distinguish whether they are satisfied or dissatisfied with their lives, broadly speaking, and hence in what region of the response scale to report their answers, those who are satisfied may not, e.g., be able to clearly determine whether to report a 7 or an 8. For this reason, people with the same underlying satisfaction may report slightly different answers across different situations. In addition to this "pure" measurement error component, we also expect reported life satisfaction to correlate positively with current mood, in line with previous studies cited in the introduction.

The affect error term u_{it} also captures measurement error, related to the inconsistency of scale use, as well as well as variation in current mood. With respect to measuring momentary affect, we do not think of the latter in terms of an error. But if the purpose is to measure long run affect, a_i^*, as in our case, we can nevertheless treat it as a measurement error as well, statistically.

To clarify the meaning of "stable" or "long-run" well-being (which we use interchangeably) in this context, we do not think of l_i^* and a_i^* as fixed individual traits, or even necessarily constant within a shorter time frame, but rather as those components that are causally related to some set of determinants pertaining to the individual's stable life circumstances, e.g. occupation, marital status and health.[8] Hence, we expect l_i^* and a_i^* to change only when such circumstances change, as compared to e.g. having had a bad night's sleep. The maintained assumption throughout this paper—and admittedly an approximation—is that these circumstances remain unchanged over the period

[8]This begs the question about how to define the long-run determinants of SWB. We leave this as a bit of an open question, but think of it loosely as such policy relevant determinants that are typically studied in the SWB literature. Formally, we can think of l_i^* and a_i^* as being realizations of some happiness (regression) functions, conditional on these determinants at time t.

studied.[9]

With l_{it} and a_{it} being demeaned measures, the model is characterized by the following covariance structure:

$$E[l_{it}l_{js}] = \begin{cases} \sigma^2_{l^*} + \sigma^2_e, i = j, t = s \\ \sigma^2_{l^*}, \quad i = j, \forall t \neq s \\ 0, \quad\quad i \neq j, \forall t, s \end{cases} \tag{3}$$

$$E[a_{it}a_{js}] = \begin{cases} \sigma^2_{a^*} + \sigma^2_u, i = j, t = s \\ \sigma^2_{a^*}, \quad i = j, \forall t \neq s \\ 0, \quad\quad i \neq j, \forall t, s \end{cases} \tag{4}$$

$$E[l_{it}a_{js}] = \begin{cases} \sigma_{l^*,a^*} + \sigma_{e,u}, i = j, t = s \\ \sigma_{l^*,a^*}, \quad i = j, \forall t \neq s \\ 0, \quad\quad i \neq j, \forall t, s \end{cases} \tag{5}$$

The model has six parameters: $\sigma^2_{l^*} > 0$ and $\sigma^2_{a^*} > 0$ represent the variance of the stable components of life satisfaction and affect, and σ_{l^*,a^*} is the covariance between these variables. The variance of the measurement errors are denoted $\sigma^2_e > 0$ and and $\sigma^2_u > 0$, whereas $\sigma_{e,u}$ is their covariance within the same measurement occasion. Thus, $\sigma_{e,u}$ captures the current-mood bias effect on reported life satisfaction. Note that the assumptions in Equations (3)–(5) imply that the error variances σ^2_e and σ^2_u are constant over time (homoskedasticity), and that the errors have no time-series dependency (no autocorrelation). We discuss these assumptions further, and show that they are plausible, in Section 5.3. Since the between-individual covariance is always zero, we henceforth drop the individual subscripts for simplicity. The within-individual covariance structure is summarized in Table A.1 in Appendix A.

The object of primary interest is the population correlation between long-run life satisfaction and affect, denoted ρ, which is defined in terms of the model parameters as

$$\rho = \frac{\sigma_{l^*,a^*}}{\sigma_{l^*}\sigma_{a^*}}. \tag{6}$$

[9]Over a given period studied, there will always be some "churning", in the sense that some individuals' circumstances change, e.g. when losing one's job or divorcing. Insofar this leads to changes in long-run SWB, it may lead to inconsistent estimates in our context. We will indirectly assess to what extent this appears to be the case, in Section 5.3.

Note that ρ has a purely descriptive interpretation, rather than coming from a causal model. A positive value of ρ can be expected if life satisfaction has a positive causal effect on affect or vice versa, if both variables are manifestations of a single underlying SWB variable, or if there is overlap in the set of (other) determinants of the two variables.

The reliability ratios (or shares of between-person variation) of life satisfaction and affect, denoted δ_l and δ_a, are also of indirect interest for estimating ρ. These are defined as

$$\delta_l = \frac{\sigma_{l^*}^2}{\sigma_{l^*}^2 + \sigma_e^2} > 0 \qquad \text{and} \qquad \delta_a = \frac{\sigma_{a^*}^2}{\sigma_{a^*}^2 + \sigma_u^2} > 0. \tag{7}$$

3.2 Estimation

First, note that the simple correlation estimator

$$\hat{\rho} = \frac{\widehat{\text{cov}}(l_t, a_s)}{\sqrt{\widehat{\text{var}}(l_t)}\sqrt{\widehat{\text{var}}(a_s)}} = \frac{\widehat{\text{cov}}(l^* + e_t, a^* + u_s)}{\sqrt{\widehat{\text{var}}(l^* + e_t)}\sqrt{\widehat{\text{var}}(a^* + u_s)}}, \tag{8}$$

is inconsistent, with probability limit

$$\text{plim}(\hat{\rho}) = \begin{cases} \frac{\sigma_{l^*,a^*} + \sigma_{e,u}}{\sqrt{\sigma_{l^*}^2 + \sigma_e^2}\sqrt{\sigma_{a^*}^2 + \sigma_u^2}}, & t = s \\ \frac{\sigma_{l^*,a^*}}{\sqrt{\sigma_{l^*}^2 + \sigma_e^2}\sqrt{\sigma_{a^*}^2 + \sigma_u^2}} < \rho, & t \neq s \end{cases} \tag{9}$$

The asympotic bias of the within-period correlation could thus be either positive or negative, due to the current-mood bias in the numerator and the attenuation bias from the error variances in the denominator. The cross-period correlation, on the other hand, is unambiguously downward-biased (asymptotically).

Our proposed candidate for a consistent estimator is instead the reliability-adjusted cross-period correlation $\tilde{\rho}$, defined as

$$\tilde{\rho} = \frac{\hat{\rho}}{\sqrt{\hat{\delta}_l}\sqrt{\hat{\delta}_a}}, \qquad t \neq s \tag{10}$$

where $\hat{\delta}_l$ and $\hat{\delta}_a$ denote the test-retest correlations

$$\hat{\delta}_l = \frac{\widehat{\text{cov}}(l_t, l_s)}{\sqrt{\widehat{\text{var}}(l_t)}\sqrt{\widehat{\text{var}}(l_s)}} \qquad \text{and} \qquad \hat{\delta}_a = \frac{\widehat{\text{cov}}(a_t, a_s)}{\sqrt{\widehat{\text{var}}(a_t)}\sqrt{\widehat{\text{var}}(a_s)}}, \tag{11}$$

defined across the same non-zero time horizon, denoted $\Delta(t,s) = |t-s| > 0$, as $\hat{\rho}$. It follows from the model assumptions, together with the Weak Law of Large Numbers and the Continuous Mapping Theorem, that $\text{plim}(\hat{\delta}_l) = \delta_l$ and $\text{plim}(\hat{\delta}_a) = \delta_a$, and hence that $\text{plim}(\tilde{\rho}) = \rho$, i.e. $\tilde{\rho}$ is a consistent estimator of ρ.[10]

As for the inference, it is not obvious how to analytically derive the variance of $\tilde{\rho}$, given the possible dependence between the three random variables $\hat{\rho}$, $\hat{\delta}_l$ and $\hat{\delta}_a$, and the cluster structure of the data (described below). Our approach is to use a non-parametric clustered bootstrap procedure, in which clusters (respondents) are re-sampled with replacement, treating the within-cluster variation as fixed. For each of 1,000 sample draws, $\tilde{\rho}$ is computed from the triplet $\{\hat{\rho}, \hat{\delta}_l, \hat{\delta}_a\}$. Thereafter, a (possibly non-symmetric) 95% confidence interval around $\tilde{\rho}$ is computed using the distances between the median and the 2.5th and 97.5th quantiles of the bootstrap distribution.

4 Data

In order to estimate the model, we conducted a survey with the aim of measuring individual well-being 1) momentarily; 2) repeatedly; 3) at random times; and 4), over a reasonably long time horizon. To these ends, we designed a longitudinal mobile phone-based survey, in which the participants received notifications by SMS, containing a link to a short web questionnaire answered directly on the phone. It was thus required that participants owned a smartphone, i.e. a phone with a web browser and a mobile Internet connection. In order to, as far as possible, capture well-being in the moment, respondents were instructed to answer as soon as possible, and each query could only be answered up to one hour after the notification.

Unlike most previous experience sampling studies, we chose not to ask about

[10]We also tried to estimate the model with the generalized methods of moments (GMM) approach, but this turned out to be difficult due to missing values and problems of numerically inverting the variance-covariance matrix. GMM estimations on a balanced subset of the data yielded very similar results, however. We also estimated ρ more parsimoniously using only the (co-)variance terms corresponding to the off-diagonal cells of the sub-matrices of Table A.1. These estimates were very similar, so we opted to present results based on reliability adjusted correlations, since the reliability ratios are of interest in themselves.

Another, and perhaps more intuitive, candidate for an estimator, is $\overline{\rho} = \widehat{\text{cor}}\left(\overline{l}, \overline{a}\right)$, where $\overline{l}$ and $\overline{a}$ denote within-means of life satisfaction and affect, respectively. It is straight-forward to show that $\overline{\rho}$ is asymptotically (in the number of individuals) biased towards zero for finite number of within-observations, however. When we computed $\overline{\rho}$ and adjusted for this finite-sample bias, we got results that were similar to those based on $\tilde{\rho}$, presented in the paper.

any contextual information, e.g. what the respondent was doing at the time, where, or with whom. The idea was to minimize response burden, in order to encourage quick responses and maintain a high participation and response rate throughout the study period.

4.1 Participants

Participants were contacted by means of recruitment letters sent to a simple random sample of 3,000 persons in the Swedish population aged 18–50. Out of the 263 people who signed up on the study's website, we restrict our sample to $n = 252$ persons who answered any questions during at least two of three survey weeks (i.e. a net participation rate of 8.4%). Participants were rewarded with two cinema tickets, administered after the first survey week.

The composition of the participating sample was as follows: 64% were female, 68% were married or cohabiting, and 51% had children living in the household. As for their occupation, 68% were working, 17% were students, and the remaining 14% were either unemployed, sick or on parental leave. The age distribution was even, and similar to that in the target population.

4.2 Survey Structure

The survey was structured as follows. For each respondent, three days were drawn randomly from each of three active survey weeks, spread evenly across a seven-week period, which started on March 14 and ended on May 1, 2016. Henceforth, we will refer to the active survey weeks as survey week one, two and three (corresponding to chronological weeks one, four and seven). The choice of a seven-week period reflects that we wanted a sufficiently long period to obtain a representative picture of individual well-being, but not too long, so as to avoid picking up long-run changes in well-being (due to changed life circumstances).

For each of the nine active survey days, queries were sent out on random times in the morning (9 am – 1 pm), in the afternoon (1 pm – 5 pm) and in the evening (5 pm – 9 pm). Each respondent thus received a total of 27 queries, containing a total of 54 questions, i.e. on average two questions per query. An affect question (described below) was included in each query, whereas other well-being questions were added randomly within individuals and survey weeks, so that each respondent had his or her unique schedule, but with everyone receiving the same set of questions each week (and hence also across the whole survey).

Participants were also required to complete a questionnaire with questions about life satisfaction and demographic/socio-economic background variables, when they signed up on the study's website one to three weeks prior to the study. A separate questionnaire was also sent out on the last day of the study, with questions about life satisfaction and an evaluation of the survey.

4.3 Response Behaviour

Out of a total of $6,804$ queries sent (252 respondents $\times$ 27 queries), $N = 5,378$ were answered, yielding an overall response rate of 79%, or 21 answers per respondent on average. The distribution of individual response rates is shown in the left panel of Figure 1. The response rate was stable over the survey period: 79% during the first week, 80% during the second, and 78% during the third. The distribution of response times, i.e. the time passed between when a query was sent and when it was answered, is shown in the right panel of Figure 1. Responses were generally provided very quickly—the mean and median response times were 633 and 170 seconds, respectively, and three-quarters of all responses were provided within 15 minutes. Hence, we are confident in interpreting our affect measures as genuinely capturing well-being "here and now".

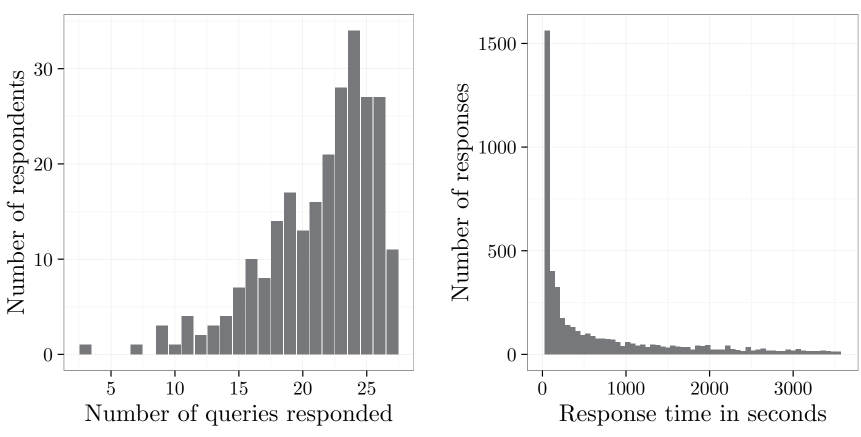

Figure 1: Distribution of response rates and response times

4.4 Well-Being Measures

A single item life satisfaction measure (henceforth SILS) was included once every survey week, based on the question "All things considered, how satisfied are you with your life as a whole nowadays?", with a response scale ranging from 0 to 10, and with endpoints labelled "extremely dissatisfied" and "extremely satisfied".[11] Life satisfaction was also measured in the sign-up and end questionnaires, using the Satisfaction With Life Scale (Diener et al., 1985), which includes five questions that are all answered on a scale from 0 to 6, with endpoints labelled "completely disagree" and "completely agree". We use the mean of the scores from these five questions (henceforth SWLS).[12] The distribution of SILS (mean = 6.8, sd = 1.9) and SWLS from sign-up (mean = 4.0, sd = 1.1) is shown in Figure 2.[13] We treat all SWB measures as cardinal.

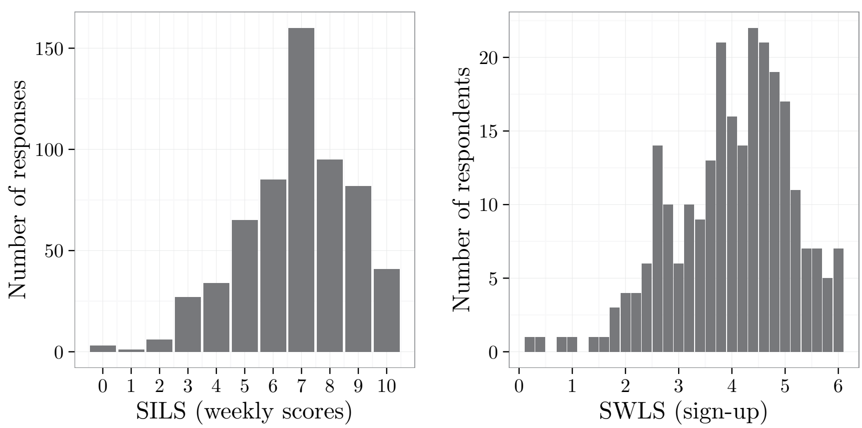

Figure 2: Distribution of life satisfaction responses

Our main measure of affective well-being is based on the question "How do you feel right now?", and was included in each of the 27 queries. The answers were elicited on a bipolar numeric response scale ranging from 0 to

[11]The survey was in Swedish, and the Swedish formulation of this question was adopted from the European Social Survey. The Swedish questionnaire and additional documentation of the survey can be accessed online from Martin Berlin's webpage, currently found at: `http://www.su.se/english/profiles/mabe7257`.

[12]The SWLS questions are "In most ways my life is close to my ideal", "The conditions of my life are excellent", "I am satisfied with my life", "So far I have gotten the important things I want in life" and "If I could live my life over, I would change almost nothing". The internal consistency of the SWLS scale in our sample, as measured by Cronbach's alpha, is equal to 0.86 (sign-up) and 0.88 (end).

[13]The correlation between SWLS at sign-up and SILS during survey week one, two and three are 0.69, 0.57, and 0.64.

10, with the endpoints labelled with a set of negative and positive adjectives: "extremely sad, displeased, depressed" and "extremely glad, pleased, happy", respectively.[14] We chose a bipolar measure in order to capture the full spectrum from negative to positive affect, as discussed in Section 2, without having to impose the extra response burden of including several specific negative and positive affect questions. Henceforth, we refer to this measure as momentary affect, or simply affect, when clear from the context. The mean and standard deviation of this variable, across all responses, are 6.5 and 1.9, respectively, and its distribution is shown in the left panel of Figure 3. For comparison, the distribution of within-individual affect means (mean = 6.5, sd = 1.2) is shown in the right panel of Figure 3.

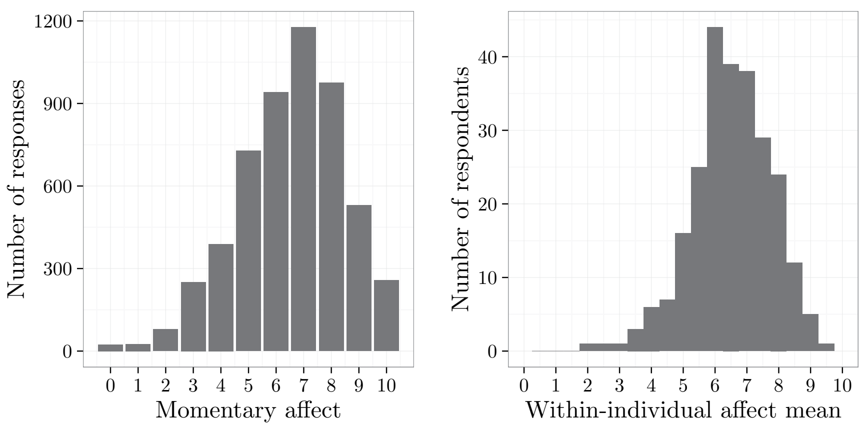

Figure 3: Distribution of momentary affect scores

A block of questions with specific emotions—happy, sad, stress, tired and pain—was included once a week, with questions phrased like "How happy do you feel right now?" (and similarly for other emotions). The response scale for these questions was unipolar, ranging from 0 to 6, and with endpoints labelled e.g. "I don't feel happy at all" and "I feel extremely happy".

Table 1 shows how the specific affect variables correlate with momentary affect, within the same measurement occasion (i.e. the same query). As can be seen from the strong correlations with happiness and sadness, momentary affect captures both positive and negative affect, as intended. The correlation with stress and tiredness is weaker, which is to be expected, since these emotions

[14]The adjectives were taken from the valence dimension of the Swedish Core Affect Scale (Västfjäll et al., 2002).

Table 1: Correlations between momentary affect and specific affect measures

	Happy	Sad	Stress	Tired	Pain
r	0.74	-0.64	-0.36	-0.35	-0.28
c.i.	(0.67, 0.80)	(-0.70, -0.57)	(-0.45, -0.27)	(-0.43, -0.25)	(-0.36, -0.20)

$N = 613$, $n = 246$. The 95% confidence intervals in parentheses are based on a non-parametric cluster-bootstrap procedure.

are more ambiguous with respect to how good or bad an experience is. The weak correlation with pain is a bit surprising, on the other hand, but it should also be noted that there is relatively little variation in this measure.

5 Results

5.1 Computation of the Correlation Estimator

Given the data structure, there are several possible ways of computing the reliability-adjusted correlation estimator $\tilde{\rho}$, proposed in Section 3.2. For SILS, the cross-period correlation $\hat{\rho}$ and the reliability ratios $\hat{\delta}_l$ and $\hat{\delta}_a$ can be computed either for the time horizon $\Delta(t, s) = 3$ weeks, or for $\Delta(t, s) = 6$ weeks, since there were two weeks of downtime between each of the three active survey weeks. We present results separately for both these cases. In each case, the triplet $\{\hat{\rho}, \hat{\delta}_l, \hat{\delta}_a\}$ is computed symmetrically, i.e. across the same time horizon, and using all available pooled observation pairs. For $\Delta(t, s) = 6$ weeks, $\hat{\rho}$, $\hat{\delta}_l$ and $\hat{\delta}_a$ are based on up to $2 \times 9 = 18$, 1 and $9 \times 9 = 81$ observation-pairs per individual, respectively (but fewer for most because of missing values). For $\Delta(t, s) = 3$ weeks, there are twice as many observations.

The reliability ratio of SWLS is computed across the sign-up and the end of the study, and hence across a time horizon of seven to nine weeks, depending on when the respondent signed up. We use SWLS–affect correlations computed across six to nine weeks, and we compute the reliability of affect across six weeks, to make the time horizons as comparable as possible. The (maximum) number of observations used per individual is the same as in the SILS case when $\Delta(t, s) = 6$ weeks.

As noted in Section 3.2, the basic motivation for computing $\hat{\rho}$ across different periods is to avoid a current-mood bias on life satisfaction judgments. Although such a bias is only present within the same measurement occasion in our model, it is possible that short-term fluctuations in affect are sluggish, and will spill over also to life satisfaction judgments that are close in time, e.g. during the same day. Rather than modelling such short-run dynamics, we choose to use only the cross-week covariance between life satisfaction and affect for computing $\hat{\rho}$. The crucial assumption for the consistency of our estimates is thus whether such dynamics persist across three or six weeks. This does in fact not appear to be a problem, as we show in Section 5.3 below.

5.2 Main Results

The results are presented in Table 2, which is structured as follows. The first column shows simple correlation estimates ($\hat{\rho}$), the second and third show reliability ratios for life satisfaction and affect ($\hat{\delta}_l$ and $\hat{\delta}_a$), and the fourth column shows the adjusted correlation estimates ($\tilde{\rho}$), obtained from the other estimates within the same row. Different time horizons and measures are organized by rows, within which samples are homogenized so as to include all individuals who contribute at least one observation pair to each estimate in that row.

Table 2: Correlations between life satisfaction and affect

	$\hat{\rho}$	$\hat{\delta}_l$	$\hat{\delta}_a$	$\tilde{\rho}$	n	N
SILS–affect	0.65				246	599
$t = s$	(0.60, 0.71)					
SILS–affect	0.42	0.72	0.30	0.91	190	26,406–
$\Delta(t, s) = 3\text{w}$	(0.34, 0.48)	(0.62, 0.81)	(0.23, 0.36)	(0.84, 0.97)		48,159
SILS–affect	0.43	0.72	0.33	0.88	163	13,203–
$\Delta(t, s) = 6\text{w}$	(0.36, 0.49)	(0.62, 0.79)	(0.26, 0.40)	(0.82, 0.95)		22,320
SWLS–affect	0.40	0.78	0.34	0.78	229	18,549–
$\Delta(t, s) = 6\text{–}9\text{w}$	(0.34, 0.46)	(0.71, 0.84)	(0.28, 0.38)	(0.71, 0.85)		29,853

Correlations on the same row are based on the same sample of individuals (n) who contribute at least one observation-pair to each correlation. The first row shows the momentary correlation between SILS and affect. The remaining rows show correlations computed across all combinations of cross-week observation-pairs (N), across three and six weeks (and up to ten weeks, for SWLS). $\hat{\rho}$ denotes the unadjusted cross-measure correlation, $\hat{\delta}_l$ and $\hat{\delta}_a$ are test-retest correlations (reliability ratios), and $\tilde{\rho}$ is the reliability-adjusted correlation. The 95% confidence intervals in parentheses are based on a non-parametric cluster-bootstrap procedure.

For comparison with subsequent estimates, the top row shows that $\hat{\rho} = 0.65$, when computed between SILS and affect within the same measurement occasion. Recall from Section 3.2 that this estimate is subject both to current-mood and attentuation bias, which go in different directions, and it is not clear *ex ante* which one of these dominates. Moving to the second row, we see that the corresponding correlation is equal to 0.42, when instead computed across a three-week period. The contrast between these two estimates thus shows that the current-mood bias is sizeable. The across-three-week estimate increases to 0.91, however, when adjusted with the reliability ratios of life satisfaction and affect. Besides the fact that this is a remarkably strong correlation in absolute terms, it also shows the huge overall impact of attenuation bias due to measurement error and temporal deviations in well-being, and that this source of bias dominates the current-mood bias, in line with our overall hypothesis. Unpacking the attenuation bias, we see that it is mainly due to the low reliability of affect, equal to 0.30, compared to the reliability of SILS, which is equal to 0.72. This is not surprising, given that momentary affect is expected to fluctuate more over a three-week period compared to life satisfaction, and our estimates in this respect are consistent with previous studies (Schimmack et al., 2002 and Eid and Diener, 2004).

Moving to the results for SILS across six weeks (third row), we see that they are very similar, with a slightly larger estimate of 0.33 for the reliability of affect, and hence a smaller adjusted correlation of 0.88. The quantitative similarity of these estimates is reassuring for our estimation strategy, as it suggests that the well-being dynamics are rather stable over the time period studied.

The results for SWLS (bottom row) differ somewhat from those for SILS. The unadjusted correlation between SWLS and affect (computed across six to nine weeks) is equal to 0.40, which is slightly weaker than the corresponding estimates for SILS. The reliability of SWLS, equal to 0.78, is higher, however, which is expected given that it is an average score of five sub-items.[15] Consequently, the adjusted correlation of 0.78 is slightly weaker compared to SILS, though still impressive in absolute terms, and stronger than correlations typically reported in previous studies.[16] We cannot say for sure why SWLS produces weaker correlations, but one candidate explanation is that the SWLS items are more retrospective and trait-like compared to the SILS question,

[15] For comparison, Diener et al. (1985) report a two-month test-retest correlation of 0.82 for SWLS, and our estimate lies in the upper range of comparable estimates reported in Pavot and Diener (1993).

[16] Interestingly, the estimate of 0.74 in Eid and Diener (2004) is close to, and within the confidence interval of, our estimate.

which is instead framed in terms of satisfaction with life "these days". Hence, an individual's SWLS score may to a larger extent be "locked in". It should also be noted that a difference between SILS and SWLS is that the latter was not measured on random occasions. We are not able to assess if this has an impact, however.

Turning to the precision of the point estimates, as indicated by the 95% confidence intervals, we see that it is unlikely that the true values are smaller than around 0.8 for SILS and 0.7 for SWLS. An informal inspection also rules out the possibility that the contrasts of main interest are produced by chance. For example, the confidence intervals of $\hat{\rho}$ and $\tilde{\rho}$ within the same row are non-overlapping throughout.

Summing up the results, our estimates that account for reliability issues are markedly stronger than those that do not, whether based on our own data, or those typically reported in previous literature. This is true both for the SILS and SWLS measures, although stronger correlations were found for SILS. Moreover, we also find a rather sizeable current-mood bias effect on SILS.

5.3 Testing for Heteroskedasticity and Autocorrelation

Our model formulation and estimation strategy rely on the assumption that the variance and covariance of the life satisfaction and affect error terms are stable. To take an example—which should be clear from Equations (10)–(11) in Section 3.2—a positive autocorrelation in affect would cause an upward (asymptotic) bias in the estimator of the reliability ratio $\hat{\delta}_a$, and hence a downward bias in the adjusted correlation $\tilde{\rho}$. Specifically, the model assumptions must hold across survey weeks, for our estimation strategy to be valid.

As noted in Section 5.2, the fact that the results for SILS are similar whether they are based on a time horizon of three or six weeks suggests that this is not a problem in practice. To assess this in more detail, we can examine the empirical variance–covariance matrix of SILS and affect, shown in Table A.2 in Appendix A. In this table, affect observations are indexed by individual measurement occasion $1, 2, \ldots, 27$, and life satisfaction by week $1, 2, 3$. Each cell on the lower and main diagonal represents the covariance of the row and and column variable, i.e. the mean of the (demeaned) within-individual cross-products, and the upper diagonal shows the corresponding correlation coefficients. The affect measurements which are simultaneous with life satisfaction are shown separately in rows/columns 28–30, to highlight the impact of current-mood bias on life satisfaction.

An informal inspection of Table A.2 suggests that there is a daily affect "shock", in the sense that affect correlates more strongly within a given day than across days. Moreover, there seems to be an autoregressive structure within days, so that affect in the morning is more correlated with affect in the afternoon, than with affect in the evening. There does not appear to be any similar dependence in the covariance across days and weeks, however, and neither does the variance appear to change systematically. For life satisfaction, the covariance between six weeks is slightly less than that between three weeks, whereas this is not true for the correlation coefficients.

Table 3: Test of homoskedasticity and autocorrelation of life satisfaction and affect

	$l_t l_s$		$a_t a_s$		$l_t a_s$	
	$t = s$	$\Delta(t,s) \geq 3$w	$t = s$	$\Delta(t,s) \geq 3$w	$t = s$	$\Delta(t,s) \geq 3$w
Intercept	3.66	2.63	3.65	1.11	2.44	1.56
	(0.39)	(0.39)	(0.22)	(0.12)	(0.30)	(0.19)
Week 2	0.17		−0.16		0.09	
	(0.35)		(0.23)		(0.35)	
Week 3	−0.19		0.09		−0.02	
	(0.40)		(0.25)		(0.37)	
$\Delta(t,s) = 6$w		−0.21		0.09		0.01
		(0.24)		(0.10)		(0.12)
Wald test, p	0.56	0.37	0.48	0.34	0.94	0.92
n	246	217	252	252	246	246
N	599	489	5,378	39,260	599	8,747

Each column shows results from regressions of week or cross-week indicators on within-individual cross-products of life satisfaction (SILS) and affect. The standard errors in parentheses are based on a robust variance-covariance matrix with individual-level clustering, and so is the Wald-test, which tests the null hypothesis that all coefficients except the intercept are zero (p-value shown).

To test the assumptions of homoskedasticity and autocorrelation more formally, we run regressions with subsets of the individual cross-products as outcomes, on indicator variables corresponding to the relevant regions in Table A.2.[17] To test for heteroskedasticity, the cross-products for $t = s$ are

[17] The individual data was used, rather than the aggregated data in Table A.2, in order to be able to compute a consistent, cluster-robust, variance-covariance matrix of the estimates. Strictly speaking, the results from these regressions do not capture the weekly (co-)variance,

regressed on dummies for week two and three (compared to week one). To test for autocorrelation, the cross-week products are regressed on a dummy for when the time horizon is six weeks (compared to three weeks). In the presence of heteroskedasticity and autocorrelation, we would expect these models to yield results that are significantly different from zero. For example, if there were a weekly affect shock that was autocorrelated across weeks, we would expect the covariance between affect during the first and second survey weeks to be stronger than that between the first and the third.

The results from these regressions are presented in Table 3. In neither case can we reject the null hypothesis of homoskedasticity and zero autocorrelation across weeks. Thus, we conclude that these issues should not be a problem for the consistency of our estimations. The stability of the cross-week covariances also suggests that the stable (latent) components of life satisfaction and affect are indeed stable during the period studied (see also footnote 9).

5.4 Robustness to Survey Context Effects

In our survey design, both the timing of the queries and the type of query (the set of questions included), are randomized within individuals. Hence, there should be little room for systematic survey context effects, as might be the case if, e.g., everyone answered the survey on the same particular day. As an additional robustness check, however, we also re-estimate the main results (for SILS) using life satisfaction and affect scores net of survey context effects. These scores are obtained as residuals from OLS estimations, in which SILS and momentary affect are regressed on a quadratic function of response time and a set of indicator variables capturing hour of the day, day of the week, survey week and type of query.

The results, shown in Table A.5 in Appendix A, are almost identical to the original results, with point estimates of the reliability-adjusted correlations computed across three and six weeks equal to 0.89 and 0.89, respectively. The similarity of the results is unsurprising in light of the low explanatory power of the survey context regressions, and the fact that the residuals correlated strongly with the original scores ($r = 0.99$ for both SILS and affect).

As an interesting sidenote, we only find statistically significant response effects for affect, with higher well-being on weekends and during evenings. It

as the data were demeaned across all weeks. The variation in means across weeks is negligible in practice, however.

is possible that our sample size is too small to detect such effects on life satisfaction, however, and the point estimates suggest that there may be a small positive weekend effect also in this case.

5.5 Results for Alternative Well-Being Measures

To assess the generality of the results presented in Section 5.2, and to address the concern that these are driven by our specific choice of affect measure, we also estimated the correlation between life satisfaction and a set of specific emotions measured once each survey week (as described in Section 4.4). The results from these estimations are presented in Appendix A, in Table A.3 for SILS, and in Table A.4 for SWLS. The latter table also presents separate estimates for the sub-item of the SWLS scale which is most comparable with SILS—the question "I am satisfied with my life" (henceforth SWLS4). For comparison with previous literature, tables A.3 and A.4 also include two affect measures derived from the specific affect questions: *net affect*, which is defined as the happy score minus the average score of sad, stress and pain; and the *u-index*, which is equal to 1 if the happy score is larger or equal to the maximum score of sad, stress and pain.[18] Since the specific affect questions were only included on three occasions for each individual, the sample sizes are somewhat smaller, with lower precision as a consequence. Hence we focus on the point estimates.

By and large, these results are in line with the main results. The happiness and sadness measures, which can be considered rather clear measures of positive and negative affect, correlate strongly and almost symmetrically with both SILS and SWLS, although somewhat stronger in the case of sadness for SWLS. In each case, the correlation is somewhat weaker in absolute terms, compared to the results based on the momentary affect measure used in the main analysis. This is expected, however, since momentary affect is bi-polar, i.e. it captures a spectrum of both positive and the negative affect. The correlation between net affect and life satisfaction ranges between 0.68, for SWLS, and 0.84, for SILS measured across three weeks. The strength of the correlation is somewhat weaker for the u-index (-0.60 for SWLS and -0.79 for SILS across three weeks), but this is perhaps not so surprising, given that it is a dummy variable, which discards variation in well-being above a certain threshold.

[18] The u-index was proposed by Krueger (2008) and in their context it was aggregated over time so that the resulting measure could vary between 0 and 1. Since our measure refers to a single point in time only, we should perhaps call it the "u-indicator" instead.

Turning to the more specific emotions tiredness, stress and pain, it can be seen that all of these correlate less strongly with both SILS and SWLS. Perhaps with the exception for pain, this is expected, in so far that these are less clear measures of positive and negative affect, as already noted in Section 4.4. Being tired is not a desirable state in general, but it is possible, e.g., to be both happy and tired after a workout or an eventful day. The fact that these correlations are weaker can be interpreted as life satisfaction exhibiting discriminant validity with respect to these measures.

Lastly, it can be noted that the adjusted correlation between SWLS4 and momentary affect is equal to 0.78, i.e. the same as between the composite SWLS measure and momentary affect. Compared to SWLS, SWLS4 appears to correlate more strongly with the other affect measures, but due to the low precision of these estimates, we cannot draw any clear conclusions.

6 Socio-Economic Correlates

It follows from our model (and random sampling within individuals), that the within-mean of momentary affect, $\overline{a}_i = \frac{1}{T_i} \sum_{T_i} a_{it}$, is a consistent estimator (in the number of within-observations T_i) of individual long-run affect a_i^*. Hence, we can use this as a summary measure of individual well-being. It is interesting to see how this measure compares with life satisfaction in terms of its socio-economic correlates. We do this by using the variables from the sign-up questionnaire (marital status, sex, presence of children in the household, employment and age) to run happiness regressions with aggregated affect, $\overline{a}_i$, and life satisfaction as outcomes, on a subset of 246 individuals for which all outcomes are available. To facilitate comparisons, all regressions are run on the z-scores of the outcome variables. We use SWLS from the sign-up questionnaire, whereas for SILS we use both the first non-missing weekly observation and a within-mean across all weeks, as separate outcomes.

The regression results are shown in Table 4, with coefficient estimates in the left panel (with standard errors in parentheses). Although the sample size is relatively small, several coefficient estimates are significantly different from zero. For example, the positive impact of marriage/cohabitation found in several other studies, is replicated across all outcomes. We also find a very strong negative impact of being long term sick or early retired, across all outcomes. Moreover, being older is associated with higher affect and life satisfaction in

Table 4: Socio-economic correlates of aggregated affect and life satisfaction

	Coefficient estimates				Ratio estimates			
	$\overline{\text{Affect}}$	SILS	$\overline{\text{SILS}}$	SWLS	$\overline{\text{Affect}}$	SILS	$\overline{\text{SILS}}$	SWLS
Married/cohab	0.32	0.52	0.46	0.53	1.00	1.00	1.00	1.00
	(0.15)	(0.14)	(0.14)	(0.15)	–	–	–	–
Female	−0.15	0.04	−0.05	0.21	−0.45	0.07	−0.12	0.40
	(0.13)	(0.12)	(0.12)	(0.12)	(0.49)	(0.27)	(0.31)	(0.31)
Children > 0	−0.07	−0.18	−0.16	−0.23	−0.22	−0.34	−0.35	−0.43
	(0.17)	(0.16)	(0.17)	(0.16)	(0.49)	(0.29)	(0.33)	(0.33)
Student	0.10	0.29	0.28	0.35	0.30	0.55	0.61	0.66
	(0.16)	(0.17)	(0.16)	(0.17)	(0.50)	(0.33)	(0.39)	(0.40)
Parental leave	0.66	0.40	0.71	0.62	2.05	0.77	1.55	1.17
	(0.26)	(0.27)	(0.25)	(0.16)	(1.67)	(0.46)	(0.89)	(0.44)
Unemployed	−0.31	−0.57	−0.40	−0.72	−0.95	−1.08	−0.87	−1.37
	(0.38)	(0.33)	(0.34)	(0.41)	(1.14)	(0.70)	(0.67)	(1.04)
Sick/retired	−1.07	−1.61	−1.44	−1.59	−3.32	−3.07	−3.13	−3.01
	(0.34)	(0.33)	(0.36)	(0.29)	(3.53)	(1.94)	(2.43)	(1.66)
Age 27–35	0.37	0.32	0.34	0.23	1.15	0.62	0.74	0.44
	(0.20)	(0.19)	(0.20)	(0.17)	(0.60)	(0.30)	(0.34)	(0.29)
Age 36–44	0.51	0.40	0.57	0.57	1.59	0.76	1.23	1.07
	(0.23)	(0.21)	(0.22)	(0.21)	(1.06)	(0.35)	(0.58)	(0.45)
Age 45–50	0.65	0.50	0.63	0.62	2.01	0.96	1.37	1.17
	(0.24)	(0.22)	(0.22)	(0.22)	(1.37)	(0.44)	(0.67)	(0.54)
R^2	0.15	0.21	0.20	0.24				

The standard errors in parentheses are based on a heteroskedasticity-robust covariance matrix, and are computed using the Delta Method for the ratio estimates. The ratios in the right panel are computed by dividing the coefficient estimates in the left panel by the "Married/cohab" coefficient. All regressions are based on the same sample of $n = 246$ individuals. The outcome variables are z-scores of, from left to right: within-mean of affect, SILS (first weekly non-missing observation), within-mean of SILS, and SWLS (from sign-up). The within-means of affect and SILS are based on an average of 21.6 and 2.4 within-observations, respectively. The reference category is male, neither married nor cohabiting, has no children in the household, is working/employed, and is between 18–26 years old. The intercept and a category for other employment that only applies to one person are omitted from the results.

our sample (in contrast with results from several other studies), and so is being on parental leave (though insignificant in one SILS model). Socio-economic variables explain more variation in SWLS than in affect, in terms of their respective R^2-values of 0.24 and 0.15, with SILS somewhere inbetween with $R^2 = 0.21$ (which is not improved by averaging the SILS scores across weeks).

An informal inspection of the magnitudes of the coefficients—which should be interpreted in terms of standard deviations of the outcome variable—seems to suggest that life satisfaction and affect are similar in terms of their correlates. It is more relevant to compare the ratios of coefficient-pairs from within the same regression, however. Such ratios are directly informative of the well-being tradeoffs between different factors, which is ultimately what matters for whether policy implications might differ as a result of using different SWB measures. Moreover, the magnitude of the ratios is not sensitive to differences in the amount of noise in the outcome variables, as is the case for z-scores.

The ratio estimates are presented in the right panel of Table 4, using the coefficient for being married or cohabiting as the denominator throughout. As an example, consider the ratio between the coefficients "Sick/early retired" and "Married/cohabiting". This estimate can be interpreted as the well-being impact of being sick or early retired (relative to being employed), measured in units equivalent to the well-being impact of being married or cohabiting (relative to not). This ratio is equal to -3.3, based on the affect estimates, -3.1 for SILS, -3.1 for averaged SILS, and -3.0, for SWLS. The estimated negative well-being impact of being long term sick or early retired is thus about three times stronger than the positive impact of being married or cohabiting, regardless of which outcome we consider. The question of whether the use of different SWB measures implies different trade-offs could be assessed in terms of whether the differences in ratios like this are statistically and economically significant across estimations.[19] Although the differences between some ratios are economically significant, we do not find any statistically significant differences. In fact, most of the ratios are themselves not significantly different from zero.[20] Our results are thus not inconsistent with life satisfaction and affect having the same determinants, but in fact we lack the statistical power

[19]Clark and Senik (2011) and Clark (2016) use the correlation between the vectors of coefficient estimates based on different SWB outcomes, to quantify the degree of similarity with respect to socio-economic correlates. This measure can obscure differences in the ratios, however, e.g. if the coefficients fall on a straight line with a non-zero intercept. Hence, we focus on the ratios instead. For comparison, we find that the coefficient vectors (excluding the intercept) from our estimations are strongly correlated—the correlation between the coefficient estimates based on mean affect, on the one hand, and those based on SILS, mean SILS and SWLS, on the other hand, are equal to 0.95, 0.98 and 0.94, respectively.

[20]The inference for ratios of regression coefficients is somewhat complicated. We rely on the asymptotic results of the Delta Method, which provides a formula for computing standard errors, based on the covariance matrix of the OLS coefficient estimates. Asymptotically, the ratios should be normally distributed, but this may not be a good approximation in our case, due to the small sample size.

to detect such discrepancies.

At this point, we also want to comment on a pair of studies by Kahneman and Deaton (2010) and Kahneman et al. (2010), concerned with discrepancies between the correlates of life satisfaction and affect. The affect measures used in both these studies refer to the previous day only, in contrast to the measure in our study, which is measured momentarily, across three different weeks during a seven-week period. Assuming that a measure of yesterday's affect captures long-run affect plus white noise (in accordance with our model), then one would expect to obtain weaker correlations between socioeconomic factors and affect, than for the correlations with life satisfaction. This is consistent with the results in Kahneman et al. (2010).

Kahneman and Deaton (2010) focus on the association between SWB and income, and find that its association with life satisfaction is stronger than the one with respect to either positive or negative affect, at least in relation to some covariates, such as being married. Moreover, they find a satiation point—beyond which there is no significant association with income—with respect to affect, but not with respect to life satisfaction. Due to lack of income data in our study (and a too small sample), we cannot replicate their analysis. However, we highlight as an important question for future research to investigate whether income has differential impacts on life satisfaction and affect, also when the affect measure is based on momentary measurement (and adjusted or aggregated so as to remove the impact of temporary fluctuations).

7 Conclusion

Summing up, we find a strong correlation between life satisfaction and long-run affect, when accounting for various reliability-related measurement issues. Potential violations of the model assumptions, in the form of heteroskedasticity and autocorrelation, do not appear be a problem for our estimation strategy, and neither do survey context effects. Moreover, our results apply to several different SWB measures, although the magnitudes of the different estimates vary somewhat. A strong correlation between life satisfaction and affect does not rule out the existence of differences in their determinants, however. Due to data limitations, we cannot provide a precise answer on this matter, but at least we do not find any clear discrepancies in the relative importance of different socio-economic correlates.

To be clear, we believe that the increased awareness of the conceptual distinction between life satisfaction and affective well-being is a good thing, and in general, we do not suggest that these should be bunched together under a generic happiness label. On the contrary, we believe that the use of a particular SWB measure should be conceptually and normatively well-motivated. Our results do suggest, however, that the empirical differences between evaluative and affective aspects of SWB might have been exaggerated in some previous research. Besides the evidence presented in this paper, there are also strong theoretical reasons to believe that life satisfaction and affect should correlate strongly in the long run—if people care about their day-to-day emotional well-being, we would expect that their long-run affect levels should be an important input for their life satisfaction judgments, and vice versa, one might expect people's life satisfaction to spill over to their day-to-day mood. Therefore, we suggest as an idea for future research to further explore the hypothesis of a one-dimensional long-run SWB dimension.

Our finding of a strong convergence between life satisfaction and affect has practical implications for the measurement of SWB. In particular, our results provide a rationale for the common practice in applied research to use life satisfaction as a summary measure of individual SWB, as it has higher (though still imperfect) reliability and is easier to measure than affect. This point also applies to policy initiatives with the aim of measuring national SWB levels. Whether the aim is to collect data on life satisfaction or affective well-being, however, we stress the value of survey designs that allow for repeated individual measurements. Already with two measurements, it is possible to estimate the reliability of the measures used, which in turn can be used to compute reliability-adjusted estimates.

Previous literature has discussed the importance of survey context effects on life satisfaction judgments (see Lucas, 2016). The present study contributes to this debate, by showing that there is a direct and non-negligible association between current mood and reported life satisfaction—i.e. a current-mood bias. We therefore caution researchers against correlating different SWB measures from the same survey occasion. Moreover, we conjecture that the current-mood bias may spill over to other subjective variables, such as subjective health. Everything else equal, this would lead to an upward bias in estimates of the impact on such variables on SWB, but imperfect reliability of other subjective variables might drive the bias in the opposite direction also in those cases.

Some limitations of our study, and suggestions for future research, should also be mentioned. First, the data available for this study do not allow us to do an in-depth assessment of whether there are systematic differences in the correlates of life satisfaction and long-run affective well-being. This topic should be a priority for future studies, since applied research on SWB, and policy implications from such research, is typically concerned with the relationship between SWB and various socio-economic variables. A crucial point is whether future studies can reveal systematic and substantial differences in the relative strength of the determinants of life satisfaction and long-run affect. If not, policy implications based on either life satisfaction or affective well-being would not differ, regardless of the strength of the direct association between the two variables.

Second, it would be interesting to study the relationship between life satisfaction and affect over a longer time period. Although the seven-week long survey period of the current study should be enough to account for moment-to-moment and day-to-day fluctuations in well-being, and provide a fairly representative sample of the experiences of most people, this may not be the case for everyone, e.g. due to the survey period coinciding with a particularly stressful, but temporary, period at work. It is a bit of an open question to what extent people's life satisfaction judgments are forward and backward looking, but at least for scales such as the SWLS, which include explicitly backward-looking questions, it would seem like a "fair" comparison with affective well-being would require the latter to be measured over a rather long period. To the extent that the results based on such long-run measurements would differ from those of the present study, we hypothesize that the satisfaction–affect association may be even stronger. It is important to note, however, that studies based on long-run data must use methods that do not conflate temporary fluctuations with long-run changes in SWB.

Finally, future studies should replicate our results in other samples, e.g. to assess whether the association between life satisfaction and affect differs across countries or across different groups within countries.

References

Adler, M. D. (2013). Happiness surveys and public policy: What's the use? *Duke Law Journal 62*(8), 1509–1602.

Busseri, M. and S. Sadava (2011). A review of the tripartite structure of subjective well-being: Implications for conceptualization, operationalization, analysis, and synthesis. *Personality and Social Psychology Review 15*(3), 290–314.

Campbell, A., P. E. Converse, and W. L. Rodgers (1976). *The quality of American life: Perceptions, evaluations, and satisfactions.* New York: Russell Sage Foundation.

Clark, A. (2016). SWB as a measure of individual well-being. In M. D. Adler and M. Fleurbaey (Eds.), *The Oxford handbook of well-being and public policy*, Chapter 18, pp. 518–548. Oxford: Oxford University Press.

Clark, A. and C. Senik (2011). Is happiness different from flourishing? Cross-country evidence from the ESS. *Revue d'Économie Politique 121*(1), 17–34.

Connolly, M. (2013). Some like it mild and not too wet: The influence of weather on subjective well-being. *Journal of Happiness Studies 14*(2), 457–473.

Deaton, A. (2012). The financial crisis and the well-being of Americans: 2011 OEP Hicks Lecture. *Oxford Economic Papers 64*(1), 1–26.

Derogatis, L. (1975). The affects balance scale. *Baltimore, MD: Clinical Psychometric Research*.

Di Tella, R. and R. MacCulloch (2006). Some uses of happiness data in economics. *The Journal of Economic Perspectives 20*(1), 25–46.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin 95*(3), 542–575.

Diener, E., R. A. Emmons, R. J. Larsen, and S. Griffin (1985). The satisfaction with life scale. *Journal of Personality Assessment 49*(1), 71–75.

Diener, E., E. Suh, R. Lucas, and H. Smith (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin 125*(2), 276–302.

Dolan, P. and D. Fujiwara (2016). Happiness-based policy analysis. In M. D. Adler and M. Fleurbaey (Eds.), *The Oxford handbook of well-being and public policy*, Chapter 10, pp. 286–318. Oxford: Oxford University Press.

Eid, M. and E. Diener (2004). Global judgments of subjective well-being: Situational variability and long-term stability. *Social Indicators Research 65*(3), 245–277.

Frey, B. S. and A. Stutzer (2002). What can economists learn from happiness research? *Journal of Economic Literature 40*(2), 402–435.

Haider, S. and G. Solon (2006). Life-cycle variation in the association between current and lifetime earnings. *The American Economic Review 96*(4), 1308–1320.

Haybron, D. M. (2016). Mental state approaches to well-being. In M. D. Adler and M. Fleurbaey (Eds.), *The Oxford handbook of well-being and public policy*, Chapter 12, pp. 347–378. Oxford: Oxford University Press.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener, and N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*, Chapter 1, pp. 3–25. New York: Russell Sage Foundation.

Kahneman, D. and A. Deaton (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences 107*(38), 16489–16493.

Kahneman, D. and A. Krueger (2006). Developments in the measurement of subjective well-being. *The Journal of Economic Perspectives 20*(1), 3–24.

Kahneman, D., A. Krueger, D. Schkade, N. Schwarz, and A. Stone (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science 306*(5702), 1776–1780.

Kahneman, D. and J. Riis (2005). Living, and thinking about it: Two perspectives on life. In N. Baylis, F. A. Huppert, and B. Keverne (Eds.), *The science of well-being*, pp. 285–301. Oxford: Oxford University Press.

Kahneman, D., D. Schkade, C. Fischler, A. Krueger, and A. Krilla (2010). The structure of well-being in two cities: Life satisfaction and experienced happiness in Columbus, Ohio; and Rennes, France. In E. Diener, D. Kahneman, and J. Helliwell (Eds.), *International differences in well-being*, pp. 16–33. New York: Oxford University Press.

Knabe, A., S. Rätzel, R. Schöb, and J. Weimann (2010). Dissatisfied with life but having a good day: Time-use and well-being of the unemployed. *The Economic Journal 120*(547), 867–889.

Kööts-Ausmees, L., A. Realo, and J. Allik (2013). The relationship between life satisfaction and emotional experience in 21 European countries. *Journal of Cross-Cultural Psychology 44*(2), 223–244.

Krueger, A. (2008). Are we having more fun yet? categorizing and evaluating changes in time allocation. *Brookings Papers on Economic Activity 2007*(2), 193–215.

Krueger, A. and D. Schkade (2008). The reliability of subjective well-being measures. *Journal of Public Economics 92*(8-9), 1833–1845.

Kuppens, P., A. Realo, and E. Diener (2008). The role of positive and negative emotions in life satisfaction judgment across nations. *Journal of Personality and Social Psychology 95*(1), 66–75.

Larson, R. and M. Csikszentmihalyi (1983). The experience sampling method. In H. T. Reis (Ed.), *New directions for methodology of social & behavioral science*, Volume 15, pp. 41–56. San Francisco: Jossey-Bass.

Lucas, R. (2016). Subjective well-being in psychology. In M. D. Adler and M. Fleurbaey (Eds.), *The Oxford handbook of well-being and public policy*, Chapter 14, pp. 403–423. Oxford: Oxford University Press.

Lucas, R. E., E. Diener, and E. Suh (1996). Discriminant validity of well-being measures. *Journal of Personality and Social Psychology 71*(3), 616.

Luhmann, M., L. C. Hawkley, M. Eid, and J. T. Cacioppo (2012). Time frames and the distinction between affective and cognitive well-being. *Journal of Research in Personality 46*(4), 431–441.

Office for National Statistics (2011). Initial investigation into subjective well-being from the opinions survey.

Pavot, W. and E. Diener (1993). Review of the satisfaction with life scale. *Psychological Assessment 5*(2), 164–172.

Robinson, M. D. and G. L. Clore (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin 128*(6), 934–960.

Schimmack, U., E. Diener, and S. Oishi (2002). Life-satisfaction is a momentary judgment and a stable personality characteristic: The use of chronically accessible and stable sources. *Journal of Personality 70*(3), 345–384.

Schimmack, U., P. Radhakrishnan, S. Oishi, V. Dzokoto, and S. Ahadi (2002). Culture, personality, and subjective well-being: Integrating process models of life satisfaction. *Journal of Personality and Social Psychology 82*(4), 582–593.

Schwarz, N. and F. Strack (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In D. Kahneman, E. Diener, and N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*, Chapter 4, pp. 61–84. New York: Russell Sage Foundation.

Schwarz, N., F. Strack, D. Kommer, and D. Wagner (1987). Soccer, rooms, and the quality of your life: Mood effects on judgments of satisfaction with life in general and with specific domains. *European Journal of Social Psychology 17*(1), 69–79.

Stone, A. and S. Shiffman (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine 16*(3), 199–202.

Suh, E., E. Diener, S. Oishi, and H. C. Triandis (1998). The shifting basis of life satisfaction judgments across cultures: Emotions versus norms. *Journal of Personality and Social Psychology 74*(2), 482–493.

Västfjäll, D., M. Friman, T. Gärling, and M. Kleiner (2002). The measurement of core affect: a Swedish self-report measure derived from the affect circumplex. *Scandinavian Journal of Psychology 43*(1), 19–31.

Watson, D., L. A. Clark, and A. Tellegen (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology 54*(6), 1063–1070.

White, M. and P. Dolan (2009). Accounting for the richness of daily activities. *Psychological Science 20*(8), 1000–1008.

Wiest, M., B. Schüz, N. Webster, and S. Wurm (2011). Subjective well-being and mortality revisited: Differential effects of cognitive and emotional facets of well-being on mortality. *Health Psychology 30*(6), 728–735.

A Tables

Table A.1: Within-individual covariance matrix of life satisfaction and affect

	l_t	l_s	a_t	a_s
l_t	$\sigma^2_{l^*} + \sigma^2_e$	$\sigma^2_{l^*}$	$\sigma_{l^*,a^*} + \sigma_{e,u}$	σ_{l^*,a^*}
l_s	$\sigma^2_{l^*}$	$\sigma^2_{l^*} + \sigma^2_e$	σ_{l^*,a^*}	$\sigma_{l^*,a^*} + \sigma_{e,u}$
a_t	$\sigma_{l^*,a^*} + \sigma_{e,u}$	σ_{l^*,a^*}	$\sigma^2_{a^*} + \sigma^2_u$	$\sigma^2_{a^*}$
a_s	σ_{l^*,a^*}	$\sigma_{l^*,a^*} + \sigma_{e,u}$	$\sigma^2_{a^*}$	$\sigma^2_{a^*} + \sigma^2_u$

Table A.2: Estimated matrix of cross-products of life satisfaction and affect

	Affect week 1									Affect week 2									Affect week 3									Affect together with SILS					
	Day 1			Day 2			Day 3			Day 4			Day 5			Day 6			Day 7			Day 8			Day 9			Affect			SILS		
	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	a_{20}	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}	a_{l_1}	a_{l_2}	a_{l_3}	l_1	l_2	l_3
a_1	3.52	*0.47*	*0.41*	*0.34*	*0.51*	*0.42*	*0.37*	*0.38*	*0.17*	*0.24*	*0.24*	*0.26*	*0.29*	*0.35*	*0.26*	*0.35*	*0.28*	*0.19*	*0.34*	*0.21*	*0.25*	*0.32*	*0.44*	*0.35*	*0.31*	*0.32*	*0.35*	*0.44*	*0.35*	*0.30*	*0.47*	*0.50*	*0.47*
a_2	1.49	3.25	*0.43*	*0.18*	*0.38*	*0.27*	*0.24*	*0.31*	*0.26*	*0.14*	*0.36*	*0.36*	*0.11*	*0.36*	*0.19*	*0.14*	*0.34*	*0.25*	*0.33*	*0.26*	*0.33*	*0.40*	*0.38*	*0.38*	*0.27*	*0.18*	*0.32*	*0.43*	*0.20*	*0.44*	*0.37*	*0.44*	*0.53*
a_3	1.41	1.35	3.21	*0.38*	*0.36*	*0.36*	*0.29*	*0.33*	*0.34*	*0.24*	*0.37*	*0.37*	*0.27*	*0.31*	*0.31*	*0.28*	*0.27*	*0.37*	*0.28*	*0.33*	*0.29*	*0.34*	*0.37*	*0.31*	*0.30*	*0.35*	*0.42*	*0.49*	*0.38*	*0.34*	*0.45*	*0.41*	*0.36*
a_4	1.19	0.58	1.32	3.66	*0.63*	*0.50*	*0.41*	*0.30*	*0.37*	*0.35*	*0.38*	*0.41*	*0.23*	*0.13*	*0.20*	*0.39*	*0.24*	*0.35*	*0.28*	*0.38*	*0.24*	*0.20*	*0.43*	*0.34*	*0.29*	*0.28*	*0.29*	*0.48*	*0.42*	*0.32*	*0.43*	*0.44*	*0.43*
a_5	1.87	1.28	1.26	2.27	3.82	*0.60*	*0.50*	*0.38*	*0.43*	*0.40*	*0.38*	*0.34*	*0.21*	*0.40*	*0.37*	*0.29*	*0.36*	*0.33*	*0.41*	*0.40*	*0.38*	*0.42*	*0.46*	*0.38*	*0.34*	*0.29*	*0.36*	*0.60*	*0.43*	*0.39*	*0.50*	*0.53*	*0.51*
a_6	1.60	0.92	1.28	2.00	2.34	4.21	*0.45*	*0.43*	*0.48*	*0.33*	*0.33*	*0.41*	*0.33*	*0.36*	*0.42*	*0.30*	*0.40*	*0.36*	*0.30*	*0.35*	*0.28*	*0.36*	*0.38*	*0.35*	*0.32*	*0.28*	*0.44*	*0.54*	*0.37*	*0.31*	*0.54*	*0.44*	*0.42*
a_7	1.24	0.80	1.01	1.56	1.86	1.70	3.63	*0.57*	*0.43*	*0.37*	*0.33*	*0.48*	*0.23*	*0.34*	*0.27*	*0.29*	*0.28*	*0.22*	*0.25*	*0.27*	*0.28*	*0.30*	*0.46*	*0.41*	*0.18*	*0.21*	*0.26*	*0.51*	*0.40*	*0.35*	*0.40*	*0.50*	*0.38*
a_8	1.34	1.09	1.20	1.09	1.49	1.74	2.07	3.95	*0.60*	*0.30*	*0.29*	*0.47*	*0.34*	*0.48*	*0.39*	*0.28*	*0.37*	*0.33*	*0.34*	*0.34*	*0.38*	*0.39*	*0.40*	*0.35*	*0.35*	*0.33*	*0.41*	*0.49*	*0.43*	*0.32*	*0.53*	*0.56*	*0.47*
a_9	0.59	0.87	1.22	1.39	1.63	1.89	1.65	2.29	3.69	*0.28*	*0.40*	*0.42*	*0.24*	*0.38*	*0.45*	*0.28*	*0.29*	*0.32*	*0.32*	*0.39*	*0.36*	*0.49*	*0.33*	*0.33*	*0.38*	*0.29*	*0.35*	*0.47*	*0.43*	*0.30*	*0.49*	*0.44*	*0.39*
a_{10}	0.89	0.48	0.81	1.23	1.41	1.22	1.29	1.07	1.06	3.76	*0.64*	*0.49*	*0.25*	*0.35*	*0.28*	*0.24*	*0.12*	*0.15*	*0.35*	*0.36*	*0.24*	*0.28*	*0.29*	*0.27*	*0.31*	*0.28*	*0.23*	*0.33*	*0.39*	*0.27*	*0.25*	*0.28*	*0.24*
a_{11}	0.75	1.18	1.15	1.35	1.35	1.23	1.19	1.00	1.43	2.20	3.18	*0.54*	*0.35*	*0.38*	*0.39*	*0.34*	*0.25*	*0.34*	*0.30*	*0.37*	*0.29*	*0.33*	*0.21*	*0.38*	*0.26*	*0.25*	*0.31*	*0.32*	*0.46*	*0.29*	*0.36*	*0.50*	*0.41*
a_{12}	0.87	1.11	1.16	1.33	1.13	1.41	1.54	1.57	1.44	1.51	1.64	2.97	*0.42*	*0.45*	*0.38*	*0.40*	*0.40*	*0.37*	*0.31*	*0.34*	*0.27*	*0.27*	*0.34*	*0.29*	*0.29*	*0.28*	*0.41*	*0.39*	*0.53*	*0.34*	*0.38*	*0.51*	*0.42*
a_{13}	0.95	0.36	0.89	0.78	0.71	1.23	0.79	1.24	0.88	0.98	1.19	1.41	3.59	*0.64*	*0.43*	*0.33*	*0.41*	*0.38*	*0.26*	*0.25*	*0.08*	*0.21*	*0.13*	*0.19*	*0.28*	*0.17*	*0.36*	*0.27*	*0.47*	*0.10*	*0.35*	*0.42*	*0.30*
a_{14}	1.23	1.18	1.04	0.43	1.48	1.37	1.17	1.71	1.29	1.23	1.27	1.37	2.13	3.54	*0.57*	*0.35*	*0.51*	*0.47*	*0.23*	*0.31*	*0.29*	*0.34*	*0.36*	*0.39*	*0.33*	*0.29*	*0.43*	*0.36*	*0.59*	*0.31*	*0.44*	*0.57*	*0.47*
a_{15}	0.86	0.64	0.98	0.72	1.40	1.67	0.99	1.49	1.64	0.97	1.33	1.22	1.46	1.88	3.48	*0.43*	*0.38*	*0.36*	*0.35*	*0.37*	*0.28*	*0.29*	*0.30*	*0.35*	*0.37*	*0.21*	*0.32*	*0.37*	*0.46*	*0.26*	*0.42*	*0.54*	*0.43*
a_{16}	1.38	0.50	1.00	1.43	1.11	1.15	0.99	1.06	1.06	0.95	1.22	1.33	1.24	1.25	1.61	3.92	*0.57*	*0.46*	*0.28*	*0.41*	*0.28*	*0.32*	*0.37*	*0.40*	*0.29*	*0.25*	*0.36*	*0.42*	*0.59*	*0.28*	*0.39*	*0.53*	*0.36*
a_{17}	0.90	1.11	0.88	0.90	1.33	1.68	1.00	1.41	1.12	0.45	0.84	1.35	1.44	1.88	1.41	2.22	3.75	*0.62*	*0.26*	*0.41*	*0.25*	*0.36*	*0.39*	*0.42*	*0.38*	*0.31*	*0.41*	*0.47*	*0.56*	*0.26*	*0.44*	*0.55*	*0.46*
a_{18}	0.62	0.78	1.15	1.26	1.15	1.35	0.79	1.17	1.13	0.51	1.13	1.08	1.27	1.63	1.18	1.59	2.16	3.36	*0.24*	*0.42*	*0.33*	*0.40*	*0.39*	*0.34*	*0.33*	*0.34*	*0.44*	*0.25*	*0.50*	*0.38*	*0.44*	*0.50*	*0.43*
a_{19}	1.10	1.10	0.93	0.92	1.43	1.11	0.90	1.26	1.11	1.19	0.96	0.93	0.93	0.76	1.16	1.03	0.90	0.80	3.30	*0.61*	*0.59*	*0.42*	*0.43*	*0.35*	*0.31*	*0.17*	*0.30*	*0.29*	*0.29*	*0.45*	*0.35*	*0.36*	*0.47*
a_{20}	0.67	0.80	1.01	1.23	1.35	1.32	0.89	1.26	1.40	1.24	1.16	1.02	0.85	1.04	1.29	1.53	1.45	1.44	2.18	3.65	*0.60*	*0.49*	*0.50*	*0.53*	*0.47*	*0.29*	*0.41*	*0.35*	*0.39*	*0.58*	*0.38*	*0.44*	*0.52*
a_{21}	0.89	1.19	0.97	0.92	1.49	1.20	1.10	1.49	1.27	0.92	1.04	0.95	0.30	1.16	1.09	1.18	1.00	1.26	2.28	2.32	4.30	*0.46*	*0.53*	*0.45*	*0.26*	*0.33*	*0.35*	*0.29*	*0.31*	*0.67*	*0.32*	*0.40*	*0.51*
a_{22}	1.08	1.33	1.12	0.69	1.43	1.24	0.94	1.28	1.68	1.01	1.08	0.83	0.72	1.11	0.97	1.10	1.21	1.26	1.47	1.71	1.72	3.31	*0.65*	*0.56*	*0.56*	*0.40*	*0.39*	*0.41*	*0.31*	*0.50*	*0.39*	*0.43*	*0.47*
a_{23}	1.62	1.35	1.28	1.50	1.70	1.55	1.60	1.55	1.22	1.12	0.73	1.13	0.45	1.29	1.09	1.46	1.42	1.33	1.61	1.88	2.21	2.28	3.96	*0.68*	*0.49*	*0.40*	*0.42*	*0.43*	*0.32*	*0.65*	*0.48*	*0.50*	*0.62*
a_{24}	1.31	1.38	1.13	1.27	1.43	1.47	1.49	1.40	1.27	1.12	1.39	0.96	0.71	1.46	1.33	1.64	1.69	1.24	1.32	2.06	1.96	2.16	2.69	4.24	*0.53*	*0.50*	*0.46*	*0.42*	*0.35*	*0.60*	*0.53*	*0.59*	*0.60*
a_{25}	1.17	0.98	1.14	1.07	1.35	1.39	0.66	1.36	1.51	1.30	0.99	1.05	1.09	1.28	1.37	1.19	1.51	1.23	1.24	1.82	1.15	2.10	1.99	2.23	4.21	*0.57*	*0.65*	*0.36*	*0.45*	*0.50*	*0.41*	*0.47*	*0.50*
a_{26}	1.18	0.65	1.13	0.99	1.11	1.14	0.82	1.31	1.18	1.08	0.84	0.92	0.63	1.03	0.80	0.99	1.28	1.21	0.63	1.12	1.34	1.45	1.57	1.92	2.37	3.97	*0.61*	*0.32*	*0.44*	*0.48*	*0.46*	*0.48*	*0.47*
a_{27}	1.12	0.98	1.26	0.89	1.17	1.53	0.82	1.30	1.11	0.77	0.92	1.17	1.18	1.44	1.01	1.21	1.37	1.33	0.93	1.22	1.27	1.21	1.40	1.65	2.35	2.07	2.86	*0.43*	*0.51*	*0.49*	*0.55*	*0.54*	*0.58*
a_{l_1}	1.65	1.47	1.68	1.84	2.28	2.10	1.93	1.98	1.77	1.30	1.12	1.27	0.97	1.31	1.38	1.68	1.83	0.90	1.05	1.32	1.25	1.40	1.74	1.72	1.49	1.22	1.41	3.91	*0.45*	*0.32*	*0.65*	*0.51*	*0.51*
a_{l_2}	1.16	0.67	1.27	1.60	1.72	1.43	1.44	1.54	1.69	1.38	1.51	1.72	1.61	2.04	1.63	2.31	2.10	1.72	1.05	1.34	1.25	1.05	1.26	1.39	1.80	1.64	1.71	1.68	3.65	*0.29*	*0.50*	*0.68*	*0.43*
a_{l_3}	1.04	1.63	1.13	1.18	1.47	1.28	1.34	1.28	1.16	1.03	1.03	1.11	0.37	1.12	0.95	1.06	0.99	1.28	1.65	2.15	2.87	1.90	2.59	2.29	2.06	1.73	1.69	1.23	1.12	3.90	*0.38*	*0.46*	*0.66*
l_1	1.68	1.21	1.53	1.53	1.84	2.09	1.43	2.02	1.76	0.96	1.20	1.23	1.21	1.58	1.49	1.49	1.70	1.62	1.23	1.38	1.33	1.31	1.93	2.07	1.64	1.71	1.79	2.45	1.77	1.38	3.68	*0.69*	*0.72*
l_2	1.74	1.53	1.47	1.64	2.18	1.74	1.93	2.18	1.78	1.04	1.71	1.73	1.46	2.09	1.97	2.16	2.07	1.80	1.30	1.57	1.69	1.42	2.01	2.37	1.90	1.80	1.86	1.99	2.54	1.79	2.61	3.85	*0.74*
l_3	1.56	1.83	1.17	1.40	1.81	1.62	1.40	1.76	1.45	0.84	1.39	1.30	1.04	1.57	1.48	1.31	1.68	1.37	1.66	1.86	2.05	1.65	2.28	2.20	1.98	1.60	1.88	1.84	1.54	2.43	2.43	2.69	3.49

The lower triangle and the main diagonal show covariances and the upper triangle shows correlations. Rows/columns 1–27 show affect observations indexed by within-individual measurement occasion, but excluding observations measured simultaneously with life satisfaction. These are instead shown separately in rows/columns 28–30, indexed by weekly life satisfaction, which is shown in rows/columns 31–33. The total sample consists of 92,829 unique within-individual cross-products of affect and life satisfaction, based on 5,378 measurements within 252 individuals. Estimates for each cell are computed using all available complete observation-pairs, within respective cell. The average number of observations per cell is 165.

Table A.3: Correlations between life satisfaction (SILS) and different affect measures, test-retest correlations, and reliability-adjusted correlations

	$\hat{\rho}$	$\hat{\delta}_l$	$\hat{\delta}_a$	$\tilde{\rho}$	n	N
SILS–happy	0.36	0.70	0.30	0.79	158	281–
$\Delta(t,s) = 3$w	(0.24, 0.46)	(0.61, 0.79)	(0.14, 0.42)	(0.57, 1.05)		558
SILS–happy	0.42	0.72	0.51	0.69	123	123–
$\Delta(t,s) = 6$w	(0.31, 0.56)	(0.61, 0.80)	(0.38, 0.64)	(0.50, 0.86)		246
SILS–sad	−0.46	0.70	0.44	−0.84	158	281–
$\Delta(t,s) = 3$w	(−0.56, −0.36)	(0.60, 0.77)	(0.34, 0.55)	(−0.95, −0.74)		558
SILS–sad	−0.47	0.72	0.46	−0.82	123	123–
$\Delta(t,s) = 6$w	(−0.57, −0.34)	(0.60, 0.80)	(0.31, 0.61)	(−1.00, −0.65)		246
SILS–stress	−0.21	0.70	0.42	−0.38	158	281–
$\Delta(t,s) = 3$w	(−0.31, −0.08)	(0.61, 0.78)	(0.31, 0.55)	(−0.54, −0.15)		558
SILS–stress	−0.11	0.72	0.44	−0.20	123	123–
$\Delta(t,s) = 6$w	(−0.28, 0.01)	(0.60, 0.82)	(0.32, 0.56)	(−0.48, 0.02)		246
SILS–tired	−0.25	0.70	0.35	−0.50	158	281–
$\Delta(t,s) = 3$w	(−0.35, −0.16)	(0.62, 0.78)	(0.23, 0.44)	(−0.68, −0.33)		558
SILS–tired	−0.28	0.72	0.32	−0.58	123	123–
$\Delta(t,s) = 6$w	(−0.38, −0.14)	(0.59, 0.79)	(0.14, 0.48)	(−0.82, −0.32)		246
SILS–pain	−0.28	0.70	0.44	−0.51	158	281–
$\Delta(t,s) = 3$w	(−0.40, −0.17)	(0.60, 0.79)	(0.28, 0.56)	(−0.74, −0.31)		558
SILS–pain	−0.23	0.72	0.46	−0.40	123	123–
$\Delta(t,s) = 6$w	(−0.35, −0.11)	(0.61, 0.81)	(0.27, 0.65)	(−0.73, −0.16)		246
SILS–net affect	0.46	0.70	0.43	0.84	158	281–
$\Delta(t,s) = 3$w	(0.35, 0.57)	(0.60, 0.79)	(0.29, 0.55)	(0.72, 0.94)		558
SILS–net affect	0.47	0.72	0.57	0.73	123	123–
$\Delta(t,s) = 6$w	(0.32, 0.56)	(0.60, 0.80)	(0.44, 0.67)	(0.53, 0.87)		246
SILS–u-index	−0.31	0.70	0.21	−0.79	158	281–
$\Delta(t,s) = 3$w	(−0.41, −0.21)	(0.60, 0.78)	(0.07, 0.30)	(−1.15, −0.58)		558
SILS–u-index	−0.32	0.72	0.36	−0.64	123	123–
$\Delta(t,s) = 6$w	(−0.41, −0.18)	(0.61, 0.81)	(0.16, 0.49)	(−0.96, −0.40)		246

Correlations on the same row are based on the same sample of individuals (n) who contribute at least one observation-pair to each correlation. Each row show correlations computed across all combinations of cross-week observation-pairs (N), across three and six weeks. $\hat{\rho}$ denotes the unadjusted cross-measure correlation, $\hat{\delta}_l$ and $\hat{\delta}_a$ are test-retest correlations (reliability ratios), and $\tilde{\rho}$ is the reliability-adjusted correlation. The 95% confidence intervals in parentheses are based on a non-parametric cluster-bootstrap procedure.

Table A.4: Correlations between life satisfaction (SWLS and sub-item "I am satisfied with my life") and different affect measures, test-retest correlations, and reliability-adjusted correlations

	$\hat{\rho}$	$\hat{\delta}_l$	$\hat{\delta}_a$	$\tilde{\rho}$	n	N
SWLS4–affect	0.39	0.74	0.34	0.78	229	18,549–
$\Delta(t,s)$ = 6–9w	(0.33, 0.46)	(0.69, 0.80)	(0.28, 0.39)	(0.69, 0.87)		29,853
SWLS–happy	0.38	0.80	0.50	0.60	158	158–
$\Delta(t,s)$ = 6–9w	(0.26, 0.48)	(0.71, 0.86)	(0.39, 0.60)	(0.42, 0.73)		316
SWLS4–happy	0.37	0.74	0.50	0.61	158	158–
$\Delta(t,s)$ = 6–9w	(0.26, 0.50)	(0.67, 0.81)	(0.38, 0.62)	(0.44, 0.80)		316
SWLS–sad	−0.43	0.80	0.46	−0.71	158	158–
$\Delta(t,s)$ = 6–9w	(−0.54, −0.32)	(0.71, 0.87)	(0.32, 0.59)	(−0.94, −0.53)		316
SWLS4–sad	−0.47	0.74	0.46	−0.80	158	158–
$\Delta(t,s)$ = 6–9w	(−0.58, −0.39)	(0.65, 0.81)	(0.32, 0.59)	(−0.97, −0.64)		316
SWLS–stress	−0.25	0.80	0.50	−0.39	158	158–
$\Delta(t,s)$ = 6–9w	(−0.38, −0.09)	(0.69, 0.88)	(0.36, 0.60)	(−0.62, −0.17)		316
SWLS4–stress	−0.27	0.74	0.50	−0.45	158	158–
$\Delta(t,s)$ = 6–9w	(−0.39, −0.16)	(0.64, 0.83)	(0.40, 0.61)	(−0.69, −0.27)		316
SWLS–tired	−0.18	0.80	0.29	−0.37	158	158–
$\Delta(t,s)$ = 6–9w	(−0.30, −0.04)	(0.66, 0.86)	(0.11, 0.43)	(−0.65, −0.07)		316
SWLS4–tired	−0.22	0.74	0.29	−0.47	158	158–
$\Delta(t,s)$ = 6–9w	(−0.32, −0.11)	(0.67, 0.83)	(0.15, 0.42)	(−0.68, −0.24)		316
SWLS–pain	−0.27	0.80	0.44	−0.45	158	158–
$\Delta(t,s)$ = 6–9w	(−0.37, −0.17)	(0.71, 0.86)	(0.28, 0.59)	(−0.60, −0.28)		316
SWLS4–pain	−0.32	0.74	0.44	−0.56	158	158–
$\Delta(t,s)$ = 6–9w	(−0.43, −0.22)	(0.65, 0.82)	(0.25, 0.58)	(−0.75, −0.38)		316
SWLS–net affect	0.47	0.80	0.60	0.68	158	158–
$\Delta(t,s)$ = 6–9w	(0.39, 0.57)	(0.69, 0.85)	(0.50, 0.68)	(0.56, 0.81)		316
SWLS4–net affect	0.49	0.74	0.60	0.74	158	158–
$\Delta(t,s)$ = 6–9w	(0.40, 0.58)	(0.65, 0.81)	(0.50, 0.68)	(0.63, 0.83)		316
SWLS–u-index	−0.31	0.80	0.34	−0.60	158	158–
$\Delta(t,s)$ = 6–9w	(−0.42, −0.23)	(0.68, 0.86)	(0.20, 0.49)	(−0.84, −0.40)		316
SWLS4–u-index	−0.33	0.74	0.34	−0.65	158	158–
$\Delta(t,s)$ = 6–9w	(−0.44, −0.22)	(0.65, 0.82)	(0.20, 0.49)	(−0.90, −0.41)		316

Correlations on the same row are based on the same sample of individuals (n) who contribute at least one observation-pair to each correlation. Each row show correlations computed across all combinations of cross-week observation-pairs (N). $\hat{\rho}$ denotes the unadjusted cross-measure correlation, $\hat{\delta}_l$ and $\hat{\delta}_a$ are test-retest correlations (reliability ratios), and $\tilde{\rho}$ is the reliability-adjusted correlation. The 95% confidence intervals in parentheses are based on a non-parametric cluster-bootstrap procedure.

Table A.5: Correlations between life satisfaction and affect, net of response effects

	$\hat{\rho}$	$\hat{\delta}_l$	$\hat{\delta}_a$	$\tilde{\rho}$	n	N
SILS–affect $t = s$	0.66 (0.59, 0.71)				246	599
SILS–affect $\Delta(t, s) = 3\text{w}$	0.41 (0.34, 0.48)	0.71 (0.62, 0.78)	0.31 (0.25, 0.37)	0.89 (0.82, 0.95)	190	26,406– 48,159
SILS–affect $\Delta(t, s) = 6\text{w}$	0.43 (0.35, 0.50)	0.72 (0.63, 0.79)	0.32 (0.26, 0.39)	0.89 (0.80, 0.96)	163	13,203– 22,320

These results replicate those for SILS in Table 2, except that they are based on life satisfaction and affect scores net of measurement context effects. These are obtained as residuals from OLS estimations in which life satisfaction and affect are regressed on a quadratic function of response time and a set of indicator variables capturing hour of the day, day of the week, survey week and type of query (which questions were included).

Study 4

Do OLS and Ordinal Happiness Regressions Yield Different Results? A Quantitative Assessment

I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be.

—Lord Kelvin, *Electrical units of measurement*

1 Introduction

Subjective well-being (SWB) measures of general life satisfaction and affective well-being are typically elicited through self-assessment questions in surveys. A characteristic of such assessments—shared with plenty of other phenomena such as subjective health, political attitudes and measures of personality[1]—is their lack of a natural measurement unit. This is reflected in the diversity of survey items used, with variations in the question wording, the number of

[1] Including all variables obtained from Likert scales, which ask respondents to express their degree of agreement with some statement.

response categories and their labelling. For example, respondents in the US General Social Survey (GSS) are asked to assess their life satisfaction using the three categories "Not very happy", "Pretty happy" and "Very happy", whereas respondents in the European Social Survey (ESS) are presented with a numeric 0–10 scale, with the two endpoints labelled "Extremely dissatisfied" and "Extremely satisfied".[2] Formally, such data are usually regarded as ordinal, in the sense that the difference between "Not very happy" and "Pretty happy", in the GSS case, not necessarily represents the same difference in underlying life satisfaction as the difference between "Pretty happy" and "Very happy". The same can be said about numerically labelled scales, as in the ESS case.

Yet, the conventional wisdom in the SWB literature has it that it matters little whether SWB data is treated as cardinal or ordinal.[3] For instance, a recent OECD publication with guidelines for measuring SWB, remarks:

> Although most subjective measures of well-being are assumed to be ordinal, rather than cardinal, evidence suggests that treating them as if they were cardinal in subsequent correlation-based analysis does not lead to significant biases: the practice is indeed common in the analysis of subjective well-being data, and there appear to be few differences between the conclusions of research based on parametric and nonparametric analyses[.] (OECD 2013, p. 189)

This finding has surfaced in several studies, mostly as a side result of sensitivity analyses comparing estimates from OLS on a numerically coded well-being variable, with those from an ordered regression model. But the conclusions regarding similarity are often based on assessments of the sign and statistical significance of the estimated regression coefficients. Formal quantitative comparisons of the coefficient magnitudes are rare. Indeed, a common remark is

[2] The questions are shown in Section 3.1.

[3] Psychologists working on SWB (and related fields such as personality psychology) often analyse sets (called instruments or scales) of questions (items) meant to capture the same phenomenon. In doing so, data are typically treated as cardinal, when averaging sub-items from the same scale or through the use of factor analysis. Rather than putting too much weight on any particular measure, a multitude measures are often considered within a given study, the idea being that these are only imperfect manifestations of underlying latent constructs. This "pragmatist" approach to measurement can appear strange to economists, for both historical and methodological reasons. Empirical economists have typically worked with variables such as incomes, prices, days of unemployment and years of education, which are all measured on a ratio scale and for which it is more obvious that cardinal methods apply. Moreover, the workhorse method of econometrics—regression analysis—is suitable for studying one outcome at a time. Economists have thus been inclined to analyse single-item measures of SWB by means of ordered regression models.

that the results are qualitatively similar (e.g. Frey and Stutzer, 2000, p. 924).

The most cited study that examines this issue in the context of SWB is Ferrer-i-Carbonell and Frijters (2004). They compare OLS and ordered logit and probit estimates from a typical within-country happiness regression, using data from the German Socio-Economic Panel, in which life satisfaction is elicited on a 0–10 numeric scale. In their interpretation of the results relating to the difference between OLS and the ordered regression estimates, they conclude that "the sign of the coefficients are the same; whether a coefficient is significant is the same; and the trade-offs between variables are roughly the same" (p. 650). They do not quantify the differences however, neither in terms of standardized coefficient estimates or the coefficient ratios (capturing the trade-offs). Using their estimates to compute the difference in the coefficient ratio between subjective health and log-income, I find that the difference is in fact quite substantial, with the ordered probit ratio deviating 30% relative to the OLS ratio. The corresponding deviation for the ratio of the coefficients for having a steady partner and log-income is −63%.[4] This casts some doubt about whether OLS and ordered regressions yield *quantitatively* similar estimates.[5]

To the best of my knowledge, the only ones to quantify differences in empirical results between OLS and ordered regressions, in the context of SWB, are Clark (2016) and Stevenson and Wolfers (2008).[6] In both cases, OLS and ordered regression models are compared in terms of the correlation between the vector of coefficient estimates from respective model. In both studies, these correlations are found to be close to one, thus supporting the conventional wisdom. The correlation of the coefficient vectors may not be an appropriate measure of model similarity, however, as explained in Section 2.5.

A more formal treatment, not explicitly linked to SWB data, is provided by Riedl and Geishecker (2014), who use simulations to study the consistency of OLS and different ordered response models that incorporate fixed effects. In their simulation examples, they find that OLS (with fixed effects) performs well

[4] My computations are based on the cross-sectional estimates in column 1 in Table 1 and columns 1 and 3 of Table 2 in Ferrer-i-Carbonell and Frijters (2004).

[5] But to their credit, Ferrer-i-Carbonell and Frijters (2004) are correct in that the inclusion of individual fixed effects or not matters more than the choice between OLS versus ordered regression.

[6] These results are not among the main results in either of these studies. In Clark (2016), the main focus is on comparing different SWB measures. In Stevenson and Wolfers (2008), the correlation estimates (together with a graphical representation) are presented in a method/robustness appendix—in a comparison of a linear probability model prediciting the outcome "Very happy", and different ordered models.

with respect to estimating coefficient ratios. With the exception of their study, the question of cardinality has not been given a systematic formal treatment, across different response scales, data sets and alternative ordinal response models. As such, it is not yet clear whether it is "kosher" for the applied happiness researcher to code reported SWB variables numerically and then use OLS or other methods relying on cardinality. The aim of this paper is to shed light on this issue.

I assess whether empirical estimates from OLS and different ordered regression models differ significantly, in both statistical and economic terms, i.e. in terms of magnitudes. To the extent that these estimates are similar, it does not prove that they are similar to the true parameters of interest, however. To address this second and, arguably, more difficult question, I simulate well-being data with known values of the true parameters. I then compare how OLS and ordered regressions perform with respect to these true parameters.

I structure my analysis around two empirical examples. The first example concerns the impact of unemployment relative to income on life satisfaction, as measured by the ratio of the coefficient estimates of unemployment and income. I estimate this ratio using three different data sets with three different life satisfaction scales: the GSS and the Panel Study of Income Dynamics (PSID) for the US, and the ESS for Sweden. This example speaks to the literature about the within-country determinants of SWB (see Dolan et al., 2008, for a review).

The second example concerns cross-country differences in mean life satisfaction. I estimate this in terms of mean-shift coefficients, standardized with the international standard deviation of life satisfaction, using data from ESS covering 21 countries. This example speaks to the literature about country-level differences in SWB, e.g. the Easterlin Paradox (Easterlin, 1974, Sacks et al., 2012), as well as policy inititatives concerned with tracking national well-being levels.

My results are somewhat mixed. I do find statistically significant differences between OLS, probit and logit estimates in the empirical examples considered, but these are typically not large in magnitude, with deviations in the order of 1–10%. The differences are found to be largest for the 11-point response scale used in the ESS, compared to the scales with fewer categories in the GSS and in the PSID. When the distribution of the error term of the ordinal model is assumed to be skewed, larger discrepancies are found throughout. The simulation results are, broadly speaking, in line with the empirical results. In short, the results do not overturn the conventional wisdom, but they paint a more nuanced picture.

The remainder of this paper is structured as follows. In Section 2, I discuss levels of measurement and parameters of interest in the context of SWB, whereafter I review the ordered regression model and propose measures for comparing results from different models. I analyse the relative impact of unemployment on life satisfaction in Section 3, and cross-country differences in life satisfaction in Section 4. In Section 5, I discuss the results further and conclude.

2 Theoretical Framework

The main purpose of "happiness economics" is, arguably, to generate knowledge about how the population distribution of SWB is affected by different allocations of scarce resources. Such knowledge is particularly useful in non-market contexts, e.g. in governments' decisions about whether to spend money on policies aimed at either reducing unemployment or improving health, or perhaps instead lower taxes. Ideally, knowledge of the well-being impact of unemployment *relative* to other factors, such as health and income, can be weighed against the respective costs associated with affecting these outcomes, in order to allocate resources in such a way that overall well-being is maximized, or at least increased. Other practical and methodological problems aside,[7] this type of analysis requires a cardinal measure of well-being. In this section, I elaborate on this point.

2.1 Measurement Level of SWB

Consider the following four statements about SWB:

1. The typical person in the US is "Pretty happy" (nominal)

2. The median Dane has higher life satisfaction than the median Belgian (ordinal)

[7] For instance, the difficulty of estimating causal well-being effects, individual heterogeneity in these effects, and how to trade off well-being between individuals, when gains and losses differ.

3. The life satisfaction difference between employed and unemployed persons is 11% larger than the life-satisfaction gap between persons with good and bad health (interval)

4. Mean life satisfaction is 14% higher in Denmark than in Belgium (ratio)

For both research and policy purposes, we want to know if statements like these are meaningful. Whether this is the case or not depends on the level of measurement of the SWB variable in question. The nomenclature established by Stevens (1946) distinguishes between four different levels of measurement of a scale. From the lowest to the highest, these are: nominal, ordinal, interval and ratio, where the latter two are typically referred to as cardinal. A variable measured on a higher measurement level contains more information, and one can thus (meaningfully) make more far-reaching statements based on such variables. Each level of measurement is associated with a set of *admissible statistics*, which are permissible to use at higher levels of measurements, but not at lower levels. The four statements above follow this hierarchy, and the labels in the parentheses describe the minimum level of measurement required for the statistic in question to be admissible.[8] For example, it is meaningful to compute a mean difference when the SWB variable is measured on an interval scale, but not on an ordinal scale. Mathematically, each scale level can also be defined in terms of which transformations preserve the information content of the variable. Table 1 summarizes these categorizations.

In the context of SWB, the central discussion concerns whether we can assume an ordinal or an interval scale. The crucial characteristic of an interval scale, compared to an ordinal scale, is that of *equidistance*, i.e. that the difference between scale points represents the same magnitude, across the whole scale.[9] With an interval scale, we can compare differences in group means, compute correlations and partial correlations (using OLS), and use the standard deviation as a measure of spread.

There are three common approaches for analysing SWB data. The first approach, which I refer to as the ordinal approach, does not assume an un-

[8]The first statement implicitly refers to the mode, whereas the third refers to a comparison of relative differences in group means.

[9]It is quite common to see cardinal treatment being justified when the observed SWB variable is continuous. This is misleading, however. Observed SWB can be continuous in the sense of having a smooth distribution with responses spread across several categories, but without being a positive affine transformation of underlying SWB, as would be the case e.g. for a concave transformation. The appropriate criterion is thus equidistance of the scale points.

Table 1: Levels of measurement

Scale	Transformation	Example statistic
Nominal	—	Mode
Ordinal	$x' = f(x)$ $f()$ positive monotone	Median
Interval	$x' = a + bx$,	Mean Standard deviation
Ratio	$x' = bx$	Gini coefficient

Based on Stevens (1946).

derlying, latent, well-being variable. Instead, it takes the response categories of the reported well-being variable, or a specific threshold among these categories, to be the objects of interest, e.g. the share of people reporting being "Very happy". It is then possible to compare groups in terms of this outcome, or to use e.g. a linear probability model to estimate partial correlations with several variables.

The advantage with this approach is that it does not require any assumptions about how large a difference in happiness is represented by going from one category to the next. However, this also happens to be the main disadvantage. Without a quantitative yardstick, we cannot know whether a difference in the share of "Very happy" between two groups represents a happiness difference that is of any practical importance. Specifically, the type of cost benefit-analysis suggested in the beginning of this section, e.g. where the benefits of unemployment reduction and health improvement are related to their respective costs, is not possible. Even if, say, the impact of bad health, reduces the probability of being "Very happy" with 20%, compared to a 10% reduction for unemployment, we cannot infer that the well-being effect of bad health is twice as strong.

The ordinal approach can still be used for establishing the existence of a positive or negative association between SWB and some other variable (e.g. that unemployment has a negative impact on life satisfaction), and for ranking well-being distributions in case of dominance.

The second approach maintains the assumption that observed (reported) well-being is ordinal, but assumes further that observed well-being is a manifestation of an underlying, or latent, cardinal variable. This is the basis of the ordered response (regression) models. The models used in most cases, at least by economists, are the ordered probit and the ordered logit. Additional assumptions of these models are discussed below, but for now, the important assumption that they have in common is that they assume the existence of an interval-scale latent variable. I refer to this as the ordered (or ordinal) response approach, or more specifically as the ordered regression model (or method).

The third approach does not make a distinction between reported and underlying well-being. Instead, it assumes that reported well-being is an interval-scale variable with some specific assignment of numeric scores to the response categories. Typically, this assignment is equidistant, so that e.g. the three categories of the GSS are assumed to correspond to well-being values of 1,2 and 3. In some cases, as for the ESS, the answers are reported on an equidistant numeric scale to begin with, which it is then natural to use.[10] I refer to this as the cardinal approach. I also refer to it more specifically as the OLS method. This is somewhat sloppy however, given that the approach first and foremost builds on the assumption of a specific cardinalization of the reported well-being variable.

In this paper, I disregard the purely ordinal approach completely and focus on comparing results obtained from the ordered response approach and the cardinal approach. In light of recent skepticism towards the use of SWB variables, it is important to stress that these latter two approaches both build on the fundamental assumption of cardinality of SWB, whether at the latent or the observed level. Bond and Lang (2014) and Schröder and Yitzhaki (2017) both make the point, albeit a bit differently, that if the assumption of ordinality of the well-being variable is instead maintained throughout, then it is possible (except under specific conditions) to obtain reversals of group mean-comparisons (or the sign of regression coefficient estimates) based on such a variable, by means of applying a sufficiently convex or concave positive monotone transformation.

This criticism does not apply to the ordered response approach and the cardinal approach, as outlined above, however. The reason is simple—arbitrary positive monotone transformations are not permissible for interval scale vari-

[10] Conceptually, it is possible to maintain the notion of a latent well-being variable with this approach, but with the equidistant-coded observed variable being regarded as a sufficiently good approximation to latent variable.

ables, only affine transformations of the form $x' = a + bx$ are.[11]

Note that there are several other important issues concerning the measurement of SWB, e.g. whether scales are comparable across individuals, groups, countries, and time periods; the influence of response styles; survey context etc. I will disregard these issues and focus on contrasting the ordered response and the cardinal approach. In particular, I assume throughout this paper that SWB scales are interpersonally comparable.

2.2 Latent SWB and Parameters of Interest

Having assumed the existence of a SWB variable measured on an interval scale, let us denote it y^*. Assume further that y^* is a linear function of K variables $x_1, \ldots, x_K$, with associated coefficients $\beta_1, \ldots, \beta_K$ and an error term ϵ. In vector form, the model is written

$$y^* = \boldsymbol{x}'\boldsymbol{\beta} + \epsilon. \tag{1}$$

The coefficient vector $\boldsymbol{\beta}$ represents marginal well-being effects of $\boldsymbol{x}$, for continuous x, and discrete well-being effects for discrete x. Importantly, the fact that y^* is measured on an interval scale implies that $\boldsymbol{\beta}$ is only defined up to a constant, i.e. the coefficients are measured on a ratio scale. Hence, a single coefficient such as β_1 only contains information about the direction of the effect of the variable x_1, but not how large the effect is. A natural way to quantify the effect is to do so in terms of the effect of another variable x_2, by computing the ratio β_1/β_2, given that $\beta_2 \neq 0$. This ratio can be interpreted as the well-being effect of x_1, measured in units equivalent to the well-being effect of x_2. When x_1 and x_2 are continuous, $-\beta_1/\beta_2$ corresponds to the marginal rate of substitution between x_1 and x_2.

β_1/β_2 thus reflects the trade-off between x_1 and x_2, and can be combined with information about the costs (or prices) associated with x_1 and x_2, for cost-benefit analysis, as in the example with health and unemployment mentioned above.[12] A ratio of particular interest is the well-being "money-metric" which

[11] Another point made by Bond and Lang (2014) is that mean-comparisons based on ordered regressions can be sensitive to whether the variance of the latent variable is estimated as a free parameter, as compared to being the same for both groups, which is the standard assumption. This issue is relevant also for my framework and was investigated in a previous version of this paper. I dropped this aspect of the analysis to keep the current paper more focussed, however, but I plan to address it in future work.

[12] Loosely speaking, such optimization problems will have nontrivial solutions if well-being is concave in the determinants, and/or if the associated costs are convex.

is obtained when x_2 is (log) income. In the first example of this paper, I focus on such a ratio, with the numerator being the coefficient associated with an indicator variable for being unemployed.

An alternative approach is to divide β_1 with the standard deviation of y^*, thus measuring the effect of x_1 in terms of standard deviations of well-being, for a given population. Such estimates have a meaningful interpretation, but are by themselves not of direct use in a cost-benefit framework. I will consider this type of estimates in the second example of this paper, concerned with cross-country differences in mean life satisfaction. It is important to keep in mind that estimates that are standardized in this way might not be comparable with estimates from other samples with different spread in the SWB variable.

2.3 Ordinal Representation

The motivation for this paper is that, although y^* is assumed to exist, it may not be observed directly. Instead, I make the common assumption that self-reported SWB, denoted y, is only an ordinal representation of y^*. Specifically, I assume that the relationship between y and y^* follows the standard ordered regression model, which combines the linear model for y^* in Equation (1), with the following relationship for mapping y^* to y:

$$
\begin{aligned}
y &= 1 \text{ if } -\infty < y^* \leq \alpha_1 \\
y &= 2 \text{ if } \alpha_1 < y^* \leq \alpha_2 \\
&\dots \\
y &= J \text{ if } \alpha_{J-1} < y^* < \infty,
\end{aligned}
$$

where J is the number of ordered response categories of y. The model is fully characterized by the specific distributional assumption on ϵ in Equation (1). The coefficients $\boldsymbol{\beta}$ and the $J-1$ threshold parameters $\alpha_1, \dots, \alpha_{J-1}$ are estimated jointly by maximum-likelihood, using numerical optimization methods.

The overall question of interest can now be formulated as follows: will OLS, based on an equidistant-coded SWB variable y, which is possibly only an ordinal representation of an interval-scale level latent variable y^*, produce estimates of β_1/β_2 or $\beta_1/\text{sd}(y^*)$ that are quantitatively similar to corresponding estimates from an ordered regression?

2.4 Distributional Assumptions

The most common assumption regarding the error term ϵ in Equation (1), is that it follows either a normal or a logistic distribution, yielding the *probit* and the *logit* model, respectively. Both of these distributions are symmetric and bell-shaped, with the logistic having somewhat fatter tails. Although these are by far the most used models, the choice is seldom motivated. A symmetric distribution can be considered a neutral assumption, in some sense. Moreover, the normal distribution underlying the probit can be rationalized by the Central Limit Theorem, if the error term is assumed to be a sum of independent variables. In general, this is a strong assumption, however. The matter is complicated further by the fact that the interpretation of the error term varies depending on the particular application and on what covariates are controlled for.

In addition to the normal and the logistic, I will therefore also consider two skewed distributions: the extreme value distribution for the minimum and the maximum, which are left- and right-skewed, respectively. When used in an ordered regression context, these distributions yield the so-called log-log and complementary log-log models and I will refer to them as the *loglog* and the *cloglog*. The distributions of these models are described further in Appendix A.[13]

2.5 Model Comparison

Regardless of whether we are interested in ratios of the form β_1/β_2 or standardized coefficients $\beta_1/\text{sd}(y^*)$, the Pearson correlation, $\text{cor}(\hat{\boldsymbol{\beta}}^A, \hat{\boldsymbol{\beta}}^B)$, between coefficient estimates from model A and B (as used e.g by Clark, 2016), is not an appropriate measure of model similarity, since it may mask meaningful differences between the two sets of estimates. To realize this, consider the example when $\hat{\boldsymbol{\beta}}^A = [2 \quad 1]'$ and $\hat{\boldsymbol{\beta}}^B = [3 \quad 2]'$. Even though the correlation is 1, the

[13] It would of course be desirable to be able to assess which ordered regression model is more plausible, especially in the light of the deviating results for the loglog and the cloglog. I have done some preliminary analyis in this direction, exploiting the fact that the probit, the loglog and the cloglog are all nested by the log-gamma distribution, and are obtained as special cases when fixing the value of an additional shape parameter. When I estimate this parameter freely, I find that the optimal fit typically lies somewhere inbetween the loglog and the probit, thus suggesting that the distribution of the error term is left-skewed (though less so than for the loglog model). The cloglog typically fits the data worst. This fact is reflected in the log-likelihood values presented in the results tables of the current paper. Further analysis using the log-gamma distribution will be incorporated in a future version of this paper.

estimates from A and B imply different marginal rates of substitution of -2 and $-3/2$, respectively. Technically, the correlation involves demeaning of the vector elements—but this operation is not admissible for coefficients β, which are on a signed ratio scale.

Instead, I compare models directly in terms of differences in particular $\hat{\beta}_1/\hat{\beta}_2$ or $\hat{\beta}_1/\text{sd}(y^*)$. In the first case, I denote difference between models A and B as $\Delta\hat{\beta}_1/\hat{\beta}_2 = \hat{\beta}_1^A/\hat{\beta}_2^A - \hat{\beta}_1^B/\hat{\beta}_2^B$. I also make comparisons in relative terms in order to facilitate comparison across variables, i.e. $\frac{\hat{\beta}_1^A/\hat{\beta}_2^A - \hat{\beta}_1^B/\hat{\beta}_2^B}{|\hat{\beta}_1^A/\hat{\beta}_2^A|}$.

3 Unemployment and Life Satisfaction

In the first empirical example, I focus on the association between unemployment and life satisfaction. This association has been studied extensively in the SWB literature, see e.g. Clark and Oswald (1994), Clark et al. (2008), Knabe and Rätzel (2011) and Winkelmann and Winkelmann (1998). It is typically found to be negative and strong in magnitude, in comparison to many other socio-economic factors.

3.1 Data

The difference between OLS and ordered regression estimates may vary depending on what life satisfaction measure is used, e.g. due to the number of response categories and if the response scale is numerically labelled. To assess this, I use three different data sets with three different scales: the GSS (Smith et al., 2015), the PSID (Institute for Social Research, U. of Michigan, 2017) and the ESS (Norwegian Centre for Research Data, 2017). In the GSS case, I use a pooled cross-section from 1972–2014, consisting of $49,350$ individuals. The PSID data is a single cross-section from 2013 (the only one including life satisfaction), consisting of $8,446$ individuals. I use the Swedish portion of the ESS data, a pooled cross-section from 2002–2014, with $11,870$ individuals.

As mentioned in the introduction, the GSS has a three-category verbal response scale, whereas the life satisfaction measure in ESS is a numeric eleven-point scale. The PSID lies inbetween, with a five-category verbal scale. The questions and response scales are as follows

- GSS: Taken all things together, how happy would you say that you are these days? Not very happy, Pretty happy, Very happy

- PSID: Please think about your life-as-a-whole. How satisfied are you with it? Are you completely satisfied, very satisfied, somewhat satisfied, not very satisfied, or not at all satisfied?

- ESS: All things considered, how satisfied are you with your life as a whole nowadays? 0–10 response scale with endpoints labelled Extremely dissatisfied and Extremely satisfied

The distribution of life satisfaction, for each data set, is shown in Table B.1 in Appendix B. In this table, the satisfaction scores are assigned numbers ranging from 1–11, but any set of equidistant numbers could be used. The dependent variable used in the OLS estimations are coded likewise, but for comparison with the ordered response models and across scales (data sets), I scale the OLS coefficient estimates by dividing with the standard deviation of the cardinally coded outcome. These coefficients can be transformed back into the original numeric scale by multiplying with the standard deviation of the respective outcomes, also reported in Table B.1.

I have coded all independent variables to be as comparable as possible, e.g. by collapsing the variables for employment status and subjective health into an equal number of categories. An important difference is that income in the PSID is reported as a continuous variable, whereas I have derived a numeric variable for the GSS and the ESS, using the midpoints of the pre-defined response categories for income. All three samples include only individuals without partial non-response in the variables used, with the exception of some categorical variables for which missing values could be assigned to an "other" category. Throughout, I use appropriate sample weights included in each of the data sets.

3.2 Empirical Results

I estimate the equation

$$y^* = \underbrace{\beta_{ue}\mathrm{I}(ue) + \beta_{inc}\log(inc) + \boldsymbol{z'\gamma}}_{=\boldsymbol{x'\beta}} + \epsilon, \tag{2}$$

where y^* is life satisfaction, $\mathrm{I}(ue)$ is an indicator of unemployment (relative to working part or full time), inc is a measure of per-spouse net household income, $\boldsymbol{z}$ is a vector of control variables with associated coefficients $\boldsymbol{\gamma}$, and ϵ

is an iid error term. The vector $\boldsymbol{z}$ consists of the following categorical control variables: other employment (vs employed in a part time or full time job), subjective health (good vs bad), marital status (married/cohabiting vs not), sex, age (25–44, 45–64 and 65+, vs 18–24), and a set of survey-year dummies when applicable.

As explained in Section 2.2, we are primarily interested in the ratio β_{ue}/β_{inc}, which measures the impact of unemployment on life satisfaction, relative to the impact of income. Henceforth, I will refer to this simply as the relative impact of unemployment. Income is entered in logarithmic form, to capture its diminishing marginal effect on life satisfaction.[14] We can therefore interpret $m = e^{|\beta_{ue}/\beta_{inc}|}$ as the m-fold change in income that would give the same change in life satisfaction as changing status between employment and unemployment.

In particular, I am interested in whether the estimate $\hat{\beta}_{ue}/\hat{\beta}_{inc}$ differs depending on whether it is estimated by OLS, under the assumption that y^* is observed (in the form of an equidistant-coded variable), or whether the ratio is estimated by an ordered regression model, under the assumption that we only have an ordered manifestation y of y^*.

Table 2 shows the estimates from the GSS. The different models are organized in columns and the first two rows show the coefficients for unemployment and log-income, which are scaled by the standard deviation of life satisfaction. For the ordered models, this standard deviation is estimated by $\mathrm{sd}(y^*) = \sqrt{\hat{\boldsymbol{\beta}}' \mathrm{var}(\boldsymbol{x})\hat{\boldsymbol{\beta}} + \mathrm{var}(\epsilon)}$, where $\boldsymbol{x}$ denotes all included covariates, with associated estimated coefficients $\hat{\boldsymbol{\beta}}$, and $\mathrm{var}(\epsilon)$ is the normalized error-term variance of the model in question. This variance is $\pi^2/3$ for the logit, 1 for the probit, and $\pi^2/6$ for the loglog and the cloglog.

The coefficient estimate for unemployment ranges between -0.36 and -0.26 standard deviations of life satisfaction—a rather sizeable impact, given that it can be interpreted in terms of the non-pecuniary impact of unemployent on life satisfaction, since it is estimated controlling for income. The coefficient estimates of log-income range between 0.09 and 0.14, which is quite small. For instance, a doubling of income, based on the OLS estimate, implies a satisfaction change of $0.12 \cdot \log(2) = 0.08$, i.e. 8% of standard deviation.

To be clear, these cross-sectional estimates should be interpreted as purely descriptive, rather than causal. This is not a major concern in this context,

[14] Layard et al. (2008) find that the marginal utility of income on life satisfaction declines somewhat faster than implied by the logarithmic form, but for simplicity I follow most of the previous literature and stick with the log-form.

though, since we are interested in contrasting results from OLS and ordered regressions, based on the same data. Presumably, the pattern of differences between OLS and ordered regression estimates would be similar when applied to similar data, in which some exogenous source of variation in unemployment or other variables is used, but this should be assessed in future research.

Table 2: Empirical results, GSS

	OLS	logit	probit	loglog	cloglog
$\hat{\beta}_{ue}$	-0.32	-0.35	-0.36	-0.30	-0.26
$\hat{\beta}_{inc}$	0.12	0.13	0.14	0.12	0.09
$\hat{\beta}_{ue}/\hat{\beta}_{inc}$	-2.68	-2.68	-2.64	-2.50	-2.80
$\Delta\hat{\beta}_{ue}/\hat{\beta}_{inc}$		0.00	0.04	0.19	-0.11
		$(-0.04, 0.05)$	$(0.03, 0.05)$	$(0.01, 0.36)$	$(-0.41, 0.15)$
R^2	0.09	0.11	0.12	0.08	0.08
Log-lik.		$-44,005$	$-43,987$	$-44,128$	$-44,336$

$n = 49,350$. The dependent variable is life satisfaction measured on a three-category verbal scale. Coefficients from ordered regressions are divided by sd(y^*). $\Delta\hat{\beta}_{ue}/\hat{\beta}_{inc}$ is the ordered regression ratio minus the OLS ratio. All estimations include year dummies and controls for sex, age, marital status, health and employment other than employed or unemployed. 95% confidence intervals are based on a non-parametric bootstrap. $R^2 = \text{var}(\hat{y}^*)/\text{var}(y^*)$ for ordered regression models.

Moving to the statistics of main interest, the coefficient ratios in the third row, we see that they range between -2.80 and -2.68. The economic interpretation of the OLS estimate is that the life satisfaction impact of unemployment is equivalent to a 15-fold income change ($e^{|-2.68|} = 14.6$). The fourth row shows the difference between the ratio estimates from the ordered regressions and OLS. The estimates of the commonly used ordered models, the logit and the probit, are very close to the OLS estimate, with the logit yielding practically the same point estimate as OLS.

Statistical inference is non-trivial in this context, since we are interested in testing for statistical differences between estimates from different estimation approaches, which can be expected to be correlated since they are based on the same data. My approach is to use a paired non-parametric bootstrap. I draw 1,000 data sets by sampling individuals with replacement, whereafter I compute equally many differences between pairs of ratio estimates from OLS and ordered regressions. I use the distances between the 2.5th and 97.5th percentiles relative to the median of this bootstrap distribution, to construct 95% confidence intervals around the original estimate of the difference in ratios. Confidence intervals for the coefficients and the ratio point-estimates are omitted, but note that these are significantly different from zero throughout.

The difference between the OLS and the probit estimate of 0.04 is statistically significant, but small in magnitude, corresponding to a 1.5% difference relative to the OLS estimate. The differences between OLS and the loglog and the cloglog are larger in magnitude and go in different directions but it is only the loglog difference of 0.19 that is statistically significant. This corresponds to a 7% deviation relative to the OLS estimate, so it is fairly small, but still non-negligible. In economic terms, the loglog estimate implies that the life satisfaction cost of unemployment corresponds to a 12-fold income change ($e^{|-2.50|} = 12.2$).

Table 3: Empirical results, PSID

	OLS	logit	probit	loglog	cloglog
$\hat{\beta}_{ue}$	−0.27	−0.28	−0.27	−0.17	−0.22
$\hat{\beta}_{inc}$	0.08	0.09	0.09	0.10	0.04
$\hat{\beta}_{ue}/\hat{\beta}_{inc}$	−3.22	−3.27	−3.11	−1.64	−5.96
$\Delta\hat{\beta}_{ue}/\hat{\beta}_{inc}$		−0.06	0.11	1.58	−2.75
		(−0.53, 0.46)	(−0.13, 0.45)	(0.64, 3.44)	(−21.14, −0.08)
R^2	0.15	0.16	0.16	0.11	0.08
Log-lik.		−202, 461	−202, 858	−202, 180	−207, 279

$n = 8,446$. The dependent variable is life satisfaction measured on a five-category verbal scale. Coefficients from ordered regressions are divided by sd(y^*). $\Delta\hat{\beta}_{ue}/\hat{\beta}_{inc}$ is the ordered regression ratio minus the OLS ratio. All estimations include controls for sex, age, marital status, health and employment other than employed or unemployed. 95% confidence intervals are based on a non-parametric bootstrap. $R^2 = \text{var}(\hat{y}^*)/\text{var}(y^*)$ for ordered regression models.

The results for the PSID are shown in Table 3. The absolute magnitudes of the coefficient estimates for both unemployment and income are smaller than for the GSS. But relatively speaking, the income estimates are smaller, so as to produce ratio estimates of greater absolute magnitude (i.e. more negative), except in the loglog case. The PSID sample is substantially smaller than the GSS sample, however, which is reflected in the precision of the differences in ratio estimates. Hence, we cannot reject the null hypotheses that the OLS estimate is equal to the logit and the probit estimates. The differences between OLS and the loglog and cloglog estimates are statistically significant, though, and the difference is clearly economically relevant, with the loglog estimate being about half as large in absolute magnitude, whereas the cloglog estimate is almost twice as large in absolute magnitude.

Table 4: Empirical results, ESS (Sweden)

	OLS	logit	probit	loglog	cloglog
$\hat{\beta}_{ue}$	−0.56	−0.46	−0.45	−0.35	−0.31
$\hat{\beta}_{inc}$	0.16	0.14	0.15	0.16	0.06
$\hat{\beta}_{ue}/\hat{\beta}_{inc}$	−3.43	−3.22	−3.07	−2.19	−4.85
$\Delta\hat{\beta}_{ue}/\hat{\beta}_{inc}$		0.21	0.36	1.24	−1.41
		(−0.27, 0.65)	(0.00, 0.70)	(0.78, 1.99)	(−6.17, 0.24)
R^2	0.16	0.15	0.16	0.09	0.09
Log-lik.		$-20,324$	$-20,332$	$-20,351$	$-20,601$

$n = 11,870$. The dependent variable is life satisfaction measured on an eleven-point numeric scale. Coefficients from ordered regressions are divided by $\text{sd}(y^*)$. $\Delta\hat{\beta}_{ue}/\hat{\beta}_{inc}$ is the ordered regression ratio minus the OLS ratio. All estimations include year dummies and controls for sex, age, marital status, health and employment other than employed or unemployed. 95% confidence intervals are based on a non-parametric bootstrap. $R^2 = \text{var}(\hat{y}^*)/\text{var}(y^*)$ for ordered regression models.

Finally, we have the results based on the Swedish portion of the ESS data in Table 4. Both the coefficient estimates for unemployment and income are larger in absolute terms, compared to both the US data sets, but the ratio estimates are more similar to the PSID than to the GSS. The differences between the ratio estimate from OLS and those from the logit and the cloglog are not statistically significant. The probit difference is statistically significant, however, and the difference in magnitude corresponds to a 10% difference relative to the OLS estimate. Taking the estimates at face value, the OLS estimate implies that the life satisfaction impact of unemployment equals a 31-fold income change, whereas the probit implies a 22-fold income change. This difference is clearly economically significant, although it also highlights the sensitivity of ratio estimates involving variables in log-form. As for the GSS and the PSID, the loglog estimate is smaller than the OLS estimate in absolute terms (i.e. less negative), and the difference is statistically and economically significant.

Summing up the results from the three data sets, I find statistically significant differences, between OLS and ordered regression estimates of the relative impact of unemployment, in six out of twelve comparisons. Five of these differences are economically relevant, but four of them are produced by the loglog and cloglog model, which are rarely used in applied work. The results are thus somewhat mixed, but at least for the GSS and the PSID, the similarity between OLS on one hand, and the logit and probit on the other hand, is striking.

All estimations include a set of socio-economic control variables, and one might ask whether the pattern of differences between OLS and ordered regression estimates are similar for these. I present estimates of these differences in Table B.2 (GSS), Table B.3 (PSID) and Table B.4 (ESS) in Appendix B.[15] Broadly speaking, the results for the other variables are in line with those for unemployment. For both the GSS an the PSID, the differences between OLS and logit and probit are small in magnitude, although some of them are statistically significant (especially for the GSS, for which the power to detect such differences is larger). The differences between OLS and the loglog and cloglog are substantial for the majority of the variables, however. In the ESS data, several of the logit and probit estimates also differ from the OLS estimates, both in terms of statistical and economic significance, as was found to be the case for unemployment. To take an example, the logit estimate for the relative impact of being married or cohabiting (as compared to not), differs by 17%, compared to the OLS estimate.

3.3 Simulation Results

Even though OLS, logit and probit produce similar empirical results, at least for the GSS and the PSID, this does not prove that these estimators are consistent with respect to the true parameters of interest. Similarity of the empirical results is a neccesary, but not a sufficient, condition for consistency of the set of estimators taken as a whole. In other words, these estimators may all be inconsistent, and their similarity may simply be a function of the data.

In order to shed light on the question of consistency, I assess how different estimators perform when applied to simulated life satisfaction data, for which the true parameter values are known. Such simulations can be done in many different ways, with respect to the assumptions that are made about the latent

[15]The coefficients for the control variables are themselves all significantly different from zero on at least a 5% significance level, with the exception of the dummy for other employment (only significant in the PSID) and the dummy for age 65+ (insignificant in the PSID). Results for the year dummies included in the GSS and the ESS estimations are not shown.

variable. The idea behind the approach taken here is to mimic the structure of the observed data as closely as possible.

I generate different data sets, each of which is based on the GSS, the PSID or the ESS, and on one of the four ordered regression models. In each case, I generate the latent life satisfaction variable y^* as in the linear model in Equation (1), but using the empirical coefficient estimates $\hat{\boldsymbol{\beta}}$ from the estimations in Section 3.2 as the true parameter values. I combine these coefficients with the actual covariate values $\boldsymbol{x}$ and a parametrically random-generated error term, in accordance with the assumed error-term distribution of the ordered regression model in question. Thereafter, I take the estimates $\hat{\boldsymbol{\alpha}}$ (estimated jointly with $\hat{\boldsymbol{\beta}}$) to be the true cutoffs, which I use to map y^* into an observed ordered variable y. For instance, in the GSS–logit case, I take $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\alpha}}$ from the estimation in the second column of Table 2, and the error term is generated according to the logistic distribution with variance $\pi^2/3$. This procedure ensures that the simulated distribution of observed satisfaction scores is similar to the actual distribution of observed scores. The approximate variance shares of the covariates and the error term, respectively, are also preserved.

In order to vary the sample size and also to induce variation in $\boldsymbol{x}$, I resample the original data in a bootstrap fashion, drawing $n = 2,000$, $n = 10,000$ or $n = 50,000$ individuals (i.e. individual $\boldsymbol{x_i}$), creating a data set that may be smaller or larger than the original one. I replicate this process $1,000$ times, for each combination of original data, model and sample size ($3 \times 4 \times 3 = 36$ combinations).

For each of these 36 sets of $1,000$ data sets, I then proceed to do the same set of estimations as in the empirical analysis above, i.e. I estimate the ratio between the coefficients for unemployment and log-income, by means of OLS (on the equidistant-coded y) and four different ordered regressions. The resulting distributions of ratio estimates include some extreme outliers, since the ratio goes to (minus) infinity when the coefficient estimate for income happens to be close to zero. I therefore present the results in terms of the median (rather than the mean) of the ratio estimates minus the true coefficient. This measure should capture the asymptotic bias, $\text{plim}(\hat{\beta}_{ue}/\hat{\beta}_{inc}) - \beta_{ue}/\beta_{inc}$, as the sample size grows. I use the median absolute deviation (rather than the standard deviation) of these estimates, as a robust measure of spread.

The simulation results based on the GSS data are shown in Table 5. Note that each cell summarizes the distribution of $1,000$ regression estimates. Start-

Table 5: Simulation results based on GSS

	OLS	logit	probit	loglog	cloglog
$\epsilon \sim$ logit					
$n = 2,000$	−0.071	−0.050	−0.028	0.053	−0.064
	(0.851)	(0.852)	(0.848)	(0.962)	(0.912)
$n = 10,000$	0.003	0.008	0.038	0.033	0.019
	(0.356)	(0.364)	(0.353)	(0.392)	(0.404)
$n = 50,000$	0.003	0.007	0.036	0.009	0.045
	(0.171)	(0.169)	(0.169)	(0.191)	(0.195)
$\epsilon \sim$ probit					
$n = 2,000$	0.008	0.004	0.034	0.130	−0.108
	(0.866)	(0.872)	(0.858)	(0.926)	(0.980)
$n = 10,000$	−0.033	−0.026	0.001	0.056	−0.017
	(0.396)	(0.400)	(0.392)	(0.421)	(0.425)
$n = 50,000$	−0.032	−0.038	0.002	0.033	−0.017
	(0.173)	(0.177)	(0.173)	(0.180)	(0.188)
$\epsilon \sim$ loglog					
$n = 2,000$	−0.104	−0.174	−0.072	0.035	−0.096
	(0.912)	(0.960)	(0.908)	(0.857)	(1.068)
$n = 10,000$	−0.087	−0.151	−0.055	0.011	−0.013
	(0.393)	(0.396)	(0.390)	(0.384)	(0.478)
$n = 50,000$	−0.079	−0.156	−0.051	0.006	−0.014
	(0.185)	(0.195)	(0.182)	(0.169)	(0.218)
$\epsilon \sim$ cloglog					
$n = 2,000$	0.067	0.134	0.098	0.175	−0.115
	(1.208)	(1.180)	(1.193)	(1.388)	(1.183)
$n = 10,000$	0.050	0.125	0.090	0.125	−0.013
	(0.532)	(0.505)	(0.525)	(0.652)	(0.471)
$n = 50,000$	0.052	0.128	0.090	0.121	0.003
	(0.243)	(0.232)	(0.239)	(0.305)	(0.229)

All simulations are based on $1,000$ replications. Data within each panel are generated using the same error term distribution and are based on fixed parameters $\boldsymbol{\beta}$ and cutoffs $\boldsymbol{\alpha}$, taken from the corresponding ordered regression estimations in Table 2. Variation in sample size is obtained by resampling the covariates $\boldsymbol{x}$ of the original data. Each cell shows the median estimate of the ratio between the coefficients for unemployment and log-income, minus the true ratio. The median absolute deviation of these estimates are shown in parenthesis. The true ratios are -2.68, -2.64, -2.50 and -2.80, for the logit, probit, loglog and cloglog, respectively.

ing with the logit-based simulations in the top panel, OLS (first column) appears to be consistent with respect to the true coefficient ratio. The logit (second column) is also consistent, which is less surprising given that the data are generated using the logistic distribution. The OLS and the logit estimates are equally precise and converge at the same rate.[16] The probit (third column) appears to be inconsistent, although the asymptotic bias is small, only 1.3% relative to the true ratio (based on the estimates with $n = 50,000$). The loglog (fourth column) appears to be consistent, whereas the cloglog (fifth column) has a small asympotic bias of 1.7% relative to the true ratio. The loglog and the cloglog have a somewhat larger spread compared to the other estimators.

Moving to the probit-based data in the second panel, we note first that the probit is consistent, as expected. The asympotic bias is −1.2% for OLS, −1.4% for the logit, 1.3% for the loglog and −0.6% for the cloglog.

For the loglog-based data, all estimators except the loglog itself are inconsistent, with varying degrees of asymptotic bias. For example, the deviation between the median logit estimate and the true ratio of −2.50, based on $n = 50,000$, is −0.156. In relative terms, the asymptotic bias is thus −6.2%. This is an economically meaningful difference, even though it is not huge. The asymptotic bias of OLS is about half as large.

There is a similar pattern for the simulations based on the cloglog. All estimators except the cloglog are inconsistent, with the logit again having the largest asymptotic bias, equal to 5.1% relative to the true ratio. In this case, OLS outperforms all misspecified ordered regression estimators.

Summing up the GSS-based simulation results, we see that no estimator is consistent across all four data-generating processes considered. The asymptotic bias is small in general, though economically meaningful in some cases. OLS does not perform worse than the ordered regression models in cases when the error-term distribution is misspecified.

The simulations based on the PSID data are shown in Table 6. As expected, all ordered regression models are consistent when their respective distributional assumptions hold. OLS appears to be inconsistent in all cases, except when the error term follows the cloglog distribution. The asymptotic bias of OLS, relative to the true ratio, varies between −2.2% and −8.5%. The logit and the probit outperform OLS when the error term follows either a logistic, a normal

[16] As expected, the median absolute deviation shrinks by a factor of approximately $\sqrt{5}$, as the sample size increases by a factor 5.

or a minimum-value (loglog) distribution, but not when it has a maximum-value (cloglog) distribution. The cloglog estimator performs best across all distributions and the loglog estimator performs worst.

Finally, the ESS simulations are shown in Table 7. OLS is inconsistent throughout, and its asymptotic bias is rather sizeable also under the standard logit and probit assumptions. For example, the asymptotic bias of OLS amounts to 20.6% of the true ratio, when the simulated data is based on the probit. This is an order of magnitude larger than the bias found in the GSS simulations. The asymptotic bias of the logit and the probit is also sizeable under the loglog and cloglog assumptions, and vice versa.

We might ask, as we did for the empirical results, whether the simulation results regarding the relative impact of unemployment generalizes to the other independent variables. I present simulation results for these variables in Appendix B, in Table B.5 (GSS), Table B.6 (PSID) and Table B.7 (ESS). I focus on the case when the simulated data is based on the probit model with $n = 50,000$.[17] These results are presented in terms of the asymptotic bias divided by the absolute value of the true coefficient (so as to maintain the sign of the bias), in order to facilitate comparisons across variables.

Restricting attention further to the performance of OLS, it turns out that the ESS-based unemployment estimate, which is by -21%, is somewhat of an outlier, both in comparison to the other data sets and in comparison with other variables within the ESS. Still, OLS performs markedly worse in the ESS-based simulations also for the other variables. The average (absolute) bias of OLS, computed across all variables except unemployment, is 0.8% for the GSS, 1.5% for the PSID and 4.3% for the ESS. This pattern is similar when the simulated data is based e.g. on the logit (results not shown). OLS thus appears to perform worse when the ordered life satisfaction variable has more categories, as in the ESS. With the exception of the unemployment estimate, the asympotic bias of OLS is not large, however, but it is hard to say whether this result generalizes to other variables in other contexts.

3.4 Summary

Summing up the results, we see a rather clear correspondence between the empirical results and the simulations. The differences between OLS and ordered

[17] I omit results related to other employment and age 65+, since the coefficients on these two variables are not significantly different from zero in all data sets.

Table 6: Simulation results based on PSID

	OLS	logit	probit	loglog	cloglog
$\epsilon \sim$ logit					
$n = 2,000$	-0.144	-0.003	0.029	0.264	0.126
	(1.556)	(1.508)	(1.506)	(1.642)	(1.562)
$n = 10,000$	-0.166	-0.039	-0.005	0.199	-0.038
	(0.753)	(0.720)	(0.710)	(0.737)	(0.785)
$n = 50,000$	-0.073	0.012	0.065	0.262	-0.013
	(0.309)	(0.287)	(0.292)	(0.308)	(0.369)
$\epsilon \sim$ probit					
$n = 2,000$	-0.170	-0.004	-0.004	0.253	0.008
	(1.625)	(1.555)	(1.556)	(1.513)	(1.637)
$n = 10,000$	-0.052	0.034	0.092	0.204	0.034
	(0.674)	(0.673)	(0.661)	(0.687)	(0.728)
$n = 50,000$	-0.119	-0.018	0.004	0.176	-0.030
	(0.340)	(0.329)	(0.321)	(0.330)	(0.353)
$\epsilon \sim$ loglog					
$n = 2,000$	-0.113	-0.041	-0.045	0.089	0.062
	(1.075)	(1.116)	(1.057)	(0.830)	(1.398)
$n = 10,000$	-0.158	-0.107	-0.091	0.011	-0.041
	(0.449)	(0.456)	(0.439)	(0.385)	(0.570)
$n = 50,000$	-0.139	-0.085	-0.061	0.005	-0.020
	(0.193)	(0.199)	(0.192)	(0.169)	(0.262)
$\epsilon \sim$ cloglog					
$n = 2,000$	2.255	2.304	2.310	3.620	1.506
	(3.098)	(3.116)	(3.018)	(3.208)	(3.100)
$n = 10,000$	0.226	0.354	0.369	0.859	0.118
	(1.991)	(1.918)	(1.946)	(2.272)	(1.888)
$n = 50,000$	-0.003	0.211	0.191	0.300	0.003
	(0.929)	(0.907)	(0.891)	(1.087)	(0.872)

All simulations are based on 1,000 replications. Data within each panel are generated using the same error term distribution and are based on fixed parameters $\boldsymbol{\beta}$ and cutoffs $\boldsymbol{\alpha}$, taken from the corresponding ordered regression estimations in Table 3. Variation in sample size is obtained by resampling the covariates $\boldsymbol{x}$ of the original data. Each cell shows the median estimate of the ratio between the coefficients for unemployment and log-income, minus the true ratio. The median absolute deviation of these estimates are shown in parenthesis. The true ratios are -3.27, -3.11, -1.64 and -5.96, for the logit, probit, loglog and cloglog, respectively.

Table 7: Simulation results based on ESS

	OLS	logit	probit	loglog	cloglog
$\epsilon \sim$ logit					
$n = 2,000$	−0.526	0.010	0.131	0.437	−0.146
	(1.174)	(0.905)	(0.893)	(0.927)	(1.087)
$n = 10,000$	−0.580	−0.012	0.084	0.372	−0.133
	(0.494)	(0.384)	(0.391)	(0.406)	(0.469)
$n = 50,000$	−0.537	0.019	0.108	0.402	−0.149
	(0.202)	(0.160)	(0.162)	(0.181)	(0.214)
$\epsilon \sim$ probit					
$n = 2,000$	−0.528	0.012	0.049	0.303	−0.160
	(1.128)	(0.907)	(0.919)	(0.919)	(1.090)
$n = 10,000$	−0.584	−0.053	0.027	0.300	−0.181
	(0.502)	(0.429)	(0.397)	(0.409)	(0.461)
$n = 50,000$	−0.632	−0.073	0.000	0.262	−0.187
	(0.217)	(0.185)	(0.174)	(0.175)	(0.208)
$\epsilon \sim$ loglog					
$n = 2,000$	−0.632	−0.303	−0.179	−0.008	−0.130
	(0.739)	(0.693)	(0.639)	(0.531)	(0.978)
$n = 10,000$	−0.645	−0.299	−0.194	0.020	−0.168
	(0.363)	(0.349)	(0.317)	(0.260)	(0.447)
$n = 50,000$	−0.640	−0.293	−0.188	0.003	−0.120
	(0.151)	(0.146)	(0.131)	(0.109)	(0.178)
$\epsilon \sim$ cloglog					
$n = 2,000$	0.679	1.161	1.068	1.857	0.544
	(2.322)	(1.835)	(1.907)	(2.142)	(1.936)
$n = 10,000$	−0.299	0.450	0.373	0.550	−0.065
	(1.303)	(1.054)	(1.037)	(1.313)	(1.077)
$n = 50,000$	−0.273	0.453	0.416	0.565	−0.007
	(0.593)	(0.454)	(0.470)	(0.591)	(0.454)

All simulations are based on $1,000$ replications. Data within each panel are generated using the same error term distribution and are based on fixed parameters $\boldsymbol{\beta}$ and cutoffs $\boldsymbol{\alpha}$, taken from the corresponding ordered regression estimations in Table 4. Variation in sample size is obtained by resampling the covariates $\boldsymbol{x}$ of the original data. Each cell shows the median estimate of the ratio between the coefficients for unemployment and log-income, minus the true ratio. The median absolute deviation of these estimates are shown in parenthesis. The true ratios are −3.22, −3.07, −2.19 and −4.85, for the logit, probit, loglog and cloglog, respectively.

regressions are small in the GSS, both in the empirical results and in the simulations. Larger differences, though not huge, are found for the ESS data in both the empirical results and the simulations. The cross-method differences in the simulation estimates for the PSID lie somewhere inbetween the GSS and the ESS, as for the empirical results. This finding is somewhat counterintutive, as one might think that the 11-point life satisfaction scale of the ESS approximates an interval scale variable better than the categorical 3-point scale of the GSS. It cannot be ruled out that these differences across the three data sets are due other differences in the data than the life satisfaction scales. This possibility does not seem very plausible, however, given the high degree of homogeneity of the independent variables used.

4 Country Differences in Life Satisfaction

In the second example, presented in this section, I estimate country-level mean shifts in life satisfaction. When estimated by OLS, this is equivalent to computing mean differences in the cardinally coded life satisfaction scores. This example is thus relevant for the literature on cross.country differences in SWB, as well as for policy initiatives concerned with national well-being levels, e.g. the OECD "Better Life Index".[18]

My approach here differs from that in the previous example, as I compare standardized coefficient estimates, rather than coefficient ratios. The estimations are based on micro-data, but the only variables included are a set of country-dummies. The coefficients from the ordered regressions are scaled by the standard deviation of the latent variable, $\mathrm{sd}(y^*)$, throughout, as described in Section 3.2. In this context, the results can thus be interpreted in terms of the international standard deviation of life satisfaction.

4.1 Data

I use all available ESS data from the 2014 wave, including $40,057$ individuals in 21 countries. The life satisfaction measure is the same as for the Swedish ESS data used in the previous example. The distribution for the international sample is shown in the rightmost column of Table B.1 in Appendix B, which also shows the international standard deviation of life satisfaction for this sample.

[18] `http://www.oecdbetterlifeindex.org`.

Table 8: Cross-country results, ESS

	$\hat{\beta}^{\text{OLS}}$	$\hat{\beta}^{\text{ordered}} - \hat{\beta}^{\text{OLS}}$			
	OLS	logit	probit	loglog	cloglog
Denmark	0.48	0.07*	0.09*	0.00*	−0.10*
Switzerland	0.35	0.03*	0.04*	−0.01	−0.09*
Norway	0.28	0.01*	0.01*	−0.01*	−0.11
Finland	0.27	−0.01	−0.01	0.02	−0.17*
Sweden	0.26	−0.01*	0.01*	0.00	−0.11*
Netherlands	0.15	−0.07	−0.07	0.05*	−0.21
Germany	0.08	0.01	0.00	−0.01*	−0.05*
Israel	0.06	0.03*	0.03	−0.02*	0.04*
Belgium	0.05	−0.05	−0.05	0.04	−0.14
United Kingdom	−0.03	0.00	−0.01	0.01*	−0.02*
Ireland	−0.13	−0.03*	−0.03*	0.05*	−0.02
Spain	−0.17	−0.01*	−0.01*	0.05*	0.02*
Poland	−0.18	0.01*	0.01*	0.04*	0.08*
Czech Republic	−0.30	−0.03	−0.02	0.10	0.03*
Slovenia	−0.36	0.01	0.01	0.10*	0.12*
France	−0.36	0.04*	0.01*	0.09*	0.12*
Estonia	−0.44	−0.01	−0.01	0.15	0.07*
Lithuania	−0.58	0.02	0.00	0.21*	0.09*
Hungary	−0.70	0.01	0.01	0.27*	0.13*
Portugal	−0.73	0.04	0.04	0.25*	0.23*
Mean abs. diff.		0.025	0.024	0.074	0.097
R^2	0.10	0.10	0.10	0.06	0.06
Log-likelihood		$-79,345$	$-79,482$	$-79,631$	$-80,011$
Rank-correlation		0.95	0.94	0.97	0.88

$n = 40,057$. * indicates that the difference w.r.t. the OLS estimate is significant on at least a 5% significance level. The first column shows coefficient estimates based on OLS, scaled with the (international) standard deviation of life satisfaction. The other columns show differences between these estimates and the estimates obtained from ordered regression models (ordered regression estimate minus OLS estimate), which are scaled by the standard deviation of the latent variable.

4.2 Empirical Results

The results are presented in Table 8. The OLS coefficients are shown in the first column, sorted in descending order. Austria is (arbitrarily) chosen as the reference country, so all the coefficients should be interpreted as the standardized difference in mean life satisfaction relative to Austria. The difference between Denmark and Portugal, the most and the least satisfied countries, respectively, spans 1.2 standard deviations. The remaining columns show the differences between the ordered regression estimates and the OLS estimates. As in the

the previous example, I test for statistical differences across models by means of a paired bootstrap. Differences that are significantly different from zero on *at least* a 5% significance level are indicated by asterisks. There are more such differences than could be expected by chance, and most differences are found between the OLS and the loglog and cloglog estimates.

Shifting attention to the magnitude of the differences, we see that they are quite small for both the logit and the probit. The largest differences between OLS and these models are found for Denmark—the difference between the logit and the OLS estimate amounts to 0.07 standard deviations, whereas the difference between the probit and the OLS estimate amounts to 0.09 standard deviations. These differences are about twice as large when expressed relative to the OLS estimate for Denmark (rather than in terms of sd-units).

The average difference, computed across the absolute values of all country mean-shifts, is only 0.025 sd-units andfor the logit and 0.024 sd-units for the probit. To the extent that we are interested in the country-ranking of life satisfaction, it is also quite insensitive to the method used. The rank-order correlation between the OLS and the logit estimates, computed by Kendall's Tau, is 0.95, whereas it is 0.94 between OLS and the probit.

In line with the results from the previous example, we find larger differences between OLS and the loglog and cloglog estimates. The loglog estimates differ by 0.074 sd-units on average and the cloglog estimates differ by 0.097 sd-units. The rank-order correlations between OLS and these estimates are 0.97 and 0.88, for the loglog and the cloglog, respectively.

In summary, the logit and the probit yield results similar to OLS, whereas the difference between OLS and the loglog and cloglog are somewhat larger, though not huge.

4.3 Simulation Results

As for the previous example concerning the relative impact of unemployment, I proceed to assess the consistency of different estimators in the cross-country context by means of simulated life satisfaction data. I use the same method as previously, the important difference here being the focus on the (standardized) coefficients, rather than ratios of coefficients. In the interest of space, I only present results for the coefficient for Sweden (relative to Austria). The results are presented in Table 9, in terms of the median simulation estimate minus the true coefficient, with the median absolute deviation in parenthesis.

Table 9: Cross-country simulation results, ESS, estimates for Sweden

	OLS	logit	probit	loglog	cloglog
$\epsilon \sim$ logit					
$n = 2,000$	−0.031	−0.004	0.009	−0.060	−0.057
	(0.092)	(0.092)	(0.098)	(0.091)	(0.079)
$n = 10,000$	−0.027	−0.001	0.009	−0.059	−0.057
	(0.042)	(0.044)	(0.046)	(0.041)	(0.037)
$n = 50,000$	−0.027	−0.001	0.009	−0.060	−0.058
	(0.017)	(0.018)	(0.019)	(0.017)	(0.015)
$\epsilon \sim$ probit					
$n = 2,000$	−0.036	−0.013	−0.001	−0.058	−0.063
	(0.089)	(0.095)	(0.096)	(0.086)	(0.080)
$n = 10,000$	−0.035	−0.016	−0.002	−0.063	−0.064
	(0.041)	(0.045)	(0.047)	(0.040)	(0.037)
$n = 50,000$	−0.033	−0.014	0.002	−0.064	−0.065
	(0.018)	(0.019)	(0.019)	(0.017)	(0.016)
$\epsilon \sim$ loglog					
$n = 2,000$	0.013	0.002	0.026	0.004	−0.072
	(0.084)	(0.089)	(0.096)	(0.080)	(0.081)
$n = 10,000$	0.011	0.000	0.026	0.001	−0.075
	(0.040)	(0.043)	(0.045)	(0.035)	(0.038)
$n = 50,000$	0.012	0.001	0.026	0.001	−0.075
	(0.017)	(0.017)	(0.018)	(0.016)	(0.015)
$\epsilon \sim$ cloglog					
$n = 2,000$	−0.005	0.017	0.023	−0.044	−0.002
	(0.098)	(0.093)	(0.101)	(0.088)	(0.079)
$n = 10,000$	−0.013	0.013	0.018	−0.047	−0.004
	(0.044)	(0.047)	(0.049)	(0.039)	(0.040)
$n = 50,000$	−0.009	0.016	0.022	−0.047	−0.002
	(0.020)	(0.020)	(0.021)	(0.019)	(0.017)

All simulations are based on 1,000 replications. Data within each panel are generated using the same error term distribution and are based on fixed parameters $\boldsymbol{\beta}$ and cutoffs $\boldsymbol{\alpha}$, taken from the corresponding ordered regression estimations in Table 8. Variation in sample size is obtained by resampling the covariates $\boldsymbol{x}$ of the original data. Each cell shows the median of the estimate of standardized mean shift of life satisfaction of Sweden, relative to Austria, with the mean absolute deviation of these estimates in parenthesis. The true coefficients are 0.26, 0.27, 0.26 and 0.15, for the logit, probit, loglog and cloglog, respectively.

As expected, all ordered regressions are consistent when their respective model assumptions are true. The fact that the estimates do not change much when the sample size increases, shows that the coefficient estimates converge much faster than the ratio estimates in the simulations of the previous example.

Turning to the OLS estimates, we see that they have a negative asymptotic bias of 0.027 sd-units, when the data is based on the logit and the probit. The true coefficients in the logit and probit simulations are 0.26 and 0.27, respectively, so the bias is around 10% in relative terms, which is economically relevant, though not huge. The magnitude of this bias is similar to the discrepancies between OLS and logit/probit in the empirical estimates. OLS performs better under the loglog and the cloglog, with a bias amounting to 0.01 and −0.01 sd-units, respectively.

Probit performs better than OLS when the logit is true, and vice versa for the logit, when the probit is true. Interestingly, and in line with some cases in the previous example, OLS outperforms the probit when the loglog is true, and both the logit and the probit are outperformed when the cloglog is true. We do not observe any cases in which OLS, the logit or the probit produce large asymptotic biases, however. The deviations of the loglog and the cloglog are sizeable in some cases, in line with both the empirical estimates and the simulation results of the previous example.

4.4 Summary

Summing up the cross-country results, I do not find large discrepancies between OLS and ordered regression estimates, or between OLS and the true parameters generated under the assumptions of the ordered models. Moreover, the patterns across empirical results and simulations appear to be consistent within different models, with e.g. the cloglog yielding the largest discrepancies.

5 Conclusion

The question motivating this paper is whether OLS and ordinal happiness regressions yield different results. At this point, it should be clear that this question fails to capture the complexity of the matter. First, because there are several possible ordinal regression models and, second, because similarity between OLS and a particular set of such models does not imply consistency with respect to the true parameters. My analysis sheds light on both of these issues.

In the first empirical example, I do indeed find that OLS and the two standard ordered models, the logit and probit, yield similar estimates of the relative impact of unemployment on life satisfaction. The largest discrepancies are found for the ESS data, but even in this case the magnitudes of the differences are rather small, though not negligible. In all three data sets considered, the loglog and the cloglog models, which are based on the assumption of a skewed error term, frequently yield estimates that are significantly different from OLS, in both a statistical and economic sense.

In the second empirical example, I show that a similar pattern holds for coefficient estimates of country-level mean-shifts in life satisfaction. The asymptotic bias of OLS when the data-generating process is consistent with the logit or the probit is in the order of 10%, relative to the true parameter.

I also compare the results of OLS and ordered regressions by means of simulated data, with known values of the true parameters. The simulation results are broadly in line with the empirical results, but provide additional insights. Generally speaking, OLS is not a consistent estimator of coefficient ratios from life satisfaction regressions, as was found to be the case in the simulations by Riedl and Geishecker (2014).

The simulations related to the first example suggest that the asymptotic bias of OLS is worse when there are more response categories, as in the ESS, compared to when there are few, as in the GSS. This result is somewhat unexpected, in light of the widespread belief that continuous variables are more suited for cardinal methods, and given that the response scale of the ESS is presented as a numerical scale. A possible explanation for this is that OLS applied to equidistant-coded scales with few categories behaves almost like a linear probability model, which in turn is robust to whether the data is ordinal. Especially so when there are few observations in some category so that they receive little weight, as is the case for the GSS, in which few people report being “Not so happy”.

As expected, the ordered models are consistent when their respective distributional assumptions are true. Logit and probit both tend to perform well when the error term follows a symmetric distribution but their asymptotic bias can be significant, and even worse than the bias of OLS, when the true distribution is in fact skewed.

The simulations related to the first example also highlight that, regardless of whether OLS or ordinal regression is used, large samples are needed to obtain precise estimates of coefficient ratios. The simulations related to the second example show that standardized coefficient estimates can be obtained with fairly high precision and without large biases, even if the model is misspecified and the sample size is moderate. The cost of using the latter type of estimates is that they cannot, by themselves, be mapped back to an economic problem.

It is premature, based on the results of this paper alone, to take a stance on whether SWB data should be analysed by means of OLS or an ordered regression approach. I would rather encourage further research, e.g. by carefully comparing the performance of these approaches in more realistic applications. However, I conclude with two general pieces of advice. First, researchers using the ordered regression approach should be aware that the assumed distribution of error term matters, and ideally, the choice should be motivated.

Second, sensitivity analyses based on comparing OLS and ordered regression estimates should do so also in quantitative terms, and not only in terms of the sign and statistical significance of the coefficients. At a minimum, coefficients should be scaled so as to be comparable, using either ratios or the latent variable standard deviation. In case the results are found to be similar, one should also be aware that this does not prove that the estimators are consistent with respect to the true parameters.

References

Bond, T. N. and K. Lang (2014). The sad truth about happiness scales. NBER Working Paper 19950, National Bureau of Economic Research.

Clark, A. (2016). SWB as a measure of individual well-being. In M. D. Adler and M. Fleurbaey (Eds.), *The Oxford handbook of well-being and public policy*, Chapter 18, pp. 518–552. Oxford: Oxford University Press.

Clark, A. and A. Oswald (1994). Unhappiness and unemployment. *The Economic Journal 104*(424), 648–659.

Clark, A. E., E. Diener, Y. Georgellis, and R. E. Lucas (2008). Lags and leads in life satisfaction: A test of the baseline hypothesis. *The Economic Journal 118*(529), F222–F243.

Dolan, P., T. Peasgood, and M. White (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology 29*(1), 94–122.

Easterlin, R. (1974). Does economic growth improve the human lot? Some empirical evidence. In *Nations and households in economic growth: Essays in honor of Moses Abramovitz*, pp. 89–125. Academic Press.

Ferrer-i-Carbonell, A. and P. Frijters (2004). How important is methodology for the estimates of the determinants of happiness? *Economic Journal 114*(497), 641–659.

Frey, B. S. and A. Stutzer (2000). Happiness, economy and institutions. *Economic Journal 110*(466), 918–938.

Genter, F. C. and V. T. Farewell (1985). Goodness-of-link testing in ordinal regression models. *Canadian Journal of Statistics 13*(1), 37–44.

Institute for Social Research, U. of Michigan (2017). Panel Study of Income Dynamics, public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI. The collection of data used in this study was partly supported by the National Institutes of Health under grant number R01 HD069609 and the National Science Foundation under award number 1157698.

Knabe, A. and S. Rätzel (2011). Quantifying the psychological costs of unemployment: The role of permanent income. *Applied Economics 43*(21), 2751–2763.

Layard, R., G. Mayraz, and S. Nickell (2008). The marginal utility of income. *Journal of Public Economics 92*(8–9), 1846–1857.

Norwegian Centre for Research Data (2017). European Social Survey rounds 1–7. Cumulative data file. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

OECD (2013). OECD guidelines on measuring subjective well-being. OECD Publishing, Paris. `http://dx.doi.org/10.1787/9789264191655-en`.

Riedl, M. and I. Geishecker (2014). Keep it simple: estimation strategies for ordered response models with fixed effects. *Journal of Applied Statistics 41*(11), 2358–2374.

Sacks, D., B. Stevenson, and J. Wolfers (2012). The new stylized facts about income and subjective well-being. *Emotion 12*(6), 1181–1187.

Schröder, C. and S. Yitzhaki (2017). Revisiting the evidence for a cardinal treatment of ordinal variables. *European Economic Review 92*, 337–358.

Smith, T. W., P. Marsden, M. Hout, and J. Kim (2015). General Social Surveys, 1972-2014 [machine-readable data file] /Principal Investigator, Tom W. Smith; Co-Principal Investigator, Peter V. Marsden; Co-Principal Investigator, Michael Hout; Sponsored by National Science Foundation. –NORC ed.– Chicago: NORC at the University of Chicago [producer]; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut [distributor]. 1 data file (57,061 logical records) + 1 codebook (3,567p.). – (National Data Program for the Social Sciences, No. 22).

Stevens, S. S. (1946). On the theory of scales of measurement. *Science 103*(2684), 677–680.

Stevenson, B. and J. Wolfers (2008). Economic growth and subjective well-being: Reassessing the Easterlin paradox. *Brookings Papers on Economic Activity 1*, 1–87.

Winkelmann, L. and R. Winkelmann (1998). Why are the unemployed so unhappy? Evidence from panel data. *Economica 65*(257), 1–15.

A Latent Error Distributions

The extreme value distribution for the minimum (with location and scale parameters normalized to 0 and 1), also called the Gumbel, or log-Weibull distribution, has cdf

$$F(x) = 1 - e^{-e^{x}},$$

The associated link function is the inverse cdf, and equal to

$$x = \log(-\log(1 - y)).$$

Because of this function, the associated ordered regression model is often called the loglog model. The maximum value distribution has cdf

$$F(x) = e^{-e^{-x}},$$

and link function

$$x = -\log(-\log(y)),$$

and the associated model is often called the complementary loglog model (or the cloglog model). The log-gamma distribution, which nests the probit, the loglog and the cloglog as special cases, is described by Genter and Farewell (1985). The loglog model and the cloglog model are available in e.g. R and Stata. The log-gamma model is implemented in the **ordinal** package in R.

The shapes of the loglog and the cloglog distributions, as well as the logit and probit for comparison, are shown in Figure A.1.

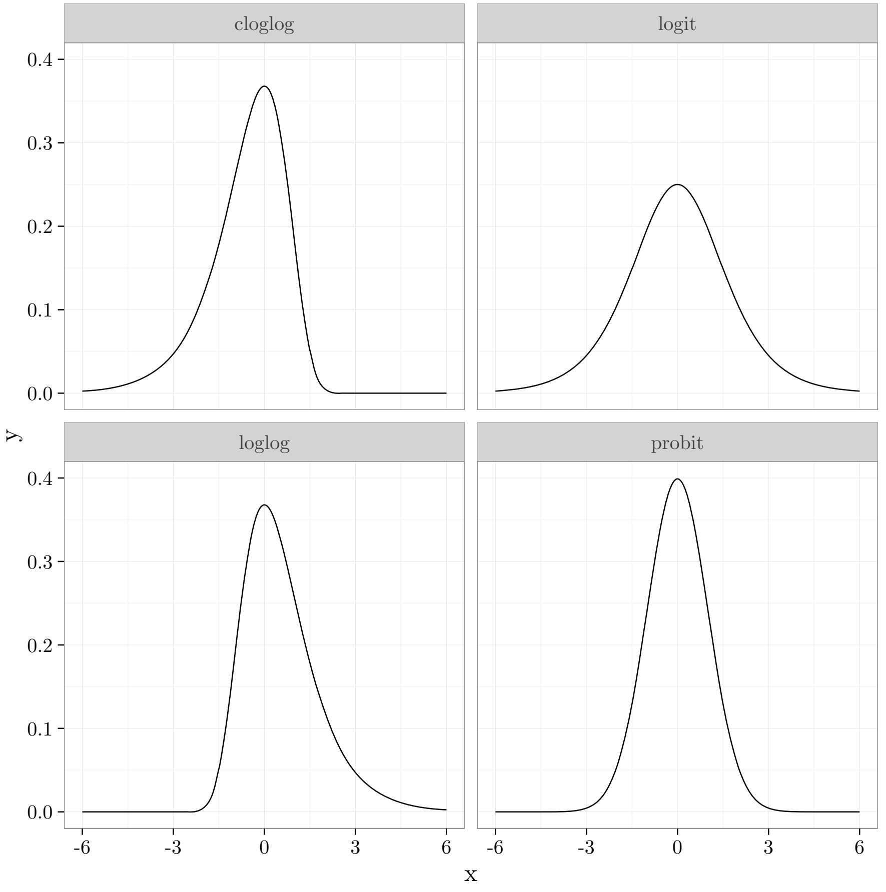

Figure A.1: Distributions of error terms of different ordered regression models

B Tables

Table B.1: Distribution of life satisfaction in GSS, PSID and ESS (%)

	GSS	PSID	ESS (SE)	ESS
1	12.1	1.2	0.3	1.2
2	56.1	3.8	0.3	0.8
3	31.8	28.1	0.7	1.9
4		46.4	1.4	3.3
5		20.5	1.9	3.9
6			5.3	10.2
7			5.5	8.4
8			16.1	17.3
9			31.3	26.4
10			22.5	15.3
11			14.8	11.4
std. dev.	0.63	0.85	1.71	2.15
n	$49,350$	$8,446$	$11,870$	$40,057$

Table B.2: Differences in estimates from ordered regressions and OLS of relative impacts of different variables on life satisfaction, GSS

	$\hat{\beta}_k/\hat{\beta}_{inc}$	$\Delta\hat{\beta}_k/\hat{\beta}_{inc}$			
	OLS	logit	probit	loglog	cloglog
Unemployed	−2.68	0.00	0.04*	0.19*	−0.11
Other employment	0.02	0.03	0.01*	−0.09	0.17
Good health	2.41	−0.04*	0.00	−0.28*	0.42*
Married or cohab.	3.88	−0.03	−0.01*	−0.61*	0.83*
Female	0.56	0.00	0.00	−0.08*	0.11*
Age 25–44	−0.93	0.00	0.01*	0.24*	−0.28*
Age 45–64	−0.93	0.01	0.01*	0.13*	−0.11
Age 65+	0.90	0.05*	0.02*	−0.18*	0.30*

* indicates that the difference w.r.t. the OLS estimate is significant on at least a 5% significance level. The first column shows coefficient ratio estimates based on OLS, using the coefficient for log-income as denominator. The other columns show differences between these estimates and the estimates obtained from ordered regression models (ordered regression estimate minus OLS estimate).

Table B.3: Differences in estimates from ordered regressions and OLS of relative impacts of different variables on life satisfaction, PSID

	$\hat{\beta}_k/\hat{\beta}_{inc}$	$\Delta\hat{\beta}_k/\hat{\beta}_{inc}$			
	OLS	logit	probit	loglog	cloglog
Unemployed	−3.22	−0.06	0.11	1.58	−2.75
Other employment	0.98	0.11	0.07	−0.44	1.33
Good health	6.62	−0.03	−0.09	−2.23*	3.48*
Married or cohab.	6.83	0.10	0.28*	−2.09*	5.77*
Female	1.64	−0.08	−0.01	−0.56*	1.43*
Age 25–44	−1.50	0.02	−0.02	0.44	−1.15
Age 45–64	−2.09	0.09	−0.01	0.51	−1.28
Age 65+	−0.30	0.17	0.07	−0.26	0.57

* indicates that the difference w.r.t. the OLS estimate is significant on at least a 5% significance level. The first column shows coefficient ratio estimates based on OLS, using the coefficient for log-income as denominator. The other columns show differences between these estimates and the estimates obtained from ordered regression models (ordered regression estimate minus OLS estimate).

Table B.4: Differences in estimates from ordered regressions and OLS of relative impacts of different variables on life satisfaction, ESS

	$\hat{\beta}_k/\hat{\beta}_{inc}$	$\Delta\hat{\beta}_k/\hat{\beta}_{inc}$			
	OLS	logit	probit	loglog	cloglog
Unemployed	−3.43	0.21	0.36*	1.24*	−1.41
Other employment	0.12	0.35	0.29	0.02	1.11
Good health	4.36	0.31*	0.29*	−1.05*	3.51*
Married or cohab.	2.48	0.43*	0.37*	−0.51*	2.60*
Female	0.39	0.06	0.07*	−0.09*	0.41*
Age 25–44	−1.36	−0.18	−0.19*	0.40*	−1.86*
Age 45–64	−1.01	−0.15	−0.09	0.30*	−1.09*
Age 65+	1.08	0.28*	0.33*	−0.30*	1.91*

* indicates that the difference w.r.t. the OLS estimate is significant on at least a 5% significance level. The first column shows coefficient ratio estimates based on OLS, using the coefficient for log-income as denominator. The other columns show differences between these estimates and the estimates obtained from ordered regression models (ordered regression estimate minus OLS estimate).

Table B.5: Asympotic bias of ratio estimates expressed as % of true parameter, with data based on the GSS and the probit model

	OLS	logit	probit	loglog	cloglog
Unemployed	−1.20	−1.46	0.07	1.23	−0.63
Good health	−0.70	−1.07	−0.34	0.00	−0.98
Married or cohab.	0.19	−0.21	0.07	−0.48	−1.39
Female	−1.21	−1.62	−0.33	0.50	−0.01
Age 25–44	−0.65	−0.56	−0.21	1.11	0.62
Age 45–64	−1.19	−1.47	−0.50	1.04	0.26

Results from simulated data based on the probit with $n = 50,000$.

Table B.6: Asympotic bias of ratio estimates expressed as % of true parameter, with data based on the PSID and the probit model

	OLS	logit	probit	loglog	cloglog
Unemployed	−3.83	−0.59	0.13	5.68	−0.97
Good health	−0.09	−0.97	−0.29	−0.98	−2.07
Married or cohab.	−3.18	−1.04	−0.17	3.73	−4.17
Female	2.84	0.54	0.60	−2.05	−1.36
Age 25–44	−1.02	0.11	−0.04	1.68	0.04
Age 45–64	0.21	0.97	0.39	1.80	0.31

Results from simulated data based on the probit with $n = 50,000$.

Table B.7: Asympotic bias of ratio estimates expressed as % of true parameter, with data based on the ESS and the probit model

	OLS	logit	probit	loglog	cloglog
Unemployed	−20.59	−2.39	0.00	8.55	−6.10
Good health	4.71	0.35	0.08	−2.23	0.32
Married or cohab.	0.84	0.19	0.25	−0.41	0.04
Female	−3.29	−1.67	−1.18	−0.22	−1.84
Age 25–44	−4.82	−0.60	0.04	2.02	−0.55
Age 45–64	−7.95	−0.66	−0.13	3.72	−0.72

Results from simulated data based on the probit with $n = 50,000$.

Swedish Institute for Social Research

Dissertation Series

If not otherwise stated, the dissertation has been submitted at Stockholm University. The dissertations in the Swedish language contain an English summary.

1. Ante Farm (1986): *A Model of the Price Mechanism*
2. Michael Tåhlin (1987): *Arbetets värde och kostnader. En studie av lönearbetets konsekvenser för individen* [The Value and Costs of Work. A Study of the Consequences of Wage Labour for the Individual]
3. Lucienne Portocarero (1987): *Social Mobility in Industrial Societies. Women in France and Sweden*
4. Lennart Erixon (1987): *Profitability in Swedish Manufacturing. Trends and Explanations*
5. Peter Hedström (1988): *Structures of Inequality. A Study of Stratification within Work Organizations* (Harvard University)
6. Jan O. Jonsson (1988): *Utbildning, social reproduktion och social skiktning* [Education, Social Reproduction, and Social Stratification]
7. Jaime Behar (1989): *Trade and Employment in Mexico*
8. Carl le Grand (1989): *Interna arbetsmarknader, ekonomisk segmentering och social skiktning. En studie av arbetslivsstrukturer, anställningsstabilitet och löneskillnader* [Internal Labour Markets, Economic Segmentation and Social Stratification]

9. Ryszard Szulkin (1989): *Privat eller offentligt? Organisationsstruktur och arbetsförhållanden under olika ägandeformer* [Private or Public? Organizational Structure and Working Conditions under Different Forms of Ownership]

10. Sten-Åke Stenberg (1990): *Vräkt ur folkhemmet. En studie av vräkningarna i Sverige under 1900-talet* [Evictions in the Welfare State]

11. Olle Lundberg (1990): *Den ojämlika ohälsan. Om klass- och könsskillnader i sjuklighet* [Inequality in Ill Health. On Class and Sex Differences in Illness]

12. Susanne Oxenstierna (1990): *From Labour Shortage to Unemployment? The Soviet Labour Market in the 1980s*

13. Sven E. Olsson (1990): *Social Policy and Welfare State in Sweden*

14. Joakim Palme (1990): *Pension Rights in Welfare Capitalism. The Development of Old-Age Pensions in 18 OECD Countries 1930 to 1985*

15. Mahmood Arai (1990): *Essays on Non-Competitive Wage Differentials*

16. Johan Fritzell (1991): *Icke av marknaden allena. Inkomstfördelningen i Sverige* [Not Solely by the Market. Income Distribution in Sweden]

17. Eugenia Kazamaki (1991): *Firm Search, Sectoral Shifts, and Unemployment*

18. Lena Schröder (1991): *Springpojkar och språngbrädor. Om orsaker till och åtgärder mot ungdomars arbetslöshet* [Dead-end Jobs and Upgrading Plans. On Reasons Behind and Programmes Against Youth Unemployment] (Uppsala universitet)

19. Olli Kangas (1991): *The Politics of Social Rights. Studies on the Dimensions of Sickness Insurance in OECD Countries* (Helsingfors universitet)

20. Göran Sidebäck (1992): *Kampen om barnets själ. Barn- och ungdomsorganisationer för fostran och normbildning 1850-1980* [The Struggle for the Soul of the Child. Child- and Youth Organizations for Rearing and Normbuilding 1850-1980]

21. Christina Axelsson (1992): *Hemmafrun som försvann. Övergången till lönearbete bland gifta kvinnor i Sverige 1968-1981* [The Housewife that Disappeared. Married Women's Transition to Paid Employment in Sweden 1968-1981]

22. Hjördis D'Agostino (1992): *Why Do Workers Join Unions? A Comparison of Sweden and OECD Countries*

23. Maria Nyström Peck (1994): *Childhood Class, Body Height and Adult Health. Studies on the Relationship between Childhood Social Class, Adult Height and Illness and Mortality in Adulthood*

24. Tomas Korpi (1994): *Escaping Unemployment. Studies in the Individual Consequences of Unemployment and Labour Market Policy*

25. Irene Wennemo (1994): *Sharing the Costs of Children. Studies on the Development of Family Support in the OECD Countries*

26. Viveca Östberg (1996): *Social Structure and Children' s Life Chances. An Analysis of Child Mortality in Sweden*

27. Stig Blomskog (1997): *Essays on the Functioning of the Swedish Labour Market*

28. Katarina Richardson (1997): *Essays on Family and Labor Economics*

29. Håkan Regnér (1997): *Training at the Job and Training for a New Job. Two Swedish Studies*

30. Kristiina Manderbacka (1998): *Questions on Survey Questions on Health* (Helsingfors universitet)

31. Helen Dryler (1998): *Educational Choice in Sweden. Studies on the Importance of Gender and Social Contexts*

32. Michael Gähler (1998): *Life After Divorce. Economic, Social and Psychological Well-being Among Swedish Adults and Children Following Family Dissolution*

33. Lena Granqvist (1998): *A Study of Fringe Benefits. Analysis Based on Finnish Micro Data* (Åbo Akademi)

34. Olof Bäckman (1998): *Longitudinal Studies on Sickness Absence in Sweden*

35. Anna Thoursie (1998): *Studies on Unemployment Duration and on the Gender Wage Gap*

36. Christian Kjellström (1999): *Essays on Investment in Human Capital*

37. Gunnar Isacsson (1999): *Essays on the Twins Approach in Empirical Labor Economics*

38. Eero Carroll (1999): *Emergence and Structuring of Social Insurance Institutions. Comparative Studies on Social Policy and Unemployment Insurance*

39. Peter Skogman Thoursie (1999): *Disability and Work in Sweden*

40. Helena Persson (1999): *Essays on Labour Demand and Career Mobility*

41. Magnus Nermo (1999): *Structured by Gender. Patterns of Sex Segregation in the Swedish Labour Market. Historical and Cross-national Comparisons*

42. Ola Sjöberg (2000): *Duties in the Welfare State. Working and Paying for Social Rights*

43. Mikael Lindahl (2000): *Studies of Causal Effects in Empirical Labor Economics*

44. Ingemar Kåreholt (2000): *Social Class and Mortality Risk*

45. Ingalill Montanari (2000): *Social Citizenship and Work in Welfare States. Comparative Studies on Convergence and on Gender*

46. Ann-Zofie E. Duvander (2000): *Couples in Sweden. Studies on Family and Work*

47. Mia Hultin (2001): *Consider Her Adversity. Four Essays on Gender Inequality in the Labor Market*

48. Carin Lennartsson (2001): *Still in Touch. Family Contact, Activities and Health Among Elderly in Sweden*

49. Per Båvner (2001): *Half Full or Half Empty? Part-time Work and Well-being Among Swedish Women*

50. Per Gillström (2001): *Fair Care. Four Essays on the Allocation and Utilization of Health Care*

51. Magnus Bygren (2001): *Career Outcomes in the Swedish Labor Market. Three Contextual Studies*

52. Björn Öckert (2001): *Effects of Higher Education and the Role of Admission Selection*

53. Susanne Alm (2001): *The Resurgence of Mass Unemployment. Studies in Social Consequences of Joblessness in Sweden in the 1990s*

54. Ann-Christin Jans (2002): *Notifications and Job Losses on the Swedish Labour Market*

55. Sara Ström (2002): *A Shared Experience. Studies on Families and Unemployment*

56. Roger Vilhelmsson (2002): *Wages and Unemployment of Immigrants and Natives in Sweden*

57. Charlotte Samuelsson (2002): *Att göra eller inte göra... Arbetslösas fritidsdeltagande, sökaktivitet, anställningsmöjligheter och tidsstruktur* [To Do or Not to Do... Unemployed's Leisure Participation, Search Activity, Job Opportunities and Time Structure]

58. Tommy Ferrarini (2003): *Parental Leave Institutions in Eighteen Post-war Welfare States*

59. Jenny Säve-Söderbergh (2003): *Essays on Gender Differences in Economic Decision-making*

60. Kenneth Nelson (2003): *Fighting Poverty. Comparative Studies on Social Insurance, Means-tested Benefits and Income Redistribution*

61. Marie Evertsson (2004): *Facets of Gender. Analyses of the Family and the Labour Market*

62. Gabriella Sjögren (2004): *Essays on Personnel Economics and Gender Issues*

63. Kent Friberg (2004): *Essays on Wage and Price Formation in Sweden*

64. Ingrid Esser (2005): *Why Work? Comparative Studies on Welfare Regimes and Individuals' Work Orientations*

65. Åsa Olli Segendorf (2005): *Job Search Strategies and Wage Effects for Immigrants*

66. Pathric Hägglund (2006): *Natural and Classical Experiments in Swedish Labour Market Policy*

67. Lars Brännström (2006): *Phantom of the Neighbourhood. Longitudinal Studies on Area-based Conditions and Individual Outcomes*

68. Helena Holmlund (2006): *Education and the Family. Essays in Empirical Labour Economics*

69. Pernilla Andersson (2006): *Four Essays on Self-Employment*

70. Johanna Kumlin (2007): *Disentangling Sex Segregation. Studies on the Roots and Routes of Labour Market Sex Segregation*

71. Anders Böhlmark (2007): *School Reform, Educational Achievement and Lifetime Income. Essays in Empirical Labor Economics*

72. Krister Sund (2007): *Teachers, Family and Friends. Essays in Economics of Education*

73. Christer Gerdes (2008): *Studying the Interplay of Immigration and Welfare States*

74. Katarina Boye (2008): *Happy Hour? Studies on Well-Being and Time Spent on Paid and Unpaid Work*

75. Lena Lindahl (2008): *Family Background and Individual Achievement. Essays in Empirical Labour Economics*

76. Richard Baltander (2009): *Education, Labour Market and Incomes for the Deaf/Hearing Impaired and the Blind/Visually Impaired*

77. Sara Brolin Låftman (2010): *Children's Living Conditions. Studies on Health, Family and School*

78. Charlotta Magnusson (2010): *Mind the Gap. Essays on Explanations of Gender Wage Inequality*

79. Lalaina Hirvonen (2010): *Essays in Empirical Labour Economics. Family Background, Gender and Earnings*

80. Martin Hällsten (2010): *Essays on Social Reproduction and Lifelong Learning*

81. Marta Lachowska (2010): *Essays in Labor Economics and Consumer Behavior*

82. Marieke Bos (2010): *Essays in Household Finance*

83. Patrik Gränsmark (2010): *Essays on Economic Behavior, Gender and Strategic Learning*

84. Elin Olsson (2011): *Social Relations in Youth. Determinants and Consequences of Relations to Parents, Teachers, and Peers*

85. Karin Halldén (2011): *What's Sex Got to Do with It? Women and Men in European Labour Markets*

86. Frida Rudolphi (2011): *Inequality in Educational Outcomes. How Aspirations, Performance, and Choice Shape School Careers in Sweden*

87. Susan Niknami (2012): *Essays on Inequality and Social Policy. Education, Crime and Health*

88. Martin Nybom (2013): *Essays on Educational Choice and Intergenerational Mobility*

89. Jenny Torssander (2013): *Equality in Death? How the Social Positions of Individuals and Families are Linked to Mortality*

90. Yerko Rojas (2014): *Childhood Social Exclusion and Suicidal Behavior in Adolescence and Young Adulthood*

91. Cecilia von Otter (2014): *Educational and Occupational Careers in a Swedish Cohort*

92. Per Olof Robling (2015): *Essays on the Origins of Human Capital, Crime and Income Inequality*

93. Hrvoje Kap (2015): *Comparative Studies of Vocational Education and Training*

94. André Richter (2016): *Essays on the Intergenerational Transmission of Disadvantage. The Role of Prenatal Health and Fertility*

95. Per Engzell (2016): *Intergenerational Persistence and Ethnic Disparities in Education*

96. Niklas Kaunitz (2017): *Workers, Firms and Welfare. Four Essays in Economics*

97. Martin Berlin (2017): *Essays on the Determinants and Measurement of Subjective Well-Being*