VOLUME 9 | ISSUE 4 | OCTOBER 2018
ISSN 1757-9597

Longitudinal and Life Course Studies: International Journal

Published by
SLLS Society for Longitudinal and Life Course Studies

Inside this issue

- Cohort anlysis of subjective wellbeing and ageing
- Health effects of work and family transitions
- Data control: anonymisation software using ALSPAC cohort data
- Sex of older siblings and psychological functioning
- Study Profile: Generations and Gender Survey

LLCS EDITORIAL BOARD

IMPACT

The journal has been accepted (2018) into the **Social Sciences Citation Index**®, maintained by Clarivate Analytics, along with their Journal Citation Reports/Social Sciences Edition, and Current Contents®/Social and Behavioral Sciences. It had been in the Emerging Sources Citation Index since 2016, and will be eligible for rating with an Impact Factor in 2019. It is also indexed by Google and by the SCOPUS citation database since 2014.

SUBSCRIPTIONS

For immediate access to current issue and preceding 2016 and 2017 issues (Volumes 7 and 8): ***Annual*** Individual rate: £20 (discounted rate £55 for 3 years**)**; ***Annual*** Library rate: £200 to register any number of library readers; ***Annual*** **Society for Longitudinal and Life Course Studies (SLLS) Membership: *Individual* £65, Student £26, Corporate £300.** SLLS members have automatic free access to the journal among many other benefits. For further information about SLLS and details of how to become a member, go to https://www.slls.org.uk/membership All issues are accessible free of charge 12 months after publication.

Back issues

Printed back issues of the journal (to end of Volume 7) are available at a price of £10 (plus postage and packaging). If you would like to purchase a full set or individual copies, visit the SLLS Bookshop at https://www.slls.org.uk/llcsjournal Depending on distance, you will receive the order within two weeks of submitting your order. **Please note:** the reselling of personal registration subscriptions and hard copies of the journal is prohibited.

LLCS Journal can be accessed online at www.llcsjournal.org

Published by the Society for Longitudinal and Life Course Studies
info@slls.org.uk www.slls.org.uk

SLLS Society for Longitudinal and Life Course Studies

SAVE THE DATE!

SLLS International Conference

hosted by the

University of Potsdam, Germany

25th – 27th September 2019

(with post-conference workshops planned for the 24th September)

University of Potsdam

The Call for Papers will go out in October 2018.

The conference fee will be discounted for SLLS members so consider joining if you have not already done so! Visit the membership page of the Society website for full details and to join www.slls.org.uk/membership

More information will be posted on the SLLS conference website as it becomes available www.sllsconference.com

www.slls.org.uk

EDITORIAL: Paths to adulthood and advances in anonymisation

Heather Joshi

This last issue of Volume 9 has cross-cutting contributions from a number of countries in Europe – including (still) the UK – and North America, and from sociology, demography, epidemiology and statistics. The empirical research is mainly confined to people in early to mid adulthood from age 18 to 40. Readers interested in childhood or later life may nevertheless find the articles of relevance. The paper on data privacy addresses a generally important question. This introduction tries to bring out some common themes.

The opening paper uses longitudinal data to address an important issue in the study of the lifecourse. In *A cohort analysis of subjective wellbeing and ageing: heading towards a midlife crisis?*, Steffen Otterbach, Alfonso Sousa-Poza and Valerie Møller set out an account of the ongoing debate about the course of human happiness as individuals pass through the lifecycle. Some argue that the profile is U-shaped with a 'natural' dip in wellbeing at some point in the middle. This not only conforms to the notion of a 'mid-life crisis' but is also claimed to be observed in great apes. Much existing evidence is based on cross-sections, which may also reflect differences between cohorts rather than the effect of the passage of time on an individual. Otterbach and colleagues contribute longitudinal evidence on individuals followed annually for eight years from 2008 to 2015 in Germany. They are from a sequence of three near-overlapping cohorts each born during three years at the start of the 1970s, 1980s and 1990s respectively. They comprise the German Family Panel, otherwise known as pairfam (yes, no capital letters!). The cohorts provide eight observations running forwards across ages roughly 16–23, 26–33 and 36–43. They report on overall life satisfaction and satisfaction with various aspects of life at each year. There is no evidence of life beyond 44 in the dataset so far, but it does provide substantial evidence for the trajectories of the average levels of satisfaction, among these cohorts, over the years in question. These trajectories are presented in three versions: as raw averages, adjusted for personal and macro-economic circumstances, and as estimates abstracting from unobserved personal attributes via fixed effects regressions. All versions are reasonably close. So are the trajectories for males and females. If the satisfaction variables are strung out from 16 to 43, it could be argued that at the evidence is generally consistent with a downward slope on the left-hand side of a life-time U shape. This needs to be qualified by the caveat that the different types of satisfaction show very differently shaped pathways towards the hypothetical mid-life milestone. Most of the decline in overall satisfaction comes between teens and twenties, with only a gentle downward slope within (and between) the cohorts followed from mid-twenties and mid-thirties. There is a similar general pattern for satisfaction with leisure activities and social life. Satisfaction with family life does drop off for all of the age groups over the years of follow-up, providing the strongest evidence for a 'nadir' around 40. Satisfaction from 'school, work or career' shows an opposite trend for the average German person in all three cohorts. This suggests that this field of research in other countries, and indeed in the attempts of official statistical agencies to chart wellbeing, should look into the various facets of life satisfaction.

This superficial account of the findings ignores the discontinuity of the profiles where the oldest year of one cohort jumps to the youngest year of the next. Cohort as well as age differences need to be understood. As the authors discuss, while individuals grow older, time inevitably proceeds. The concept of the life course admits the inextricability of the process of individuals growing up or growing old from the march of historical time. This paper brings a lot of thought, as well as evidence, to the question posed in its title, but it is only the beginning of an answer.

Our second paper, by Juli Simon Thomas, *Health effects of work and family transitions* also looks at mid-life. The focus is on self-reported health and depressive symptoms, at age 40, as they may be predicted by earlier life events. Events under scrutiny are those involving the gain or loss of a job or of a partner, and their timing and coincidence. These mid-life outcomes and their predictors are measured in the American cohort born in 1957–1964, the NLSY79. The outcomes are similar to, but distinct from, the subjective wellbeing studied by Otterbach and colleagues. The life events, especially when divorce and job loss coincide, might be experienced as crises, but they may take place at any point from 18 to 40, not necessarily at 'mid-life'. While life events are adjusted-out of two versions of the German age-trajectories, here their importance as triggers of poor health at 40 is brought to the fore. In this approach, differences between men and women are more evident than in the life satisfaction results from Germany. Job loss and divorce have stronger negative associations with health at 40 than the positive associations with marriage and employment entry. Marriage has a stronger 'protective effect' against depression for women than men. Although the more recent events show greater impact, events before age 26 show persistent traces. The main message is that the effects of transitions are 'riddled with intersectionalities'; troubles tend to come together. This complements other work by this author on transitions in life domains in USA and their outcomes. Her study of families in the Panel Study of Income Dynamics, *Dimensions of family disruption: coincidence and impacts on children's attainments* (published in *Longitudinal and Life Course Studies*, 2018, Vol 9, no 2, pp.157–187), also includes residential mobility alongside partnership and employment transitions as predicting high school and college attainments of the adults' offspring.

The third article in this number addresses a completely different but also very important issue in research on longitudinal data, the preservation of informant privacy. Although this feature of data management is usually taken for granted, breaches of confidentiality have the potential for a crisis, not only for the dataset in which they occur, but in damaging public confidence in the basis for giving information to surveys, longitudinal surveys in particular. Demetris Avraam, Andy Boyd, Harvey Goldstein and Paul Burton summarise their contribution in their title: *A software package for the application of probabilistic anonymisation to sensitive individual-level data: a proof of principle with an example from the ALSPAC birth cohort study*. Recognising the need to balance disclosure control with retaining data utility, this team have combined expertise in data management, epidemiology and statistics to develop a tool for data custodians of longitudinal studies. It builds on an existing practice of perturbing data to make it safe to release, but gives analysts enough information to allow for the artificial errors introduced into the file anonymised for secondary analysis. They show worked examples on a dataset about asthma in this proof of principle exercise. The fruition of this project should help longitudinal studies 'to maintain participant trust and to share data securely and effectively while meeting ever more stringent data protection requirements'.

The notion of resilience to stress appears in the first article as one theory for why the life satisfaction may recover, if it does, after a mid-life nadir. It takes centre stage in the research note on national register data from Sweden by Scott Montgomery and colleagues, *Sex of older siblings and stress resilience*. Their evidence comes from a rigorous assessment administered to young men in Sweden facing compulsory military service as to whether they had the capability to cope with the stress of combat. This is by far better quality data on psychological resilience than can normally be collected in multi-purpose surveys. It is thus of interest even though it does not extend much beyond age 20 or to young women, or beyond Sweden. The question addressed in this research note is how the assessed ability to cope under stress is related to antecedent family background, in particular the presence of siblings and their sex, for which there is also evidence in linked registers. The authors add, to the (not unmixed) literature on the advantage of being first-born, the finding that older brothers present more of a challenge than older sisters. To put it much more crudely than the authors, 'Who's afraid of a big, bad brother?' On another note, I noticed another point in common with Otterbach et al.: in both papers the useful literature reviews include animal studies.

The potential to study relationships among life events, and between them and wellbeing outcomes, is one of many applications offered by the *Gender and Generations Survey* featured in this issue's study profile, by Anne Gauthier, Susana Cabaço and Tim Edery from the Netherlands Interdisciplinary Demographic Institute. It is part of the Gender and Generations Programme funded by the European Union. This valuable component of the European longitudinal data infrastructure runs the GGS surveys in 21 countries (in east and west Europe) as well as hosting harmonised histories on childbearing and marital events from UK, USA, and Uruguay. The GGS surveys start in these countries and at dates mostly in the first decade of the 21st century (varying by country) and following individuals, initially aged 18–79, at three-year intervals. They hence represent birth cohorts from the last three-quarters of the 20th century. The study profile concentrates on those countries where there are two waves of data already available to external researchers. A new round of surveys is expected to begin in 2020. The focus of data collection, and analysis to date, has been on family dynamics, ranging across the life course from leaving home, through childbearing to care of the elderly. It is a resource of growing potential for internationally comparative research and information in our field. Readers of this journal who are not yet familiar with it are recommended to take a look, and to consider making use of it.

Longitudinal and Life Course Studies 2018 Volume 9 Issue 4 Pp 382–411 ISSN 1757-9597

A cohort analysis of subjective wellbeing and ageing: heading towards a midlife crisis?

Steffen Otterbach Institute for Health Care & Public Management, University of Hohenheim, Germany
steffen.otterbach@uni-hohenheim.de
Alfonso Sousa-Poza Institute for Health Care & Public Management, University of Hohenheim, Germany
Valerie Møller Quality of Life Studies, Institute of Social and Economic Research (ISER), Rhodes University, South Africa

(Received March 2018 Revised July 2018) http://dx.doi.org/10.14301/llcs.v9i4.509

Abstract

Using eight waves from the German Panel Analysis of Intimate Relationships and Family Dynamics (pairfam), we analyse how different domains of subjective wellbeing evolve within seven years (2008–2015) in three different cohorts born 10 years apart (1971/73, 1981/83, and 1991/93). This study contributes to the ongoing debate about subjective wellbeing following a U-shaped pattern over the life course. In four domains our results show the first half of such a U-shaped pattern: on average, general life satisfaction – as well as satisfaction with leisure time, social contacts and friends, and family – declines substantially between the ages of 15 and 44, with the most significant decrease taking place at a young age (early 20s). Nevertheless, trajectories among the three cohorts differ markedly, indicating that, ceteris paribus, responses on subjective wellbeing differ greatly between cohorts born just a decade apart. The results further indicate that the two older cohorts assess family life and social contacts more favourably than the youngest cohort.

Keywords

Subjective wellbeing domains; life satisfaction; ageing, longitudinal data, pairfam survey; Germany

Introduction

Although a large body of psychological literature addresses the midlife crisis (see Brim, 1992; Wethington, 2000), its existence is frequently questioned (e.g. Chiriboga, 1997; McCrae & Costa, 1990). Nevertheless, much popular discourse acknowledges a period of unhappiness, stress, personality changes and difficulties encountered around the age of 40. Wethington (2000), for example, provides evidence that over a quarter of all Americans report having experienced a crisis at midlife.[1] Blanchflower and Oswald (2008) also show for a large sample from the UK Labour Force Survey that the incidence of depression and anxiety follows an inverse U-shape and peaks at around the age of 46. Much research in several disciplines on the evolution of subjective wellbeing (SWB) across the lifespan also documents a U-shaped relation between SWB and age, with the minimum generally encountered around middle age (Bauer, Cords, Sellung, & Sousa-Poza, 2015; Blanchflower & Oswald 2008; Lang, Llewellyn, Hubbard, Langa, & Melzer, 2011; López Ulloa, Møller, & Sousa-Poza, 2013).

Studies on the U-shaped relation between SWB and age tend to use either cross-sectional data or panel data from existing surveys. In their seminal paper, Blanchflower and Oswald (2008) analysed a cross-sectional sample of over 500,000 individuals in the United States and Europe. In the United States, depending on the specification, males reach their minimum life satisfaction at between 36 and 53 years of age, whereas women reached a minimum at 39. In Europe, wellbeing reached a minimum at around 45. Much of the related literature relied on data from long-running panels, such as the British Household Panel (Cheng, Powdthavee, & Oswald, 2017; Clark, 2007; Clark & Oswald, 1994; McAdams, Lucas, & Donnellan, 2012), the German Socio-Economic Panel (Cheng et al., 2017; Frijters & Beatton, 2012; Gwozdz & Sousa-Poza, 2010; Kassenboehmer & Haisken-DeNew, 2012; Van Landeghem, 2008, 2012), the U.S. General Social Survey (Easterlin, 2006; Easterlin & Sawangfa, 2007) or the Panel Survey of Household Income Labour Dynamics in Australia (Cheng et al., 2017; Frijters & Beatton, 2012). Although much of the evidence points to a U-shape, conflicting evidence exists. Depending on the data used, the definition of wellbeing, estimation technique, and choice of covariates, several different forms can be observed. As pointed out by López Ulloa et al. (2013, p. 240), “it is difficult to say with certainty whether the relationship between age and well-being across the lifespan is linear or convex”.

Much of this controversy can be attributed to the fact that, ideally, the analysis of SWB across the lifespan should be conducted using long-running panels that follow representative individuals over the entire lifetime (Frijters & Beatton, 2012). The main advantage of such data is the ability to *directly* control for ‘cohort effects,’ the potential differences between the SWB of individuals born at a certain point in time under particular circumstances and those born at different times (Schilling, 2005). Unfortunately, such data are rarely available, but some data sets do exist (such as the British National Child Development Study and the British Cohort Study). To our knowledge, Galambos, Fang, Krahn, Johnson, and Lachmann (2015) take the longest perspective into account and use happiness data from the Edmonton Transitions Study, which followed over a 25-year period a group of individuals from working- and middle-class neighbourhoods in a large western city in Canada.

The aim of this paper is to analyse ageing and subjective wellbeing using cohort data that encompass all ages between 15 and 43. Specifically, we analyse how different SWB domains evolve within seven years in three different cohorts born 10 years apart (1991–1993, 1981–1983 and 1971–1973). Although our three cohorts do not follow individuals throughout their entire life, following them over seven years has the distinct advantage over existing studies that we can analyse large samples of a single cohort over a relatively long timespan.

Our contribution is thus twofold: first, by analysing specific cohorts, we are able not only to take cohort effects directly into account but also to assess how strong such cohort effects may be. Although past research has documented the existence of cohort effects (e.g. Blanchflower & Oswald, 2008; Clark, 2007; Gwozdz & Sousa-Poza, 2010), by actually following different cohorts across time we are able to get a much clearer picture of these cohort effects. Second, by focusing on several life satisfaction domains, we are able to shed light on the trajectories of global life satisfaction across the lifespan. Thus, an analysis on global life satisfaction masks developments in specific domains that could provide an answer to the origins of changes in global life satisfaction. The influence of different domains will most probably not only change across the lifespan, but may also compensate each other (Theuns, Baran, van Vaerenbergh, Hellenbosch, & Tilinouine, 2012; Theuns, Hofmans, & Verresen, 2007). Yet, with a few notable exceptions (Easterlin, 2006; Easterlin & Sawangfa, 2007; McAdams et al., 2012), little research takes a disaggregated approach, i.e. analyse the development of specific domains across time. None to our knowledge analyse domains with longitudinal data and, in particular, with a cohort approach taken in this study.

Conceptual framework

Several theories have been put forward in order to explain how wellbeing progresses through the lifecycle, and also why a midlife crisis may occur (see the literature review in López Ulloa et al., 2013). According to one socioeconomic theory, younger individuals may have higher expectations than their elders, which may not be met, leading to a drop in wellbeing in younger years. This decline continues as long as aspirations are not being met. In a related train of thought, problems can occur at the midlife transition around age 40 when an individual perceives personal growth as stymied or thwarted (Levinson & Levinson, 1996). The gerontology literature has also highlighted this process whereby older individuals learn to adapt to their strengths and weaknesses and thus have more realistic aspirations, which can raise wellbeing as they age. According to Argyle (2001) happiness increases slightly with age, mainly due to a declining goal-achievement gap. This thus offers an explanation as to why wellbeing rises after middle age. Similarly, Charles and Carstensen's (2009) socio-emotional selectivity theory emphasises that, with passing time and shrinking time horizons, individuals experience more life satisfaction as age increases because they spend more time in activities that contribute more directly to their wellbeing.

There is also some evidence that happy people live longer, which could also increase wellbeing in older age. In their meta-analysis, Howell, Kern, and Lyubomirsky (2007) show that probability of living longer increases by 14% for individuals with high wellbeing compared to those with low wellbeing. In a survey of people living in industrial countries, happier people enjoy an increased longevity of between 7.5 and 10 years, a strong effect comparable to smoking or not (Veenhoven, 2008).

An alternative suggestion is that the midlife crisis is a response to the realisation of approaching death (Jaques, 1965), although the increase in life expectancy well beyond what is considered middle age has rendered this explanation somewhat obsolete (Wethington, 2000). However, evolution may also play a role. In their study that analyses the wellbeing of 508 great apes, Weiss, King, Inoue-Murayama, Matsuzawa, and Oswald (2012) show that a midlife crisis also appears to exist among these species. One possible explanation is that evolutionary selection of individuals that have a higher wellbeing at young and old ages may take place, as "these individuals, being satisfied at stages of their life where they have fewer resources to improve their lot, would be less likely to encounter situations that could be harmful to them or their kin." (Weiss et al., 2012, p. 19950).

Media coverage of the midlife crisis may also accentuate this 'crisis', i.e. personal experiences around middle age may be influenced by "social commentators and media pundits, in search of opportunities to market information as products, arouse 'moral insecurities' that evoke a culture of fear. These fears create panic over aging, even when life is going well" (Wethington, 2000, p. 88).

It must also be stressed that some theories from different disciplines *do not* posit a midlife crisis. The most prominent economic theory is the "life cycle hypothesis" (Modigliani & Brumberg, 1954) which, simply stated, assumes that individuals try to smooth consumption across the lifecycle and in doing so try to maintain a constant utility (i.e. wellbeing) level. Taken at face value, one would thus not expect changes in wellbeing across the lifecycle. The assumptions underlying this theory are, however, quite stringent and loosening them gives rise to more differentiated results. However, as pointed out by Blanchflower and Oswald (2008), "textbook economic analysis is not capable […] of producing unambiguous predictions about the pattern of well-being through life" (Blanchflower & Oswald, 2008, p. 1735). There are also psychological theories that primarily stress the stability of wellbeing across time. A prominent theory is the set point theory, which argues that individuals are born with a predisposition to a certain level of happiness, based on genetics and personality (e.g. Brickman, Coates, & Janoff-Bulman, 1978; Clark & Georgellis, 2012). Changes in wellbeing should thereby only be temporary, and always revert back to a baseline level that is determined biologically. Also known as "hedonic adaptation", this is a process whereby "individuals return to baseline levels of happiness following a change in life circumstances" (Lucas, 2007, p. 75). Even as early as 1999, Diener and Lucas (1999, p. 227) argued that "the influence of genetics and personality suggests a limit on the degree to which policy can increase subjective well-being […] Changes in the environment, although important for short-term well-being, lose salience over time through processes of adaptation, and have small effects on long-term subjective well-being".

In conclusion, one can state that there are numerous, yet often contradictory, theories from several disciplines that explain the passage of wellbeing across the lifecycle. As pointed out by Weiss et al. (2012), there is still little convergence of explanations about the origins of the midlife crisis.

Methods and data

The strand of literature on the relation between happiness and age – also referred to as the mysterious U-shaped relation (Frijters & Beatton, 2012) or the age-happiness puzzle (Li, 2016) – is characterised by a broad discussion on appropriate methodology. In general, this discussion reflects the different views on whether the focus of analytical interest should be happiness over the life course *per se* (Baetschmann, 2014; Easterlin, 2006; Glenn, 2009) or an isolated pure age effect net of all other influences and life-course events (Blanchflower & Oswald, 2008). Adherents of the Easterlin tradition point out that such events as leaving school, securing a first job and subsequent job promotions, getting married, having children, getting divorced, being widowed, experiencing a health decrease and even becoming frail at a particular life stage are natural features of the life course. They therefore argue that these immanent life course events should not be controlled away (Hellevik, 2017). Glenn (2009), for example, in his response to Blanchflower and Oswald (2008), argues that the U-shape is merely the result of using inappropriate control variables. Likewise, Kassenboehmer and Haisken-DeNew (2012) emphasise the importance of controlling for unobserved heterogeneity and taking into account time-invariant individual fixed-effects. Using data from the German Socio-Economic Panel (SOEP) Study, these authors conclude that the U-shape becomes flat once fixed-effects are controlled for. Conversely, Frijters and Beatton (2012), in an analysis of three well-known panel data sets (the SOEP, the Household, Income and Labour Dynamics in Australia (HILDA) Survey, and the British Household Panel Survey (BHPS)), show that the U-shape is deepened by the addition of control variables commonly used in life satisfaction analyses.

In the discussion of appropriate control variables, it is generally agreed that controlling for cohort effects is central (Baetschmann, 2014; Blanchflower & Oswald, 2008; Glenn, 2009), reflecting the fact that individuals born at a certain point in time and under particular circumstances may differ in subjective wellbeing from those born at different times. However, the linear dependency of age, cohort and time creates a problem of multi-dimensionality. That is, whereas in a cross-sectional setting, age perfectly corresponds to birth year, in a longitudinal setting, it is a linear combination of cohort and time. As a result, any attempt to construct broader categories of age and cohort to allow for some variation (e.g. Oswald, 2008) creates more or less serious problems of multi-collinearity (Glenn, 2009). Hence, in the age-period-cohort conundrum, simultaneous identification of these three effects is impossible. In fact, Baetschmann (2014) even argues that this isolated pure and under-identified age effect is uninteresting and its interpretation unmeaningful simply because 'it is not possible to become older without proceeding in time' (p. 397).

Methods

In line with this literature, we use three different but related methods applied to each cohort separately. First, we specify a simple Ordinary Least Squares (OLS) model using the age groups as categorical dummy variables with no additional control variables. The predictions from such a model are equal to the unconditional means of SWB over the age groups. As no control variables are included (i.e. the natural features of the life course are not controlled away), the argument that the observed trends in subjective wellbeing are a mere result of (inappropriate) control variables does not hold for this approach. Second, following the strand of literature arguing that consideration of control variables is essential, we next examine whether the observed trends in SWB are confounded by the inclusion of other influences on SWB. Thus, we estimate OLS regressions using commonly accepted control variables, as well as a health measure (see Frijters & Beatton, 2012). We also include Gross Domestic Product (GDP) per capita and unemployment rates as macro-economic control variables on the federal state level. Here, standard errors are adjusted for within-person clustering of observations. Lastly, we estimate fixed-effects regressions, which enable us to hold unobserved heterogeneity constant, again including the same set of time-variant socio-economic control variables and federal-level macro controls. We thus limit our OLS analyses to time-variant control variables in order to focus on how model predictions change conditional on fixed effects when all else is equal. Thereby we acknowledge that controlling for unobserved heterogeneity might be particularly important. Our models can be expressed as follows:

$$SWB_{it} = \beta X_{it} + \gamma Z_{kt} + \phi_t + \alpha_i + \varepsilon_{it}$$

where SWB_{it} is a measure of subjective wellbeing (overall life satisfaction or domain satisfaction), X_{it} is a vector of the time-variant control variables, and Z_{kt} is a vector of the time-variant macroeconomic control variables (GDP per capita and unemployment rates) on the federal state level. Once we control for these latter, ϕ_t corresponds to the age groups of the respective cohorts and captures any remaining time-specific (survey wave) effects, thus identifying any potential non-linear age effects. While other studies used second or higher degree polynomials (e.g. Li, 2016; Wooden & Li, 2014; Frijters & Beatton, 2012) to approximate and thereby smoothly interpolate the relationship between age and subjective wellbeing, we model age in a non-parametric way. Thus, we do not make any assumption about the functional form of the underlying relationship and allow the effect of age on SWB to vary in the most flexible way. Random errors are denoted by ε_{it}, and individual fixed effects by α_i, which in the pooled OLS estimation, is restricted to zero (i.e., excluded from the model). The X_{it}, and Z_{kt} vectors are also restricted to zero in estimations of the mean.

In principle, the ordinal feature of our dependent SWB variables would require a non-linear estimation method such as ordinal logit. However, as in many other studies (e.g. Wooden & Li (2014) or Kassenboehmer & Haisken-DeNew (2012)) we treat our dependent variables as cardinal (i.e. as a discrete ratio scale). For the ease of interpretation and comparability to other studies we follow Ferrer-i-Carbonell and Frijters (2004) who show that it makes only minor differences regarding the results whether ordinality or cardinality are assumed. All estimations are carried out for both men and women, as well as for the three cohorts separately. The results are presented graphically as the unconditional mean and the model predictions over the age groups, respectively. Because individuals grow older as time proceeds, in this specific setting, age and time are non-separable dimensions.

Sample

The analyses are based on release 8.0 data (Brüderl et al., 2017) from the first eight waves of the German Panel Analysis of Intimate Relationships and Family Dynamics (pairfam)[2], a longitudinal nation-wide survey aimed at providing an empirical data base for the study of partnership and family dynamics. Begun in 2008 and collected annually ever since, at baseline, pairfam surveyed about 12,000 randomly selected respondents (anchor persons) among three cohorts born 1971–73 (4,054 individuals), 1981–83 (4,010 individuals) and 1991–93 (4,338 individuals). Although corrected panel attrition rates stabilised around 10% after wave 3, by wave 6 more than half of the original pairfam sample had been lost (Brüderl et al., 2017). From wave two on pairfam is complemented by the Demographic Differences in Life Course Dynamics in Eastern and Western Germany (DemoDiff) panel study, which follows closely the design of pairfam but only samples the cohorts 1971–1973 and 1981–1983. Initiated and funded by the Max Planck institute for Demographic Research, DemoDiff has been fully integrated in pairfam and from wave 5 onwards its respondents are regarded as regular pairfam respondents (Brüderl et al., 2017). The cohort-sequential design of the study with its adjacent segments regarding the three age groups is illustrated in figure A.1. Data were collected by mode of computer-assisted personal interviewing (CAPI) among respondents living in private households in Germany with sufficient language skills to follow the German-speaking interview. As can be seen in graph A.1 in the appendix, the cohorts do not (yet) overlap. However, the cohorts "touch each other", which allows us to assess with reasonable confidence whether or not trajectories for the different cohorts differ.

Measurement of subjective wellbeing

In addition to a wealth of variables describing family and partnership dynamics, pairfam offers rich information on several domains of wellbeing and satisfaction. In particular, at the beginning of the interview, respondents are asked, 'How satisfied are you with the following domains of your life?': (i) school, education, career; (ii) leisure activities, hobbies, interests; (iii) friends, social contacts; and (iv) family. The interview concludes with the question, 'All in all, how satisfied are you with your life at the moment?' All these satisfaction domains are surveyed on an 11-point scale ranging from 0 (very dissatisfied) to 10 (very satisfied). Our choice of domains is thus primarily data driven, and we acknowledge that several important domains (e.g. satisfaction with income or health) are not covered by our analysis.

We use single-item measures for subjective wellbeing. It could be argued that multi-item measures such as the satisfaction with life scale (Diener, Emmons, Larsen, & Griffin, 1985) consisting of multiple questions provide advantageous psychometric properties to cover the multidimensional aspects of subjective wellbeing compared to single-item scales (Jovanovic 2016). However, it has been shown that single item measures for life satisfaction and subjective wellbeing have strong correlations with and good reliability compared to multi-item measures of life satisfaction (Robustellie & Whisman, 2016).

Covariates

The analysis does, however, include a parsimonious set of socio-economic covariates that are widely used as standard control variables, as well as a measure of health (Frijters & Beatton, 2012). The explanatory variables are marital status ('married' or 'not married'), number of children, and self-rated health status within the last four weeks. This latter is measured on a five-point scale ('poor', 'suboptimal', 'satisfactory', 'good', 'excellent') that is then recoded into a binary good/poor health dummy based on the first and last two categories, respectively, with satisfactory health as the reference category. Further control variables include being unemployed ('yes' or 'no'), not being in the labour force ('yes' or 'no'), and the natural logarithm of equivalised net household income, which is adjusted to household structure according to the modified Organization of Economic Co-operation and Development (OECD) scale. We also control for whether another person was present during the interview. Finally, to capture wealth and period effects related to the business cycle, we also include GDP per capita and unemployment rates on the federal state level. As Baetschmann (2014) points out, capturing these effects is particularly important when the observation period is short but encompasses the European economic crises. Summary statistics describing the SWB domains and all covariates are given in table 1.

Results

Figure 1 depicts the results for life satisfaction (with corresponding tables for all figures provided in the appendix and regression results for the full samples provided in a supplementary appendix). Although a cursory glance at the first wave results for each cohort suggests a downward movement in life satisfaction (fixed-effects predictions of 7.78, 7.47 and 7.48 for the 1991/93, 1981/83 and 1971/73 cohorts, respectively), the changes *within* a cohort suggest that a strong decline in life satisfaction takes place only in the youngest cohort. In the other two cohorts, the relation remains quite flat. This drop in the youngest cohort is quite large, about 0.37, 0.32, and 0.14 points within seven years for the unconditional and conditional values, respectively. For the middle cohort, life satisfaction remains quite flat, and the oldest cohort experiences a slight decrease. Life satisfaction thus declines substantially between the ages of 15 and 24 (covered by the young cohort) and then remains relatively flat until the age of 44 (covered by the middle and old cohorts). This is in accordance with the left-hand side of the U-shaped relation between age and life satisfaction. As can be seen by the confidence intervals, most changes in these two older cohorts are not significant. This pattern is similar for men and women but more pronounced for males. When comparing life satisfaction (OLS and fixed-effects results) in the eighth wave of the first cohort with the first wave of the second cohort, we do not observe a major 'jump' in life satisfaction, implying that cohort effects are most probably quite negligible between these two cohorts. This being the case, it appears that the strong decline in life satisfaction in young years levels out at about the age of 24 or 25.

A different pattern emerges, however, for satisfaction with school, education and career (see figure 2), which increases in younger years by about 0.24 and 0.14 points for the fixed-effects predictions and unconditional mean, respectively, but changes less sharply in the two older cohorts (differences insignificant). Nevertheless, we observe a marked cohort effect between the young and middle cohorts, with a large and significant drop in the fixed-effects predictions and unconditional means by 0.41 and 0.36 points, respectively. This pattern is more or less equal for both genders.

Figure 3 shows the results for satisfaction with leisure activities, hobbies and interests, the first of which shows a steep decline in the early years that tends to level off around the late 20s. Although this finding holds true for both men and women, no strong cohort effects are observable in this domain. This pattern is similar to the analysis of both amount and use of leisure time by McAdams et al. (2012) using BHPS data.

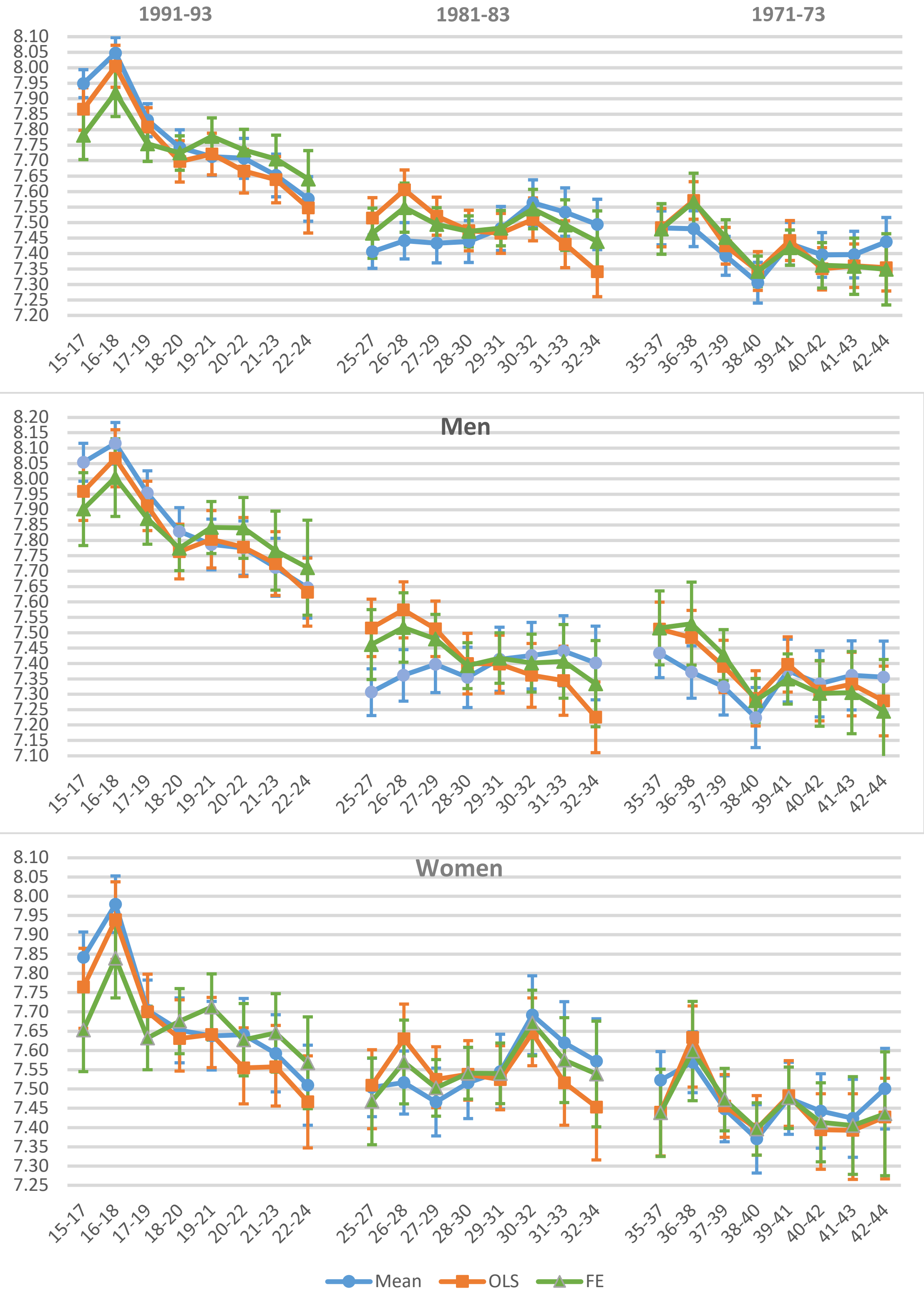
A strong downward trend is also apparent in all cohorts and for both men and women with respect to satisfaction with social contacts and friends (see figure 4). Within all cohorts, this domain drops significantly by between 1.07 (unconditional mean of the young cohort) and 0.45 points (fixed-effects predictions of the oldest cohort) in a pre-midlife decline that is also reported by McAdams et al. (2012) in their analysis of the domain social life. In this domain, assessments are more favourable among the two older cohorts, signalling a slight cohort effect between the middle and old cohort, especially in the female sample.

Table 1. Descriptive statistics: number of observations, means, standard deviations

	Cohorts pooled			Cohort 1991–93			Cohort 1981–83			Cohort 1971–73		
Variable	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD
Life satisfaction	65,236	7.57	1.69	21,842	7.82	1.53	20,812	7.46	1.74	22,582	7.42	1.78
Job satisfaction	64,952	7.23	2.16	21,817	7.43	2.02	20,709	7.12	2.25	22,426	7.15	2.20
Satisfaction with leisure	65,264	7.04	2.15	21,857	7.64	1.94	20,826	6.83	2.15	22,581	6.66	2.23
Satisfaction with social contacts	65,278	7.74	1.96	21,862	8.33	1.67	20,831	7.56	1.99	22,585	7.34	2.07
Satisfaction with family	65,254	8.38	1.81	21,860	8.49	1.67	20,824	8.34	1.86	22,570	8.31	1.88
Unemployed	65,309	0.05	0.22	21,867	0.03	0.16	20,843	0.08	0.27	22,599	0.05	0.22
Not in labour force	65,309	0.30	0.46	21,867	0.59	0.49	20,843	0.19	0.40	22,599	0.11	0.31
Married	65,309	0.33	0.47	21,867	0.01	0.08	20,843	0.32	0.47	22,599	0.63	0.48
Number of chidlren	65,299	0.79	1.11	21,865	0.02	0.17	20,841	0.68	0.97	22,593	1.62	1.19
Self-rated health	65,240	3.74	0.98	21,845	3.87	0.98	20,813	3.74	0.98	22,582	3.62	0.97
Good health	65,240	0.66	0.47	21,845	0.70	0.46	20,813	0.66	0.47	22,582	0.62	0.49
Satisfactory health	65,240	0.22	0.41	21,845	0.19	0.39	20,813	0.21	0.41	22,582	0.24	0.43
Bad health	65,240	0.12	0.33	21,845	0.11	0.31	20,813	0.12	0.33	22,582	0.14	0.34
Net equivalised h'hold income	52,478	1542.38	1080.19	14,683	1320.73	980.75	17,891	1556.56	1003.36	19,904	1693.14	1184.41
Ln net equivalised h'hold income	52,478	7.19	0.58	14,683	7.01	0.64	17,891	7.21	0.56	19,904	7.30	0.53
Unemployment rate	65,300	6.73	2.84	21,867	6.35	2.68	20,838	7.01	2.94	22,595	6.83	2.85
BIP per capita	65,300	32082.08	7042.32	21,867	33107.71	6829.36	20,838	31564.44	7254.06	22,595	31566.90	6937.96
Year 2008	65,309	0.19	0.39	21,867	0.20	0.40	20,843	0.19	0.39	22,599	0.18	0.38
Year 2009	65,309	0.16	0.37	21,867	0.16	0.37	20,843	0.16	0.37	22,599	0.16	0.37
Year 2010	65,309	0.14	0.35	21,867	0.14	0.35	20,843	0.14	0.34	22,599	0.14	0.34
Year 2011	65,309	0.12	0.33	21,867	0.12	0.33	20,843	0.12	0.33	22,599	0.12	0.33
Year 2012	65,309	0.11	0.31	21,867	0.11	0.31	20,843	0.11	0.31	22,599	0.11	0.32
Year 2013	65,309	0.10	0.30	21,867	0.10	0.30	20,843	0.10	0.30	22,599	0.10	0.30
Year 2014	65,309	0.09	0.29	21,867	0.09	0.28	20,843	0.09	0.29	22,599	0.09	0.29
Year 2015	65,309	0.08	0.28	21,867	0.08	0.27	20,843	0.08	0.28	22,599	0.09	0.28

Data: German Panel Analysis of Intimate Relationships and Family Dynamics (pairfam), 2008–2015.

Figure 1. Overall life satisfaction, unconditional mean and model predictions from OLS and FE estimations including 95% CIs

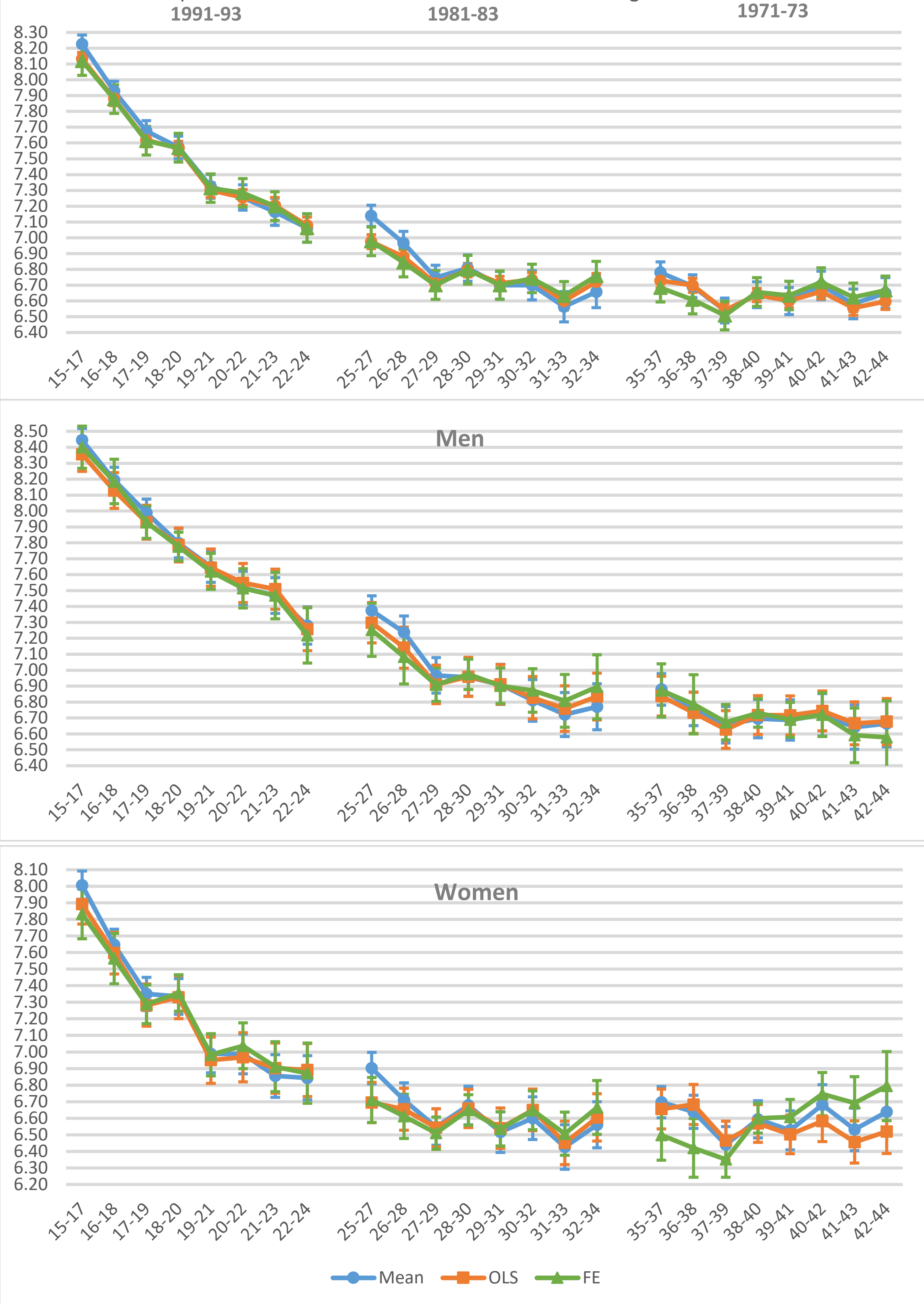


Note: Model predictions include marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalised net household income, whether another person was present during interview, GDP per capita, and unemployment rate as control variables. Full sample (men and women) consists of 21,842, 20,812, 22,582 (unconditional mean) and 14,670, 17,869, 19,888 (OLS and FE predictions) observations for the three birth cohorts, respectively.

Figure 2. Satisfaction with job, education, career, unconditional mean and model predictions from OLS and FE estimations including 95% CIs

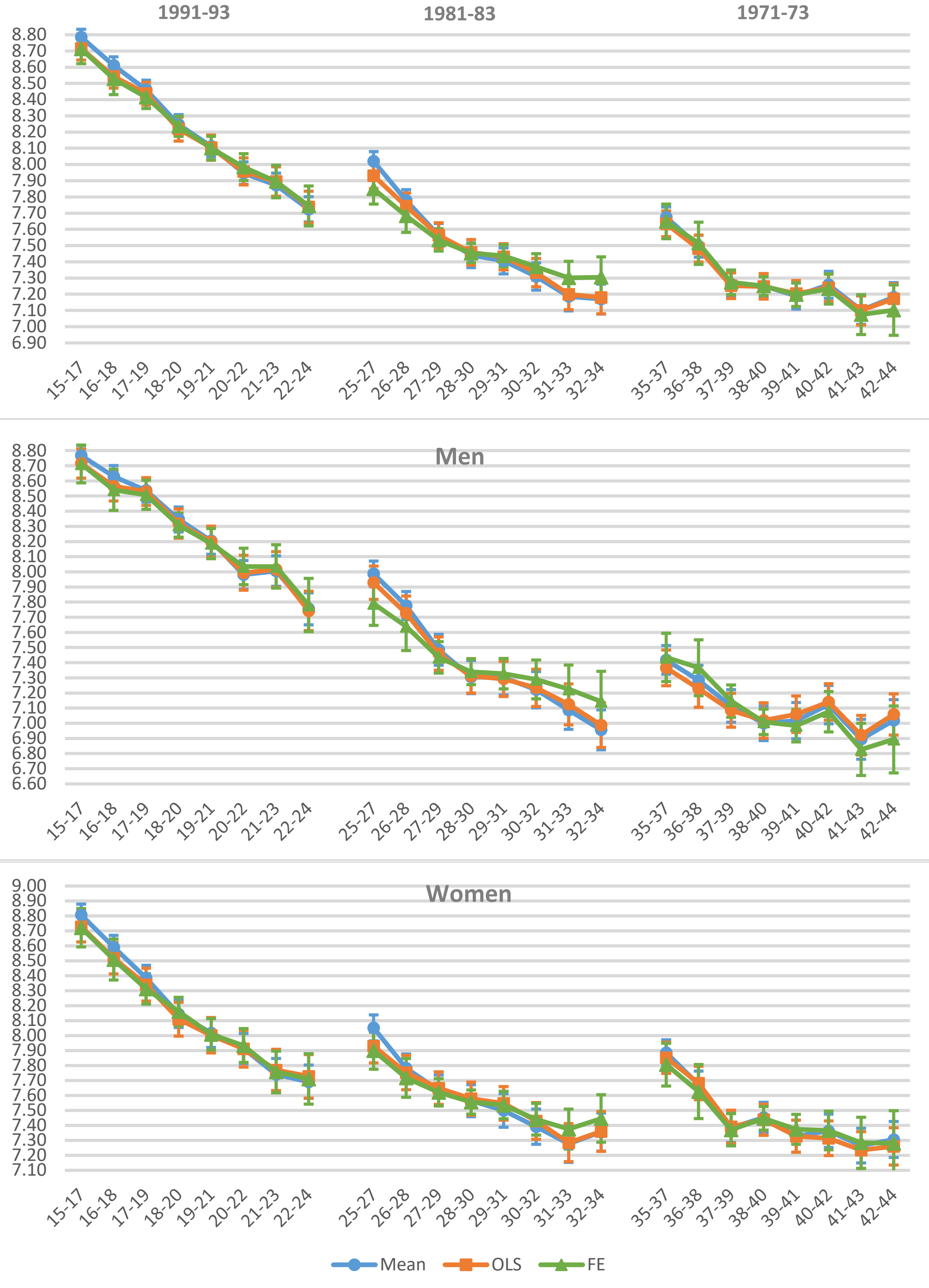
Note: Model predictions include marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalised net household income, whether another person was present during interview, GDP per capita, and unemployment rate as control variables. Full sample (men and women) consists of 21,817, 20,709, 22,426 (unconditional mean) and 14,640, 17,763, 19,758 (OLS and FE predictions) observations for the three birth cohorts, respectively.

Figure 3. Satisfaction with leisure activities, hobbies, interests, unconditional mean and model predictions from OLS and FE estimations including 95% CIs



Note: Model predictions include marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalised net household income, whether another person was present during interview, GDP per capita, and unemployment rate as control variables. Full sample (men and women) consists of 21,857, 20,826, 22,581 (unconditional mean) and 14,671, 17,864, 19,879 (OLS and FE predictions) observations for the three birth cohorts, respectively.

Figure 4. Satisfaction with friends, social contacts, unconditional mean and model predictions from OLS and FE estimations including 95% CIs



Note: Model predictions include marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalized net household income, whether another person was present during interview, GDP per capita, and unemployment rate as control variables. Full sample (men and women) consists of 21,862, 20,831, 22,585 (unconditional mean) and 14,675, 17,863, 19,880 (OLS and FE predictions) observations for the three birth cohorts, respectively.

Figure 5 then graphs the results for satisfaction with family life, which exhibits a marked downward trend within each cohort. The differences in unconditional means and fixed-effects predictions between the first wave of the young cohort and the eighth wave of the old cohort are 0.59 and 0.43 points, respectively, signalling a sharp and significant decrease as midlife approaches. Particularly strong and significant cohort effects are also observable between the middle and old cohort, about 0.35 and 0.44 points for the unconditional means and the fixed-effects predictions, respectively. This general pattern of declining satisfaction with family life is very similar for both men and women; however, the cohort effects differ: the female sample is characterised by a large and significant cohort effect between the young and middle cohorts but the male sample, by a large and significant effect between the middle and old cohort. As with social contacts and friends, assessments of family life are more favourable among the two older cohorts.

Some methodological concern

Potential concerns in our analysis could be (i) non-random response; (ii) attrition; and (iii) panel conditioning. First, as shown in table 1, item non-response could be an issue particularly with respect to household income. However, it is not uncommon that respondents do not want to reveal their income and non-response rates of about 20% are quite common (Sousa-Poza & Henneberger, 2000). Missing information on household income is an even more severe problem in the youngest cohort because respondents at the age of around 15+ years are likely to live with their parents and probably have no information about parents' and household income. As a check of whether missing information biases to our results, we re-estimate our OLS and fixed-effects regressions without the household income variable. The predictions of these regressions do not differ in any notable way from our main specification.

Second, the continuous decline in sample size could raise concerns regarding panel attrition. However, it is important to note that more than 50% of our regression samples are included in all waves. As a robustness check, we use a balanced panel and demonstrate that the unconditional mean and model predictions from a balanced versus an unbalanced panel slightly differ in levels but not in trends. In addition, following Wooden and Li (2014) we include a variable indicating whether a respondent does not participate in wave *t+1* to test and control for potential selectivity bias (see also Verbeek & Nijman (1992)). Re-estimating our regressions including this variable does not notably change our results.

Third, some of the patterns produced in this study could be influenced by panel conditioning effects, i.e. the possibility that the duration a person spends in a panel affects the way the person responds to certain questions. Wooden and Li (2014), using Australian HILDA data, find very little evidence that average life satisfaction is affected by the duration of individual stays in the panel. Likewise, in their analysis of the big five personality traits using the SOEP, Lucas and Donnellan (2011) show that panel conditioning effects are present but small in size. However, Kassenboehmer and Haisken-DeNew (2012) demonstrate that time in the panel effects are more pronounced among German SOEP respondents (see also Baird, Lucas, & Donnellan, 2010). The usual way to analyse this effect is with refreshment samples (e.g. Baird et al, 2010; Lucas & Donnellan, 2011; Wooden & Li, 2014). Unfortunately, the cohort design of pairfam has no refreshments and it is thus not possible to assess panel conditioning effects in a comprehensive way. Past research has shown that, if panel conditioning effects exist, they are small and always negative, i.e. life satisfaction declines with the duration in the panel. Baird et al. (2010), for example, show that, on a 10-point life satisfaction scale, with each additional year in the SOEP survey, life satisfaction scores decline by only about .03 points. In order to get a rough indication of whether our results are being influenced by panel conditioning, we run our OLS estimates and include a variable describing the length of time that a respondent stays in the panel. Although we do not have a refreshment sample, not all respondents participated in all waves, i.e. some respondents interrupted their participation. Of the 52.427 observations in the regression sample on life satisfaction, about 43% missed out at least one wave. The results of these regressions do not change the conclusions of this paper.[3]

Figure 5. Satisfaction with family, unconditional mean and model predictions from OLS and FE estimations including 95% CIs

Note: Model predictions include marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalised net household income, whether another person was present during interview, GDP per capita, and unemployment rate as control variables. Full sample (men and women) consists of 21,860, 20,824, 22,570 (unconditional mean) and 14,674, 17,862, 19,868 (OLS and FE predictions) observations for the three birth cohorts, respectively.

Discussion and conclusions

Using data from three cohorts born 10 years apart and applying three different but related analytical methods, this study provides evidence that SWB decreases from the late teens to about middle age. This decline is very pronounced for certain SWB domains, notably satisfaction with social contacts and friends, and satisfaction with leisure activities, hobbies and interests. There is also a downward trend in general life satisfaction. One of our most important findings is that the largest declines take place in the youngest cohort between the ages of 15 and 23. Although we are unaware of studies that document such changes in life satisfaction among young adults[4], Goldbeck, Schmitz, Besier, Herschbach, and Henrich (2007) do provide evidence of a relatively large decline in life satisfaction among German adolescents aged 11 to 16. As a result, they argue that decreasing life satisfaction has to be considered as a developmental phenomenon. Our results indicate that these developmental changes among adolescents, which are also recorded for other countries (Proctor, Linley, & Maltby, 2009), continue into young adulthood. They thus support the notion that 'emerging adulthood', the transitional developmental stage between late adolescence and adulthood that occurs between ages 18 and 24 (Arnett, 2004) is a stress-arousing and anxiety-provoking period because of the many diverse tasks and expectations it entails.

What could be causing the large drop in life satisfaction in the youngest cohort? In order to shed some light on this question we decomposed the difference in life satisfaction between the eighth and first survey years using a standard Oaxaca-Blinder decomposition (Blinder, 1973; Oaxaca, 1973). We try to explain the drop in life satisfaction by nesting the four domains into the analysis. Our results[5] show that about 99% of the decline in life satisfaction can be explained by these four domains in this young cohort. Of the four domains, satisfaction with friends and satisfaction with leisure are equally influential and account for 85% of the decline in life satisfaction. Satisfaction with family plays a relatively less significant role, and job satisfaction has an attenuating effect, i.e. actually increases life satisfaction. Taken at face value, these results point to the important role that changing social structures and time allocation (e.g. most notably available time for leisure activities) have in shaping the wellbeing of young adults.

The advantage of using cohort data such as ours is the ability to directly control (i.e. observe) cohort effects. Perhaps the most striking result in our study is the size of these cohort effects; that is, the difference in reported SWB of similarly aged individuals in different cohorts. For example, the unconditional mean and the conditional mean of general life satisfaction from the fixed-effects model (i.e. once numerous socio-demographic factors, macro-level variables, and unobserved fixed-effects are controlled) among individuals aged 22–24 in the youngest cohort is 0.17 points higher than that of individuals aged 25–27 in the middle cohort. This discrepancy points to stark inter-cohort differences in SWB response behaviour. What is particularly intriguing is that these cohort effects arise even though the cohorts are only a decade apart. Their identification thus highlights the necessity of adequately controlling for cohort effects during any analysis of multi-cohort SWB data. It is also worth noting that not all domains exhibit an equally strong cohort effect. For instance, differences among cohorts in the satisfaction with family life domain is particularly striking, with older cohorts (*ceteris paribus*) having higher levels of satisfaction. This apparent increased dissatisfaction in the young cohort may have implications for the claim that the rise in non-traditional attitudes towards family, as well as an increased belief in gender egalitarianism, could be negatively affecting satisfaction with family life (Lye & Biblarz, 1993; Taniguchi & Kaufman, 2013).

Admittedly, it could be argued that comparing the eighth wave with the first wave of an adjacent cohort fails to take into account that data collection occurred in different years. However, although this point is valid, our controls for annual GDP and unemployment may well capture much of the time effect. It is also highly unlikely that further time-dependent variables can explain some of the extremely large (and even gender and domain-specific) differences between cohorts.

It should also be noted that most studies based on long-running panel data (e.g. the BHPS or SOEP) can only indirectly or inadequately control for cohort effects, primarily because of the small size of the (narrowly defined) cohorts and their relatively short duration in the panel. For example, in a comparable research setting (i.e. survey years between 2008 and 2015) using the SOEP, the average sample size for individuals born 1971–1973 with no missing values of overall life satisfaction is a mere 823 per survey year, with an average panel duration of 4.8 years.

No doubt as further waves of pairfam are collected, a more precise analysis of individual cohort SWB trajectories will become possible and shed more light on how SWB evolves across time and generations.

Acknowledgements

This paper uses data from the German Family Panel pairfam, coordinated by Josef Brüderl, Karsten Hank, Johannes Huinink, Bernhard Nauck, Franz Neyer and Sabine Walper. pairfam is funded as long-term project by the German Research Foundation (DFG).

References

Argyle, M. (2001). *The psychology of happiness*. London: Routledge.

Arnett, J. (2004). *Emerging adulthood: The winding road from the late teens through the twenties*. New York: Oxford University Press.

Baetschmann, G. (2014). Heterogeneity in the relationship between happiness and age: evidence from the German Socio-Economic panel. *German Economic Rev*iew, *15*(3), 393–410. https://doi.org/10.1111/geer.12015

Baird, B. M., Lucas, R. E., & Donnellan, M. B. (2010). Life Satisfaction Across the Lifespan: Findings from Two Nationally Representative Panel Studies. *Social Indicators Research, 99*(2), 183–203. https://doi.org/10.1007/s11205-010-9584-9

Bauer, J. M., Cords, D., Sellung, R., & Sousa-Poza, A. (2015). Effects of different life events on life satisfaction in the Russian Longitudinal Monitoring Survey. *Economics Letters*, *129*, 91–94. https://doi.org/10.1016/j.econlet.2015.02.016

Blanchflower, D. G. & Oswald, A. J. (2008). Is well-being U-shaped over the life cycle? *Social Science & Medicine*, *66*, 1733–1749. https://doi.org/10.1016/j.socscimed.2008.01.030

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, *8*, 436–455. https://doi.org/10.2307/144855

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology*, *36*(8), 917–927. https://doi.org/10.1037/0022-3514.36.8.917

Brim, O. G. (1992). *Ambition: how we manage success and failure throughout our lives*. New York: Basic Books.

Brüderl, J., Hank, K., Huinink, J., Nauck, B., Neyer, F. J., Walper, S., ... Wilhelm, B. (2017). The German Family Panel (pairfam). GESIS Data Archive, Cologne. ZA5678 Data file Version 8.0.0, https://doi.org/10.4232/pairfam.5678.8.0.0

Charles, S. & Carstensen, L. (2009). Socioemotional selectivity theory. In H. Reis & S. Sprecher (Eds.), *Encyclopedia of human relationships* (pp. 1578–1581). Thousand Oaks: Sage Publications. https://doi.org/10.4135/9781412958479.n519

Cheng, T. C., Powdthavee, N., & Oswald, A. J. (2017). Longitudinal Evidence for a Midlife Nadir in Human Well-being: Results from Four Data Sets. *Economic Journal*, *127*(599), 126–142. https://doi.org/10.1111/ecoj.12256

Chiriboga, D. A. (1997). Crisis, challenge, and stability in the middle years. In M. E. Lachman & J. B. James (Eds.), *Multiple paths of midlife development* (pp. 293–322). Chicago: University of Chicago Press.

Clark, A. (2007). *Born to be mild? Cohort effects don't (fully) explain why well-being is U-shaped in age*. IZA Discussion paper No. 3170.

Clark, A. & Georgellis, Y. (2012). Back to Baseline in Britain: Adaptation in the British Household Panel Survey. *Economica*, *80*, 496–512. https://doi.org/10.1111/ecca.12007

Clark, A. & Oswald, A. (1994). Unhappiness and unemployment. *Economic Journal*, *104*(424), 648–659. https://doi.org/10.2307/2234639

Diener, E., & Lucas, R. E. (1999). Personality and subjective well-being. In D. Kahneman, E. Schwarz, & N. Diener (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 213–229). New York: Russell Sage.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Easterlin, R. A. (2006). Life cycle happiness and its sources: Intersections of psychology, economics, and demography. *Journal of Economic Psychology*, *27*(4), 463–482. https://doi.org/10.1016/j.joep.2006.05.002

Easterlin, R. A. & Sawangfa, O. (2007). *Happiness and domain satisfaction: theory and evidence*. IZA Discussion Paper No. 2584, 2007: 1–35.

Ferrer-i-Carbonell, A. & Frijters, P. (2004). How Important is methodology for the estimates of the determinants of happiness? *Economic Journal*, *114*(497), 641–659. https://doi.org/10.1111/j.1468-0297.2004.00235.x

Frijters, P. & Beatton, T. (2012). The mystery of the U-shaped relationship between happiness an age. *Journal of Economic Behavior & Organization*, *82*, 525–542. https://doi.org/10.1016/j.jebo.2012.03.008

Galambos, N. L., Fang, S., Krahn, H. J., Johnson, M. D., & Lachman, M. E. (2015). Up, not down: The age curve in happiness from early adulthood to midlife in two longitudinal studies. *Developmental Psychology*, *51*(11), 1664–1671. https://doi.org/10.1037/dev0000052

Glenn, N. (2009). Is the apparent U-shape of well-being over the life course a result of inappropriate use of control variables? A commentary on Blanchflower and Oswald. *Social Science & Medicin*e, *69*, 481–485. https://doi.org/10.1016/j.socscimed.2009.05.038

Goldbeck, L., Schmitz, T. G., Besier, T., Herschbach, P., & Henrich, G. (2007). Life satisfaction decreases during adolescence. *Quality of Life Research*, *16*, 969–979. https://doi.org/10.1007/s11136-007-9205-5

Gwozdz, W. and Sousa-Poza, A. (2010). Ageing, health and life satisfaction for the oldest old: an analysis for Germany. *Social Indicators Res*earch, *97*, 397–417. https://doi.org/10.1007/s11205-009-9508-8

Hellevik, O. (2017). The U-shaped age-happiness relationship: real or methodological artifact? *Quality and Quantity*, *51*, 177–197. https://doi.org/10.1007/s11135-015-0300-3

Howell, R. T., Kern, M. L., & Lyubomirsky, S. (2007). Health benefits: Meta-analytically determining the impact of well-being on objective health outcomes. *Health Psychology Review*, *1*, 83–136. https://doi.org/10.1080/17437190701492486

Huinink, J., Brüderl, J. Nauck, B., Walper, Castiglioni, L., & Feldhaus, M. (2011). Panel analysis of intimate relationships and family dynamics (pairfam): Conceptual framework and design. *Zeitschrift für Familienforschung – Journal of Family Resea*rch, *23*, 77–101.

Jaques, E. (1965). Death and the midlife crisis. *International Journal of Psychoanalysis*, *46*, 502–514.

Jovanovic, V. (2016). The validity of the Satisfaction with Life Scale in adolescents and a comparison with single-item life satisfaction measures: a preliminary study. *Quality of Life Research*, *25*, 3173–3180. https://doi.org/10.1007/s11136-016-1331-5

Kassenboehmer, S. C. & Haisken-DeNew, J. P. (2012). Heresy or enlightment? The well-being age U-shape effect is flat. *Economics Letters*, *117*, 235–238. https://doi.org/10.1016/j.econlet.2012.05.013

Lang, I.A., Llewellyn, D.J., Hubbard, R.E., Langa, K.M., & Melzer, D. (2011). Income and the midlife peak in common mental disorder prevalence. *Psychological Medicine*, *41*(7), 1365–1372. https://doi.org/10.1017/S0033291710002060

Levinson, D. J. & Levinson, J. D. (1996). *The seasons of a woman's life*. New York: Knopf.

Li, N. (2016). Multidimensionality of longitudinal data: Unlocking the age-happiness puzzle. *Social Indicators Research*, *128*(1), 305–320. https://doi.org/10.1007/s11205-015-1032-4

López Ulloa, B. F., Møller, V., & Sousa-Poza, A. (2013). How does subjective well-being evolve with age? A literature review. *Journal of Population Ageing, 6*(3), 227–246. https://doi.org/10.1007/s12062-013-9085-0

Lucas, R. E. (2007). Adaptation and the set point model of subjective well-being. *Current Directions in Psychological Science, 16*(2), 75–79. https://doi.org/10.1111/j.1467-8721.2007.00479.x

Lucas, R. E. & Donnellan, B. M. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology*, *101*(4), 847–861. https://doi.org/10.1037/a0024298

Lye, D. N. & Biblarz, T. J. (1993). The effects of attitudes toward family life and gender roles on marital satisfaction. *Journal of Family Iss*ues, *14*(2), 157–188. https://doi.org/10.1177/019251393014002002

McAdams, K. K., Lucas, R. E., & Donnellan, B. M. (2012). The role of domain satisfaction in explaining the paradoxical association between life satisfaction and age. *Social Indicators Re*search, *109*, 295–303. https://doi.org/10.1007/s11205-011-9903-9

McCrae, R. & Costa, P. (1990). *Personality in adulthood*. New York: Guilford. https://doi.org/10.1007/s11205-011-9903-9

Modigliani, F., & Brumberg, R. (1954). Utility analysis and the consumption function: an interpretation of cross-section data. In K. Kurihara (Ed.), *Post-Keynesian Economics*. New Brunswick: Rutgers University Press.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, *14*, 693–709. https://doi.org/10.2307/2525981

Proctor, C. L., Linley, A. & Maltby, J. (2009). Youth life satisfaction: A review of the literature. *Journal of Happiness Studies*, *10*, 583–630. https://doi.org/10.1007/s10902-008-9110-9

Robustelli, B. L. & Whisman, M. A. (2016). Gratitude and Life Satisfaction in the United States and Japan. *Journal of Happiness Studies*, *19*(1), 41–55. https://doi.org/10.1007/s10902-016-9802-5

Schilling, O. K. (2005). Cohort- and age-related decline in elder's life satisfaction: is there really a paradox? *European Journal of Ageing*, *2*(4), 254–263. https://doi.org/10.1007/s10433-005-0016-7

Sousa-Poza, A. & Henneberger, F. (2000). Wage Data Collected by Telephone Interviews: An Empirical Analysis of the Item Nonresponse Problem and its Implications for the Estimation of Wage Functions. *Swiss Journal of Economics and Statistics*, *136*(1), 79–98.

Taniguchi, H. & Kaufman, G. (2013). Gender role attitudes, troubles talk, and marital satisfaction in Japan. *Journal of Social and Personal Relationships*, *31*, 7975–7994.

Theuns, P., Baran, B., van Vaerenbergh, R., Hellenbosch, G., & Tiliouine, H. (2012). A cross-cultural experimental approach to the contribution of health, religion and personal relations to subjective satisfaction with life as a whole. *Psicológica*, *33*, 591–608.

Theuns, P., Hofmans, J., & Verresen N. (2007). A functional measurement inquiry on the contribution of different life domains to overall subjective well-being. *Teori & Modelli*, n.s., XII, 1–2, 181–189.

Van Landeghem, B. (2008). *Human well-being over the life cycle: longitudinal evidence from a 20-year panel*. LICOS Discussion Paper No. 213/2008, Katholieke Universiteit Leuven, 2008. https://doi.org/10.2139/ssrn.1360731

Van Landeghem, B. (2012). A test for convexity of human well-being over the lifecycle: longitudinal evidence from a 20-year panel. *Journal of Economic Behavior and Organization*, *81*(2), 571–585. https://doi.org/10.1016/j.jebo.2011.08.001

Veenhoven, R. (2008). Healthy happiness: effects of happiness on physical health and the consequences for preventive health care. *Journal of Happiness Studies*, *9*, 449–469. https://doi.org/10.1007/s10902-006-9042-1

Verbeek, M., & Nijman, T. (1992). Testing for selectivity bias in panel data models. *International Economic Review*, *33*(3), 681–703. https://doi.org/10.2307/2527133

Weiss, A., King, J. E., Inoue-Murayama, M., Matsuzawa, T., & Oswald, A. J. (2012). Evidence for a midlife crisis in great apes consistent with the U-shape in human well-being. *Proceedings of the National Academy of Sciences*, *109*(49), 19949–19952. https://doi.org/10.1073/pnas.1212592109

Wethington, E. (2000). Expecting Stress: Americans and the "Midlife Crisis". *Motivation and Emotion*, *24*(2), 85–103. https://doi.org/10.1023/A:1005611230993

Wooden, M. & Li, N. (2014). Panel conditioning and subjective well-being. *Social Indicators Research*, *117*, 235–255. https://doi.org/10.1007/s11205-013-0348-1

Endnotes

1. Wethington (2000) does, however, point out that this high prevalence is partly due to respondents' tendency to use the term 'midlife crisis' quite broadly and to encompass events that occurred any time between the ages of about 30 to 70. Quoting Wethington (2000, p. 99): "It also implies a parsimonious explanation for why beliefs that the midlife crisis is a common risk of aging are so persistent. Almost any event or feeling socially symbolic of aging can qualify as a midlife crisis, if the definition is very elastic."
2. A detailed description of this study can be found in Huinink et al. (2011).
3. The results of these robustness tests are available upon request.
4. Interestingly, Galambos et al. (2015) in their longitudinal study actually observe an *increase* in happiness during young adulthood. Two points, however, must be stressed when comparing this study with ours. First, Galambos et al. (2015) follow a small group of individuals from working- and middle-class neighborhoods in a large western city in Canada, i.e. their sample cannot be compared with ours. Second, and importantly, they collect happiness data with a three-point scale, which does not allow for much variability in responses.
5. Available upon request.

Appendix

Table A.1. Overall life satisfaction, unconditional mean and model predictions from OLS and FE estimations including standard errors

	Full sample						Men						Women					
	Mean		OLS		FE		Mean		OLS		FE		Mean		OLS		FE	
	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE
Cohort 1991-93																		
15–17	7.95	0.02	7.87	0.04	7.78	0.04	8.05	0.03	7.96	0.05	7.90	0.06	7.84	0.03	7.76	0.05	7.65	0.05
16–18	8.05	0.03	8.00	0.03	7.92	0.04	8.12	0.03	8.07	0.05	8.00	0.06	7.98	0.04	7.94	0.05	7.84	0.05
17–19	7.83	0.03	7.81	0.03	7.75	0.03	7.95	0.04	7.91	0.04	7.87	0.04	7.70	0.04	7.70	0.05	7.63	0.04
18–20	7.74	0.03	7.70	0.03	7.72	0.03	7.83	0.04	7.76	0.05	7.77	0.04	7.65	0.04	7.63	0.05	7.68	0.04
19–21	7.71	0.03	7.72	0.03	7.78	0.03	7.79	0.04	7.80	0.05	7.84	0.04	7.64	0.05	7.64	0.05	7.71	0.04
20–22	7.71	0.03	7.67	0.04	7.73	0.03	7.78	0.05	7.78	0.05	7.84	0.05	7.64	0.05	7.56	0.05	7.63	0.05
21–23	7.65	0.04	7.64	0.04	7.70	0.04	7.71	0.05	7.72	0.05	7.77	0.07	7.59	0.05	7.56	0.05	7.65	0.05
22–24	7.58	0.04	7.55	0.04	7.64	0.05	7.65	0.05	7.63	0.06	7.71	0.08	7.51	0.05	7.47	0.06	7.57	0.06
Cohort 1981–83																		
25–27	7.41	0.03	7.51	0.03	7.47	0.04	7.31	0.04	7.52	0.05	7.46	0.06	7.50	0.04	7.51	0.05	7.47	0.06
26–28	7.44	0.03	7.61	0.03	7.55	0.04	7.36	0.04	7.57	0.05	7.52	0.06	7.52	0.04	7.63	0.05	7.57	0.06
27–29	7.43	0.03	7.52	0.03	7.49	0.03	7.40	0.05	7.51	0.05	7.48	0.04	7.47	0.04	7.53	0.04	7.50	0.04
28–30	7.44	0.03	7.47	0.03	7.47	0.03	7.35	0.05	7.40	0.05	7.39	0.04	7.52	0.05	7.54	0.04	7.54	0.03
29–31	7.48	0.04	7.46	0.03	7.48	0.03	7.41	0.05	7.40	0.05	7.42	0.04	7.55	0.05	7.52	0.04	7.54	0.04
30–32	7.56	0.04	7.51	0.04	7.54	0.03	7.43	0.06	7.36	0.05	7.40	0.05	7.69	0.05	7.65	0.05	7.67	0.04
31–33	7.53	0.04	7.43	0.04	7.49	0.04	7.44	0.06	7.34	0.06	7.41	0.06	7.62	0.05	7.52	0.05	7.57	0.06
32–34	7.49	0.04	7.34	0.04	7.44	0.05	7.40	0.06	7.23	0.06	7.33	0.07	7.57	0.06	7.45	0.06	7.54	0.07
Cohort 1971–73																		
35–37	7.48	0.03	7.48	0.03	7.48	0.04	7.43	0.04	7.51	0.04	7.52	0.06	7.52	0.04	7.44	0.04	7.44	0.06
36–38	7.48	0.03	7.57	0.03	7.57	0.05	7.37	0.04	7.48	0.05	7.53	0.07	7.57	0.04	7.63	0.04	7.60	0.07
37–39	7.39	0.03	7.42	0.03	7.45	0.03	7.32	0.05	7.39	0.04	7.43	0.04	7.45	0.04	7.46	0.04	7.47	0.04
38–40	7.31	0.03	7.34	0.03	7.34	0.03	7.22	0.05	7.29	0.05	7.28	0.04	7.37	0.05	7.40	0.04	7.40	0.03
39–41	7.43	0.04	7.44	0.03	7.42	0.03	7.38	0.05	7.40	0.05	7.35	0.04	7.48	0.05	7.48	0.05	7.48	0.04
40–42	7.40	0.04	7.35	0.03	7.36	0.04	7.33	0.05	7.31	0.05	7.30	0.05	7.44	0.05	7.39	0.05	7.41	0.05
41–43	7.40	0.04	7.36	0.04	7.36	0.05	7.36	0.06	7.33	0.05	7.31	0.07	7.42	0.05	7.39	0.05	7.41	0.06
42–44	7.44	0.04	7.35	0.04	7.35	0.06	7.36	0.06	7.28	0.06	7.24	0.09	7.50	0.05	7.43	0.05	7.44	0.08

Note: OLS and fixed-effects models include the following control variables: marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalized net household income, whether another person was present during the interview, GDP per capita, and unemployment rate.

Table A.2. Satisfaction with school, education, career, unconditional mean and model predictions from OLS and FE estimations including standard errors

	Full sample						Men						Women					
	Mean		OLS		FE		Mean		OLS		FE		Mean		OLS		FE	
	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE
Cohort 1991–93																		
15–17	7.36	0.03	7.34	0.04	7.25	0.05	7.37	0.04	7.37	0.06	7.33	0.08	7.35	0.04	7.30	0.06	7.19	0.08
16–18	7.35	0.03	7.35	0.05	7.26	0.06	7.42	0.05	7.45	0.06	7.40	0.09	7.28	0.05	7.24	0.07	7.14	0.08
17–19	7.37	0.04	7.36	0.04	7.36	0.04	7.46	0.05	7.40	0.06	7.44	0.06	7.28	0.05	7.31	0.07	7.29	0.06
18–20	7.50	0.04	7.45	0.05	7.51	0.04	7.56	0.05	7.48	0.06	7.53	0.05	7.44	0.06	7.41	0.07	7.48	0.06
19–21	7.51	0.04	7.48	0.05	7.54	0.05	7.59	0.06	7.60	0.06	7.62	0.06	7.42	0.06	7.35	0.07	7.43	0.07
20–22	7.49	0.04	7.42	0.05	7.48	0.05	7.53	0.06	7.52	0.07	7.54	0.08	7.45	0.06	7.33	0.07	7.41	0.07
21–23	7.48	0.05	7.52	0.05	7.55	0.06	7.57	0.06	7.63	0.07	7.61	0.09	7.40	0.07	7.41	0.07	7.46	0.08
22–24	7.50	0.05	7.45	0.05	7.49	0.07	7.60	0.07	7.54	0.07	7.55	0.11	7.39	0.07	7.35	0.08	7.39	0.09
Cohort 1981–83																		
25–27	7.14	0.04	7.13	0.05	7.08	0.06	7.13	0.05	7.24	0.06	7.18	0.08	7.14	0.05	7.03	0.06	6.99	0.08
26–28	6.99	0.04	7.10	0.04	7.03	0.06	7.01	0.05	7.20	0.06	7.14	0.08	6.98	0.06	7.01	0.06	6.92	0.08
27–29	7.09	0.04	7.13	0.04	7.12	0.04	7.26	0.06	7.33	0.06	7.34	0.06	6.94	0.06	6.95	0.06	6.92	0.05
28–30	7.13	0.04	7.12	0.04	7.11	0.04	7.26	0.06	7.27	0.06	7.27	0.05	7.02	0.06	6.99	0.06	6.98	0.05
29–31	7.13	0.05	7.08	0.04	7.11	0.04	7.30	0.07	7.21	0.06	7.24	0.05	6.99	0.07	6.97	0.06	7.00	0.05
30–32	7.10	0.05	7.05	0.05	7.08	0.05	7.19	0.07	7.09	0.07	7.13	0.07	7.03	0.07	7.01	0.07	7.05	0.07
31–33	7.17	0.05	7.11	0.05	7.16	0.06	7.24	0.07	7.17	0.07	7.18	0.08	7.13	0.07	7.07	0.07	7.15	0.08
32–34	7.27	0.05	7.16	0.05	7.26	0.07	7.30	0.08	7.16	0.07	7.25	0.09	7.23	0.08	7.15	0.07	7.26	0.09
Cohort 1971–73																		
35–37	7.18	0.03	7.17	0.04	7.25	0.06	7.21	0.05	7.21	0.06	7.32	0.08	7.15	0.05	7.12	0.06	7.19	0.08
36–38	7.08	0.04	7.21	0.04	7.18	0.07	7.04	0.05	7.15	0.06	7.17	0.09	7.11	0.05	7.25	0.06	7.17	0.10
37–39	7.09	0.04	7.14	0.04	7.13	0.04	7.11	0.06	7.18	0.06	7.22	0.05	7.07	0.06	7.12	0.06	7.06	0.06
38–40	7.12	0.04	7.13	0.04	7.10	0.03	7.16	0.06	7.21	0.06	7.17	0.05	7.10	0.06	7.07	0.06	7.05	0.05
39–41	7.18	0.04	7.16	0.04	7.13	0.04	7.29	0.06	7.31	0.05	7.24	0.05	7.09	0.06	7.04	0.06	7.05	0.06
40–42	7.21	0.05	7.11	0.04	7.11	0.05	7.30	0.07	7.25	0.06	7.20	0.06	7.14	0.06	6.99	0.06	7.04	0.08
41–43	7.21	0.05	7.10	0.04	7.10	0.06	7.28	0.07	7.20	0.06	7.17	0.09	7.16	0.07	7.03	0.06	7.05	0.10
42–44	7.22	0.05	7.06	0.05	7.08	0.08	7.30	0.07	7.18	0.06	7.17	0.11	7.15	0.07	6.97	0.06	7.02	0.12

Note: OLS and fixed-effects models include the following control variables: marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalized net household income, whether another person was present during the interview, GDP per capita, and unemployment rate.

Table A.3. Satisfaction with leisure activities, hobbies, interests, unconditional mean and model predictions from OLS and FE estimations including standard errors

	Full sample						Men						Women					
	Mean		OLS		FE		Mean		OLS		FE		Mean		OLS		FE	
	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE
Cohort 1991–93																		
15–17	8.23	0.03	8.13	0.04	8.12	0.05	8.45	0.04	8.36	0.06	8.40	0.07	8.01	0.04	7.89	0.06	7.83	0.08
16–18	7.93	0.03	7.88	0.04	7.88	0.05	8.19	0.04	8.13	0.06	8.19	0.07	7.65	0.05	7.60	0.06	7.56	0.08
17–19	7.68	0.03	7.62	0.04	7.61	0.04	7.99	0.04	7.93	0.05	7.93	0.05	7.35	0.05	7.28	0.06	7.29	0.06
18–20	7.57	0.04	7.57	0.04	7.57	0.04	7.80	0.05	7.79	0.05	7.78	0.05	7.33	0.06	7.33	0.07	7.36	0.06
19–21	7.33	0.04	7.30	0.05	7.31	0.04	7.65	0.05	7.64	0.06	7.62	0.06	6.99	0.06	6.95	0.07	6.98	0.06
20–22	7.26	0.04	7.26	0.05	7.28	0.05	7.52	0.05	7.55	0.06	7.51	0.06	6.99	0.06	6.97	0.08	7.04	0.07
21–23	7.16	0.04	7.20	0.05	7.20	0.05	7.47	0.06	7.51	0.06	7.47	0.07	6.85	0.07	6.90	0.08	6.91	0.08
22–24	7.06	0.05	7.08	0.05	7.06	0.07	7.28	0.06	7.26	0.07	7.22	0.09	6.84	0.07	6.89	0.08	6.87	0.09
Cohort 1981–83																		
25–27	7.14	0.03	6.97	0.04	6.98	0.05	7.38	0.05	7.30	0.07	7.25	0.08	6.90	0.05	6.69	0.06	6.71	0.07
26–28	6.97	0.04	6.88	0.05	6.84	0.05	7.24	0.05	7.14	0.07	7.08	0.09	6.71	0.05	6.65	0.06	6.61	0.07
27–29	6.75	0.04	6.71	0.04	6.70	0.04	6.97	0.06	6.91	0.06	6.91	0.05	6.55	0.06	6.54	0.06	6.51	0.05
28–30	6.81	0.04	6.79	0.04	6.80	0.03	6.95	0.06	6.96	0.06	6.97	0.05	6.68	0.06	6.66	0.06	6.65	0.05
29–31	6.70	0.04	6.71	0.04	6.70	0.04	6.91	0.06	6.91	0.06	6.90	0.06	6.52	0.06	6.54	0.06	6.54	0.05
30–32	6.70	0.05	6.73	0.05	6.74	0.05	6.81	0.07	6.83	0.07	6.87	0.07	6.60	0.07	6.65	0.06	6.65	0.06
31–33	6.56	0.05	6.60	0.05	6.63	0.05	6.72	0.07	6.76	0.07	6.81	0.08	6.43	0.07	6.45	0.07	6.51	0.07
32–34	6.66	0.05	6.72	0.05	6.76	0.06	6.77	0.07	6.83	0.08	6.90	0.10	6.56	0.07	6.60	0.07	6.66	0.08
Cohort 1971–73																		
35–37	6.78	0.04	6.73	0.04	6.68	0.06	6.88	0.05	6.84	0.06	6.87	0.09	6.70	0.05	6.66	0.06	6.50	0.08
36–38	6.69	0.04	6.70	0.05	6.61	0.07	6.76	0.05	6.73	0.07	6.79	0.09	6.64	0.05	6.68	0.06	6.42	0.09
37–39	6.54	0.04	6.54	0.04	6.51	0.04	6.66	0.06	6.63	0.06	6.67	0.06	6.44	0.06	6.47	0.06	6.35	0.06
38–40	6.64	0.04	6.64	0.04	6.66	0.03	6.70	0.06	6.72	0.06	6.73	0.05	6.59	0.06	6.57	0.06	6.60	0.05
39–41	6.60	0.04	6.60	0.04	6.63	0.04	6.69	0.06	6.72	0.06	6.69	0.06	6.53	0.06	6.50	0.06	6.61	0.05
40–42	6.70	0.05	6.66	0.04	6.72	0.05	6.72	0.07	6.74	0.06	6.72	0.07	6.68	0.06	6.58	0.06	6.75	0.07
41–43	6.58	0.05	6.56	0.05	6.62	0.06	6.64	0.07	6.67	0.07	6.59	0.09	6.53	0.07	6.46	0.06	6.69	0.08
42–44	6.65	0.05	6.60	0.05	6.67	0.08	6.66	0.07	6.68	0.07	6.58	0.12	6.64	0.07	6.52	0.07	6.79	0.11

Note: OLS and fixed-effects models include the following control variables: marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalized net household income, whether another person was present during the interview, GDP per capita, and unemployment rate.

Table A.4. Satisfaction with friends and social contacts, unconditional mean and model predictions from OLS and FE estimations including standard errors

	Full sample						Men						Women					
	Mean		OLS		FE		Mean		OLS		FE		Mean		OLS		FE	
	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE
Cohort 1991–93																		
15–17	8.79	0.02	8.72	0.04	8.71	0.05	8.77	0.03	8.72	0.05	8.71	0.06	8.81	0.04	8.73	0.05	8.72	0.07
16–18	8.61	0.03	8.54	0.04	8.53	0.05	8.63	0.04	8.57	0.05	8.54	0.07	8.59	0.04	8.52	0.05	8.51	0.07
17–19	8.46	0.03	8.44	0.04	8.42	0.04	8.54	0.04	8.53	0.05	8.51	0.05	8.39	0.04	8.34	0.06	8.31	0.05
18–20	8.25	0.03	8.22	0.04	8.24	0.03	8.35	0.04	8.32	0.05	8.31	0.04	8.14	0.05	8.11	0.06	8.16	0.05
19–21	8.11	0.03	8.10	0.04	8.10	0.04	8.20	0.05	8.20	0.05	8.19	0.05	8.02	0.05	8.00	0.06	8.01	0.05
20–22	7.95	0.04	7.96	0.04	7.98	0.04	7.98	0.05	7.99	0.06	8.03	0.06	7.91	0.05	7.91	0.06	7.93	0.06
21–23	7.87	0.04	7.90	0.05	7.90	0.05	8.01	0.05	8.02	0.06	8.03	0.07	7.74	0.06	7.77	0.07	7.76	0.07
22–24	7.72	0.04	7.74	0.05	7.74	0.06	7.75	0.05	7.74	0.07	7.78	0.09	7.69	0.06	7.73	0.07	7.71	0.09
Cohort 1981–83																		
25–27	8.02	0.03	7.93	0.04	7.85	0.05	7.99	0.04	7.93	0.06	7.79	0.07	8.05	0.04	7.93	0.06	7.90	0.06
26–28	7.78	0.03	7.74	0.04	7.68	0.05	7.78	0.05	7.73	0.06	7.64	0.08	7.78	0.05	7.75	0.06	7.72	0.07
27–29	7.56	0.04	7.56	0.04	7.53	0.03	7.49	0.05	7.46	0.06	7.43	0.05	7.64	0.05	7.65	0.06	7.62	0.05
28–30	7.44	0.04	7.46	0.04	7.45	0.03	7.30	0.06	7.31	0.06	7.34	0.04	7.56	0.05	7.58	0.06	7.56	0.04
29–31	7.41	0.04	7.43	0.04	7.44	0.03	7.30	0.06	7.29	0.06	7.33	0.05	7.50	0.06	7.55	0.06	7.53	0.05
30–32	7.31	0.04	7.33	0.04	7.37	0.04	7.22	0.06	7.23	0.06	7.29	0.07	7.39	0.06	7.43	0.06	7.44	0.05
31–33	7.19	0.05	7.20	0.05	7.30	0.05	7.09	0.06	7.12	0.07	7.23	0.08	7.28	0.06	7.28	0.07	7.37	0.07
32–34	7.17	0.05	7.18	0.05	7.30	0.06	6.96	0.07	6.99	0.08	7.14	0.10	7.35	0.07	7.36	0.07	7.45	0.08
Cohort 1971–73																		
35–37	7.67	0.03	7.63	0.04	7.65	0.05	7.42	0.05	7.37	0.06	7.43	0.08	7.89	0.04	7.85	0.05	7.81	0.07
36–38	7.50	0.03	7.48	0.04	7.51	0.07	7.28	0.05	7.23	0.06	7.37	0.09	7.67	0.05	7.68	0.06	7.63	0.09
37–39	7.26	0.04	7.25	0.04	7.27	0.04	7.11	0.06	7.09	0.06	7.15	0.05	7.38	0.05	7.39	0.06	7.37	0.06
38–40	7.25	0.04	7.25	0.04	7.25	0.03	7.00	0.06	7.02	0.06	7.01	0.04	7.45	0.05	7.44	0.05	7.45	0.04
39–41	7.19	0.04	7.20	0.04	7.20	0.04	7.02	0.06	7.06	0.06	6.98	0.05	7.33	0.05	7.33	0.05	7.37	0.05
40–42	7.26	0.04	7.24	0.04	7.23	0.05	7.12	0.06	7.14	0.06	7.08	0.07	7.36	0.06	7.31	0.06	7.37	0.07
41–43	7.10	0.04	7.10	0.05	7.07	0.06	6.89	0.07	6.92	0.07	6.83	0.09	7.26	0.06	7.23	0.06	7.28	0.09
42–44	7.18	0.05	7.17	0.05	7.10	0.08	7.02	0.07	7.06	0.07	6.89	0.11	7.31	0.06	7.26	0.06	7.28	0.11

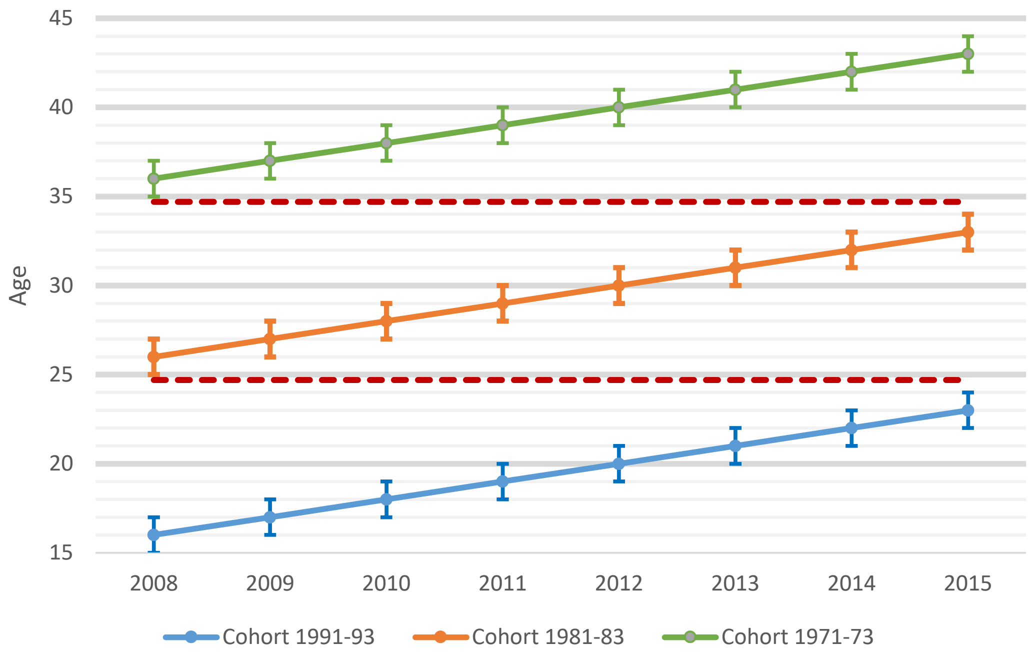
Note: OLS and fixed-effects models include the following control variables: marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalized net household income, whether another person was present during the interview, GDP per capita, and unemployment rate.

Table A.5. Satisfaction with family, unconditional mean and model predictions from OLS and FE estimations including standard errors

	Full sample						Men						Women					
	Mean		OLS		FE		Mean		OLS		FE		Mean		OLS		FE	
	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE
Cohort 1991–93																		
15–17	8.75	0.03	8.71	0.04	8.65	0.04	8.76	0.03	8.69	0.05	8.62	0.06	8.74	0.04	8.73	0.06	8.68	0.06
16–18	8.61	0.03	8.59	0.04	8.53	0.04	8.60	0.04	8.57	0.05	8.52	0.06	8.61	0.04	8.62	0.06	8.55	0.06
17–19	8.47	0.03	8.45	0.04	8.41	0.03	8.49	0.04	8.46	0.05	8.41	0.04	8.45	0.04	8.44	0.06	8.42	0.05
18–20	8.45	0.03	8.40	0.04	8.41	0.03	8.47	0.04	8.46	0.05	8.47	0.04	8.43	0.05	8.33	0.06	8.35	0.04
19–21	8.34	0.03	8.32	0.04	8.36	0.03	8.32	0.05	8.31	0.05	8.36	0.05	8.36	0.05	8.32	0.06	8.37	0.05
20–22	8.36	0.04	8.36	0.04	8.42	0.04	8.31	0.05	8.33	0.06	8.41	0.06	8.40	0.05	8.38	0.06	8.43	0.06
21–23	8.32	0.04	8.30	0.05	8.35	0.04	8.28	0.05	8.29	0.06	8.36	0.06	8.36	0.06	8.31	0.07	8.34	0.06
22–24	8.28	0.04	8.27	0.05	8.33	0.05	8.18	0.06	8.20	0.06	8.27	0.08	8.38	0.06	8.34	0.06	8.39	0.07
Cohort 1981–83																		
25–27	8.53	0.03	8.73	0.04	8.66	0.05	8.38	0.04	8.63	0.05	8.51	0.07	8.69	0.04	8.81	0.05	8.78	0.06
26–28	8.47	0.03	8.57	0.04	8.52	0.05	8.38	0.05	8.52	0.05	8.40	0.07	8.55	0.04	8.60	0.05	8.62	0.07
27–29	8.33	0.03	8.38	0.04	8.39	0.03	8.22	0.05	8.26	0.05	8.26	0.05	8.44	0.05	8.49	0.05	8.51	0.04
28–30	8.32	0.04	8.33	0.04	8.35	0.03	8.18	0.05	8.19	0.05	8.24	0.04	8.45	0.05	8.45	0.05	8.46	0.04
29–31	8.25	0.04	8.26	0.04	8.28	0.03	8.10	0.06	8.09	0.06	8.14	0.05	8.38	0.05	8.40	0.05	8.40	0.04
30–32	8.23	0.04	8.17	0.04	8.19	0.04	8.06	0.06	7.98	0.06	8.03	0.06	8.39	0.05	8.34	0.05	8.33	0.05
31–33	8.15	0.04	8.07	0.04	8.10	0.05	7.99	0.06	7.89	0.06	7.95	0.07	8.30	0.06	8.24	0.06	8.25	0.07
32–34	8.19	0.04	8.09	0.05	8.15	0.06	7.94	0.07	7.82	0.07	7.95	0.09	8.40	0.06	8.34	0.06	8.33	0.09
Cohort 1971–73																		
35–37	8.54	0.03	8.63	0.03	8.59	0.04	8.47	0.04	8.65	0.05	8.65	0.06	8.60	0.04	8.59	0.05	8.53	0.06
36–38	8.42	0.03	8.37	0.04	8.34	0.05	8.37	0.05	8.34	0.05	8.39	0.07	8.46	0.04	8.40	0.05	8.30	0.08
37–39	8.30	0.03	8.27	0.04	8.28	0.03	8.23	0.05	8.20	0.05	8.23	0.05	8.37	0.05	8.34	0.05	8.32	0.05
38–40	8.27	0.04	8.29	0.04	8.31	0.03	8.25	0.05	8.29	0.05	8.30	0.04	8.28	0.05	8.29	0.05	8.32	0.04
39–41	8.18	0.04	8.22	0.04	8.24	0.03	8.21	0.06	8.26	0.05	8.23	0.04	8.16	0.05	8.20	0.05	8.25	0.05
40–42	8.23	0.04	8.23	0.04	8.26	0.04	8.26	0.06	8.27	0.06	8.25	0.06	8.20	0.05	8.20	0.05	8.26	0.06
41–43	8.16	0.04	8.20	0.04	8.22	0.05	8.09	0.06	8.13	0.06	8.11	0.07	8.21	0.05	8.25	0.05	8.30	0.07
42–44	8.16	0.04	8.21	0.04	8.22	0.06	8.15	0.06	8.19	0.06	8.14	0.09	8.17	0.06	8.22	0.06	8.29	0.09

Note: OLS and fixed-effects models include the following control variables: marital status, number of children, self-rated health, employment status (being unemployed, not in the labour force), the natural logarithm of equivalized net household income, whether another person was present during the interview, GDP per capita, and unemployment rate.

Figure A.1: Sequential cohort design of pairfam data, 2008-2015



Supplementary material

Table S.1. Overall life satisfaction, OLS, and FE estimates, full sample

	1991–93				1981–83				1971–73			
	OLS		FE		OLS		FE		OLS		FE	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Unemployed	-1.073***	0.10	-0.775***	0.09	-0.887***	0.07	-0.459***	0.07	-0.886***	0.09	-0.547***	0.09
Not in labour force	0.017	0.03	-0.033	0.04	0.071*	0.04	0.141***	0.04	-0.054	0.06	0.008	0.06
Married	0.404**	0.18	0.331*	0.17	0.480***	0.04	0.064	0.05	0.483***	0.04	0.183**	0.07
Number of children	-0.087	0.11	0.061	0.12	0.042*	0.02	0.015	0.04	0.064***	0.02	0.012	0.04
Good health	0.652***	0.03	0.327***	0.03	0.686***	0.03	0.333***	0.03	0.674***	0.03	0.267***	0.03
Bad health	-0.455***	0.06	-0.407***	0.05	-0.504***	0.06	-0.434***	0.05	-0.719***	0.06	-0.459***	0.05
Ln equivalised h'hold income	0.204***	0.03	0.080***	0.03	0.456***	0.04	0.186***	0.04	0.559***	0.04	0.327***	0.05
Others present during interview	-0.019	0.05	0.072	0.05	0.105***	0.04	0.084**	0.04	0.108***	0.04	0.132***	0.03
Unemployment rate	-0.020**	0.01	0.051***	0.02	-0.015*	0.01	-0.008	0.02	-0.008	0.01	-0.018	0.02
BIP per capita	-0.000	0.00	-0.000	0.00	-0.000**	0.00	-0.000	0.00	0.000	0.00	0.000	0.00
Year 2009	0.138***	0.04	0.138***	0.04	0.091**	0.04	0.082**	0.04	0.088**	0.04	0.088**	0.04
Year 2010	-0.059	0.04	-0.027	0.04	0.006	0.04	0.028	0.04	-0.058	0.04	-0.028	0.04
Year 2011	-0.169***	0.05	-0.057	0.06	-0.040	0.05	0.006	0.06	-0.140***	0.04	-0.137**	0.06
Year 2012	-0.145***	0.05	-0.003	0.06	-0.049	0.05	0.016	0.06	-0.042	0.04	-0.060	0.06
Year 2013	-0.201***	0.05	-0.047	0.07	-0.004	0.05	0.078	0.07	-0.133***	0.05	-0.118	0.07
Year 2014	-0.228***	0.05	-0.077	0.07	-0.085	0.05	0.026	0.07	-0.123***	0.05	-0.120	0.08
Year 2015	-0.319***	0.06	-0.141*	0.08	-0.173***	0.05	-0.027	0.08	-0.129***	0.05	-0.130	0.09
Constant	6.254***	0.26	6.803***	0.37	4.031***	0.31	6.218***	0.38	2.729***	0.36	4.960***	0.46
Number of observations		14,670		14,670		17,869		17,869		19,888		19,888
Number of groups				3,607				4,174				4,384
R^2 overall				0.080				0.144				0.188
R^2 within				0.059				0.044				0.042
R^2 between				0.083				0.169				0.253

Note: *** p<0.01, ** p<0.05, * p<0.1

Table S.2. Satisfaction with school, education, career, OLS, and FE estimates, full sample

	1991–93				1981–83				1971–73			
	OLS		FE		OLS		FE		OLS		FE	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Unemployed	-3.433***	0.15	-2.912***	0.16	-2.322***	0.11	-1.800***	0.12	-2.403***	0.11	-1.566***	0.13
Not in labour force	-0.250***	0.04	-0.315***	0.05	-0.315***	0.06	-0.158***	0.06	-0.644***	0.08	-0.410***	0.08
Married	-0.364	0.26	-0.317	0.24	0.172***	0.05	-0.057	0.06	0.042	0.05	-0.163*	0.08
Number of children	-0.780***	0.15	-0.328*	0.17	-0.124***	0.03	-0.048	0.05	0.058***	0.02	0.015	0.06
Good health	0.571***	0.04	0.278***	0.05	0.623***	0.04	0.213***	0.04	0.507***	0.04	0.165***	0.04
Bad health	-0.204***	0.07	-0.112	0.07	-0.254***	0.07	-0.168***	0.07	-0.446***	0.06	-0.261***	0.06
Ln equivalised h'hold income	0.117***	0.03	-0.027	0.04	0.475***	0.04	0.074	0.05	0.677***	0.05	0.277***	0.06
Others present during interview	-0.203***	0.07	0.002	0.07	-0.099*	0.05	-0.021	0.05	-0.058	0.05	-0.011	0.04
Unemployment rate	-0.047***	0.01	0.007	0.03	-0.016	0.01	0.007	0.02	-0.016	0.01	-0.025	0.03
BIP per capita	-0.000**	0.00	-0.000	0.00	-0.000***	0.00	-0.000	0.00	-0.000	0.00	-0.000	0.00
Year 2009	0.015	0.05	0.010	0.05	-0.027	0.05	-0.055	0.06	0.044	0.05	-0.079	0.05
Year 2010	0.020	0.06	0.107*	0.06	-0.002	0.06	0.034	0.06	-0.022	0.05	-0.124**	0.05
Year 2011	0.110*	0.06	0.257***	0.08	-0.011	0.06	0.031	0.08	-0.034	0.05	-0.153**	0.08
Year 2012	0.143**	0.06	0.285***	0.09	-0.048	0.06	0.025	0.08	-0.005	0.06	-0.125	0.09
Year 2013	0.088	0.07	0.232**	0.09	-0.080	0.07	0.000	0.09	-0.061	0.06	-0.145	0.10
Year 2014	0.184***	0.07	0.302***	0.10	-0.017	0.07	0.077	0.10	-0.062	0.06	-0.156	0.11
Year 2015	0.111	0.07	0.242**	0.11	0.032	0.07	0.172	0.11	-0.106*	0.06	-0.172	0.13
Constant	7.069***	0.30	7.834***	0.46	4.123***	0.39	6.962***	0.50	2.314***	0.41	5.877***	0.62
Number of observations		14,640		14,640		17,763		17,763		19,758		19,758
Number of groups				3,607				4,168				4,374
R^2 overall				0.115				0.139				0.148
R^2 within				0.078				0.050				0.035
R^2 between				0.129				0.192				0.202

Note: *** p<0.01, ** p<0.05, * p<0.1

Table S.3. Satisfaction with leisure activities, hobbies, interests, OLS, and FE estimates, full sample

	1991–93				1981–83				1971–73			
	OLS		FE		OLS		FE		OLS		FE	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Unemployed	-0.150	0.11	-0.036	0.12	-0.049	0.08	0.146*	0.08	-0.372***	0.10	-0.152	0.10
Not in labour force	0.033	0.04	0.091**	0.05	-0.051	0.06	0.072	0.05	-0.164**	0.08	-0.005	0.07
Married	-0.543**	0.26	-0.281	0.25	-0.007	0.06	-0.140**	0.06	0.143**	0.06	-0.138	0.09
Number of children	-0.550***	0.12	-0.313*	0.16	-0.316***	0.03	-0.347***	0.05	-0.177***	0.03	-0.393***	0.06
Good health	0.517***	0.05	0.225***	0.04	0.638***	0.05	0.281***	0.04	0.725***	0.04	0.231***	0.04
Bad health	-0.287***	0.07	-0.177***	0.07	-0.089	0.07	-0.082	0.06	-0.269***	0.07	-0.103**	0.05
Ln equivalised h'hold income	0.182***	0.03	0.060*	0.03	0.163***	0.04	-0.047	0.05	0.056	0.05	-0.019	0.05
Others present during interview	0.024	0.06	-0.051	0.07	-0.108**	0.05	-0.072	0.05	-0.046	0.05	-0.014	0.04
Unemployment rate	-0.019	0.01	0.021	0.02	-0.021*	0.01	-0.029	0.02	-0.032**	0.01	0.023	0.02
BIP per capita	-0.000**	0.00	0.000*	0.00	-0.000**	0.00	-0.000	0.00	-0.000	0.00	0.000	0.00
Year 2009	-0.253***	0.05	-0.242***	0.05	-0.095*	0.05	-0.134**	0.05	-0.028	0.05	-0.076	0.05
Year 2010	-0.514***	0.05	-0.505***	0.06	-0.264***	0.05	-0.277***	0.06	-0.186***	0.05	-0.177***	0.06
Year 2011	-0.566***	0.06	-0.549***	0.07	-0.181***	0.06	-0.180**	0.07	-0.090	0.06	-0.028	0.08
Year 2012	-0.832***	0.06	-0.806***	0.08	-0.264***	0.06	-0.278***	0.08	-0.126**	0.06	-0.050	0.09
Year 2013	-0.877***	0.06	-0.835***	0.09	-0.241***	0.06	-0.236***	0.09	-0.069	0.06	0.035	0.10
Year 2014	-0.929***	0.07	-0.920***	0.09	-0.376***	0.07	-0.345***	0.10	-0.171***	0.06	-0.061	0.11
Year 2015	-1.056***	0.07	-1.058***	0.10	-0.254***	0.07	-0.218**	0.11	-0.131**	0.07	-0.017	0.13
Constant	6.985***	0.31	6.880***	0.44	6.131***	0.39	7.874***	0.45	6.558***	0.47	7.038***	0.60
Number of observations		14,671		14,671		17,864		17,864		19,879		19,879
Number of groups				3,607				4,173				4,382
R^2 overall				0.060				0.046				0.018
R^2 within				0.055				0.023				0.012
R^2 between				0.075				0.057				0.025

Note: *** p<0.01, ** p<0.05, * p<0.1

Table S.4. Satisfaction with friends and social contacts, OLS, and FE estimates, full sample

	1991–93				1981–83				1971–73			
	OLS		FE		OLS		FE		OLS		FE	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Unemployed	-0.197*	0.11	-0.078	0.11	-0.137*	0.08	0.130*	0.07	-0.353***	0.11	0.008	0.10
Not in labour force	0.064*	0.04	0.068*	0.04	0.059	0.06	0.112**	0.05	-0.026	0.07	-0.002	0.06
Married	-0.457*	0.24	-0.197	0.21	0.045	0.05	-0.035	0.05	0.160***	0.06	-0.136*	0.08
Number of children	-0.551***	0.11	-0.420***	0.14	-0.111***	0.03	-0.321***	0.04	-0.039	0.02	-0.197***	0.05
Good health	0.435***	0.04	0.177***	0.04	0.467***	0.04	0.143***	0.04	0.524***	0.04	0.152***	0.03
Bad health	-0.178***	0.06	-0.113**	0.06	-0.106	0.07	-0.045	0.05	-0.189***	0.06	-0.046	0.05
Ln equivalised h'hold income	0.101***	0.03	-0.013	0.03	0.192***	0.04	-0.042	0.04	0.090*	0.05	0.008	0.05
Others present during interview	-0.055	0.05	-0.025	0.06	-0.121**	0.05	-0.042	0.04	-0.083*	0.05	0.009	0.04
Unemployment rate	-0.022**	0.01	0.014	0.02	-0.013	0.01	-0.013	0.02	0.006	0.01	0.038	0.02
BIP per capita	-0.000*	0.00	0.000	0.00	-0.000***	0.00	-0.000	0.00	0.000	0.00	0.000***	0.00
Year 2009	-0.172***	0.04	-0.185***	0.04	-0.188***	0.05	-0.166***	0.05	-0.153***	0.05	-0.135***	0.05
Year 2010	-0.279***	0.04	-0.298***	0.05	-0.368***	0.05	-0.315***	0.05	-0.382***	0.05	-0.376***	0.05
Year 2011	-0.497***	0.05	-0.477***	0.06	-0.473***	0.05	-0.396***	0.06	-0.387***	0.05	-0.399***	0.07
Year 2012	-0.612***	0.05	-0.613***	0.07	-0.501***	0.05	-0.412***	0.07	-0.430***	0.05	-0.453***	0.08
Year 2013	-0.759***	0.06	-0.729***	0.08	-0.598***	0.06	-0.481***	0.08	-0.396***	0.06	-0.417***	0.09
Year 2014	-0.820***	0.06	-0.818***	0.09	-0.732***	0.06	-0.548***	0.09	-0.534***	0.06	-0.575***	0.11
Year 2015	-0.978***	0.06	-0.969***	0.10	-0.752***	0.07	-0.544***	0.10	-0.463***	0.06	-0.548***	0.13
Constant	8.103***	0.28	8.107***	0.44	6.801***	0.37	8.607***	0.42	6.623***	0.45	6.429***	0.57
Number of observations		14,675		14,675		17,863		17,863		19,880		19,880
Number of groups				3,607				4,173				4,384
R^2 overall				0.060				0.023				0.007
R^2 within				0.059				0.036				0.020
R^2 between				0.072				0.022				0.007

Note: *** p<0.01, ** p<0.05, * p<0.1

Table S.5. Satisfaction with family, OLS, and FE estimates, full sample

	1991–93				1981–83				1971–73			
	OLS		FE		OLS		FE		OLS		FE	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Unemployed	-0.406***	0.11	-0.266***	0.10	-0.247***	0.08	0.016	0.08	-0.295***	0.10	0.029	0.08
Not in labour force	-0.039	0.04	-0.015	0.04	0.010	0.05	0.109***	0.04	-0.053	0.07	0.007	0.06
Married	0.256	0.18	0.413**	0.20	0.477***	0.05	0.255***	0.05	0.722***	0.05	0.336***	0.08
Number of children	-0.063	0.10	0.009	0.12	0.125***	0.03	0.096***	0.04	0.109***	0.02	0.140***	0.05
Good health	0.327***	0.04	0.130***	0.04	0.375***	0.04	0.111***	0.03	0.382***	0.04	0.115***	0.03
Bad health	-0.297***	0.07	-0.146**	0.06	-0.206***	0.06	-0.038	0.05	-0.302***	0.06	-0.110**	0.05
Ln equivalised h'hold income	0.110***	0.03	0.011	0.03	0.208***	0.04	0.047	0.04	0.064	0.04	-0.007	0.04
Others present during interview	0.161***	0.05	0.106**	0.05	0.177***	0.04	0.115***	0.04	0.218***	0.04	0.109***	0.04
Unemployment rate	-0.007	0.01	0.047**	0.02	-0.017	0.01	-0.009	0.02	0.023*	0.01	0.038**	0.02
BIP per capita	-0.000	0.00	0.000	0.00	-0.000***	0.00	-0.000	0.00	-0.000	0.00	-0.000	0.00
Year 2009	-0.112***	0.04	-0.118***	0.04	-0.168***	0.04	-0.142***	0.05	-0.255***	0.04	-0.249***	0.04
Year 2010	-0.256***	0.05	-0.234***	0.05	-0.351***	0.04	-0.273***	0.05	-0.351***	0.04	-0.309***	0.04
Year 2011	-0.310***	0.05	-0.238***	0.06	-0.403***	0.05	-0.312***	0.06	-0.336***	0.04	-0.278***	0.06
Year 2012	-0.387***	0.05	-0.286***	0.07	-0.479***	0.05	-0.389***	0.07	-0.402***	0.05	-0.347***	0.07
Year 2013	-0.346***	0.06	-0.231***	0.07	-0.566***	0.05	-0.477***	0.07	-0.393***	0.05	-0.331***	0.08
Year 2014	-0.405***	0.06	-0.297***	0.08	-0.665***	0.06	-0.561***	0.09	-0.430***	0.05	-0.371***	0.09
Year 2015	-0.434***	0.06	-0.317***	0.09	-0.648***	0.06	-0.512***	0.10	-0.419***	0.05	-0.363***	0.10
Constant	7.953***	0.29	7.861***	0.34	7.372***	0.37	8.479***	0.42	7.283***	0.38	8.142***	0.49
Number of observations	14,674		14,674		17,862		17,862		19,868		19,868	
Number of groups			3,607				4,173				4,382	
R^2 overall			0.020				0.048				0.064	
R^2 within			0.018				0.018				0.019	
R^2 between			0.024				0.061				0.092	

Note: *** p<0.01, ** p<0.05, * p<0.1

Longitudinal and Life Course Studies 2018 Volume 9 Issue 4 Pp 412–432 ISSN 1757-9597

Health effects of work and family transitions

Juli Simon Thomas Harvard University, USA
ejuliast@gmail.com

(Received March 2018 Revised September 2018) http://dx.doi.org/10.14301/llcs.v9i4.507

Abstract

Disruptive life events, including transitions in work or family structure, affect health. Research often focuses on one transition rather than thinking of an event framework in which respondents experience multiple transitions across qualitatively distinct domains. This paper contributes original evidence on the effects of event interaction, transition timing, and multiple occurrences of events on health outcomes. I look at employment loss, employment gain, marriage, and divorce as instances of disruptive transitions or instability in the life course; I analyse these events' effects on self-rated health and depression at ages 40 and 50. I show that employment losses and divorces have significant negative effects on health, and employment gains and marriages show smaller positive effects or null effects. Higher counts of transitions lead to stronger effects on health. Respondents who are older at event occurrence show larger negative effects, suggesting that work and family instability at early ages is not as detrimental to health as such instability at later ages. These results show that there are similarities across work and family domains in effects on health outcomes; moreover, experiencing several transitions can lead to overlaps in effects that might lessen or worsen health outcomes overall.

Keywords

Health outcomes; disruptions; transitions; divorce; job loss

Background

Disruptive life events affect one's health: for example, job insecurity (Ali & Avison, 1997; Ferrie, 2001) and marital dissolution (Prigerson, Maciejewski & Rosenheck, 1999) have been shown to affect a variety of health outcomes. However, this knowledge is gained by focusing on one event at a time rather than thinking of events within a framework in which respondents experience multiple transitions across qualitatively distinct domains, which fits more closely with people's lived experiences. I bring together a focus on transitions in the domains of work and family, rather than considering these domains to be separate entities and analysing them as such. These domains have been shown to be co-incident (e.g. a job loss prompting a divorce (Charles & Stephens, 2004; Sayer, England, Allison & Kangas, 2011)) and are therefore a good starting point to explore these ideas. The goal is to see if transitions across domains evoke similar health effects, which would suggest that a framework that considers both domains (e.g. considering these events together as an example of disruptions or instability in the life course rather than analysing them as separate domains) is beneficial.

A secondary goal is to see how transitions interact when they are co-incident in a specified timeframe. I look at employment loss and gain as well as marriage and divorce, and I analyse the effects of these events on self-rated health and depression at age 40. By juxtaposing losses and gains across these two domains, I aim to disentangle the potential perceived benefits of gains and the expected negative effects of losses while maintaining cross-domain interaction. Since all of these transitions are frequently coincident across domains in people's lives, they form part of the life course context, and considering them in the same framework matters.

Beyond transitions being co-incident with each other, each event can occur more than once, and events can occur in early or later adulthood; in this paper, I consider these possibilities as well. I draw from life course theory when adding the timing and co-incidences of events in adulthood. I also build on Wheaton's (1990) contextual approach to stress effects, research on cumulative risk and resilience (Evans, Li & Whipple, 2013; Masten, 2013; Masten, 2014; Masten & Monn, 2015), and the social epidemiology of stress exposure across the life course (Ben-Schlomo & Kuh, 2002; Bronfenbrenner 2004; Marmot 2005; Seeman & Crimmins, 2001).

The case for considering multiple transitions

The events considered in this paper coexist in the span of a life course, occurring on the path or trajectory of a person's life (Elder 1998; Elder, Johnson & Crosnoe, 2003; Wheaton & Gotlib, 1997). Employment gains and losses could be considered 'linked' in that a person who experiences one often experiences the other at some point in the life course as well, as are marriages and divorces. But individuals also experience events in both domains within their lives, and this combining of events over time is what creates different life course pathways. When looking at one event's occurrence, I consider the other events as potentially co-occurring across a specific age span. As such, I analyse each transition both as a single event and as an event that could co-occur with other transitions in the life course.

Primarily based in sociology and public health, the concept of cumulative (dis)advantage suggests that disruptions in early life shape or inform future choices and opportunities (DiPrete & Eirich, 2006). Indeed, it has been shown that early childhood disadvantage affects adult health (Repetti, Taylor & Seeman 2002; Turner, Thomas & Brown, 2016) and mortality (Hayward & Gorman, 2004), and that childhood health affects adult socioeconomic outcomes (Palloni, 2006); this connection is complex (e.g. Link & Phelan 1995) and riddled with intersectionalities (Adler & Ostrove, 1999). Also related, in studies of health outcomes, the concept of allostatic load is "a measure of the cumulative physiological burden exacted on the body through attempts to adapt to life's demands" (Seeman, McEwen, Rowe, & Singer, 2001); models of weathering (Geronimus, 1996) and accelerated aging (also called age-as-leveler; see e.g. Hayward, Miles, Crimmins, & Yang, 2000) are similar efforts to understand the biological wear and tear of cumulative (dis)advantage and possible resilience (Lowe, Rhodes & Waters 2015). All of these conceptual frameworks assume that previous events in life matter and perhaps persist in affecting future outcomes; it is possible for transitions to happen as a result of other transitions or be otherwise connected, especially when looking across a longer age span. Situated in life course theory, Wheaton (1990) calls this a person's "role history" prior to an event (p. 209).

The ideas of cumulative disadvantage, differential biological wear and tear (allostasis, weathering, accelerated aging), and role history suggest that, when considering the effects of one transition, other transitions matter as well, in some way. However, these are large bodies of literature, and this paper sets out neither to prove nor disprove these conceptual theories. Instead, I use these theories to underscore that it is important to consider the fact that other events matter when considering the effects of a single transition. Particularly in the realms of work and family, we find events that frequently occur in people's lives, and it is possible that they will co-occur. Thus, I posit that considering experiences of transitions across qualitatively distinct domains, and multiple occurrences of all disruptive transitions, matters for health outcomes. I remain agnostic on the reasons why multiple transitions occur.

Transitions as positive, negative, normative, or disruptive

It seems likely that the nature of the transition matters – one would guess that marriage is more of a positive event as compared to employment loss, generally speaking. Life course theory contends that some events are normative, such as completion of schooling, first marriage, or retirement (Riley & Riley, 1994; Uhlenberg & Mueller, 2004). Some events are non-normative, such as job loss; these events are sometimes unexpected (McLeod & Almazan, 2004) and could be considered turning points (Wheaton, 1997; Wheaton & Gotlib, 1997). Life course theory states that timing matters (Elder, Johnson, & Crosnoe, 2003; Mayer, 2009), and as such, events that are normative at some times are non-normative when off-time (McLeod & Almazan, 2004). I refer to all the events considered in this paper as *disruptive transitions*, though I do not imply negative effects with this terminology. Indeed, some of the events could be beneficial, and the timing of the events could dictate the effects' directionality. However, they do have the potential to disrupt the life course, and therefore I use this terminology. The work and family transitions considered in this paper are specific instances of disruptive transitions; there are many other transitions that could be considered within this framework, including but not limited to residential moves and health-related events (e.g., diagnoses of chronic illnesses, transition into parenthood). I focus on work and family transitions due to their likelihood of occurrence and co-occurrence. By comparing and contrasting transitions that are likely to occur and co-occur, I set the stage for considering other transitions. I address timing by analysing events in early versus later adulthood to see if there are differences in effects.

Effects of employment transitions

Loss of employment has been widely shown to affect physical and mental health negatively (Ali & Avison, 1997; Backhans & Hemmingsson, 2011; Bambra, 2011; Bartley, Ferrie & Montgomery, 2006; Burgard, Brand & House, 2009 Kasl & Jones, 2000; Kessler, House & Turner, 1987; Paul & Moser, 2009; Sleskova et al., 2006). It has been linked to higher mortality rates (Bambra, 2011; Morris, Cook & Shaper, 1994), especially during recessionary periods (Noelke & Beckfield, 2014). There is less work on the effects of employment gain on health. Some scholars have considered the possibility of re-employment in the aftermath of a job loss, and there is evidence that the damage of a job loss can be repaired by a subsequent job gain (Bartley, Ferrie & Montgomery, 2006; Kessler, Turner & House, 1989). However, Ali & Avison (1997) show that both single and married mothers transitioning into employment experience feelings of distress (single mothers for financial reasons and married mothers for increased caregiving stress). More generally, employment instability has been shown to be detrimental to health (Ferrie, Shipley, Marmot, Stansfeld, & Davey Smith, 1998; Ferrie, 2001; Frech & Damaske, 2012), implying that employment losses and gains are intricately linked. It is worth noting that work environment matters: poor working conditions have been linked to lower health outcomes (Bambra, 2014).

Effects of marital transitions

Marriage is generally shown to aid one's health (Frech & Williams, 2007; Uecker, 2012), even leading to lower mortality (Koball, Moiduddin, Henderson, Goesling, & Besculides, 2010). On the other hand, divorce has been shown to negatively impact one's health (Blekesaune, 2008; Hughes & Waite, 2009; Prigerson, et al., 1999). Some authors find that this connection holds especially true for women (Liu & Umberson, 2008) despite marital selection effects by health status (Cheung & Sloggett, 1998), though others find similar effects for men and women (Blekesaune, 2008). There is variation in the effects of marriage and divorce, in that leaving a harmful marriage can actually be beneficial (Booth & Amato, 2001), staying in a dysfunctional marriage can be damaging to one's health (Coyne & DeLongis, 1986; Hawkins & Booth, 2005; Kiecolt-Glaser & Newton, 2001), effects can vary based on the number of times one has entered and exited marital statuses (Blekesaune, 2008), and personal characteristics matter (Frech & Williams, 2007; Waldron, Hughes & Brooks, 1996). I focus on the average overall effects for both events.

Data and methods

I use data from the National Longitudinal Survey of Youth, 1979 (NLSY79) (U.S. Bureau of Labor Statistics, 2012). The NLSY79 is a cohort study of 12,686 respondents in the United States who were first interviewed in 1979 and followed annually through 1994, biannually since then (most recently in 2012). The sample began with youth born between 1957 and 1964; from a random sample of housing units in selected U.S. areas and a random sample of members of the military from Department of Defense records, participants were first screened and then assigned to sample groups. The NLSY79 is one of the most widely used longitudinal studies in the United States for its thoroughness and length. These factors also make it a good choice for looking at the multiple influences of work and family transitions over the life course.

The NLSY79 includes an insightful module of questions about the respondent's health that was given when the respondent was forty years old. Initially implemented to look at health limitations on work, this module was changed to provide a wealth of baseline health information at middle age. The main variables I use to assess health are from this module: a measure of self-rated health and a measure of depression. Self-rated health is assessed on a five-point Likert scale (excellent, very good, good, fair, poor), from which I create a binary indicator of good health (excellent, very good, good) or bad health (fair, poor). Self-rated health has been established as a good measure of overall health (Idler & Benyamini, 1997), and creating a binary measure has ample precedent (Case & Paxson, 2005). The depression score is created using the CES-D[1] set of questions; a score of 16 or above is considered an indicator of depression and a score of less than 16 on this scale means the respondent is not depressed. I use this cut-off to create a binary measure of depression, as suggested in NLSY79 documentation.

The health module includes measures of physical and mental health based on the SF-12[2] summary scale. Correlations between self-rated health and the included measure of physical health, and between the CES-D depression outcomes and the measure of mental health are high (0.61 and 0.66, respectively). I therefore use the self-rated health and CES-D measures for my analyses, as the latter is repeated at various intervals (necessary for sensitivity analyses) and the former encompasses the respondents' viewpoints. When possible, I point out where results for the measures chosen for this paper deviate from results from physical and mental health scores.

I start with the full sample of 12,686 respondents. I remove respondents who did not complete the health module's questions about self-rated health and the battery of questions that encompass the CES-D at age 40 (n=4,328); most of these (n=4,223) do not complete these questions because they are not eligible for the interview due to age requirements or they were dropped from the sample in those years because they were part of the military or poor-white oversample (Bureau of Labor Statistics, 2016). I remove individuals who are missing either measure due to question refusal or choice to skip the questions (n=105). I do not impute health values for the refusal/chosen skip individuals since these are my outcome variables (Wooldridge, 2006). I remove those who did not complete physical and mental health scores and a question about health limitations on work at age 18 that is used as a baseline health score in some models (n=111). My final sample includes 8,247 respondents.

Age at transition may matter (Kasl & Jones, 2000). To investigate possible age-based variation, I look at disruptive transitions experienced between the ages of 18 and 25 and again between the ages of 26 and 40. This roughly divides events into occurring during early versus mid-life working and relationship histories. This decision is also based on the average age of attaining a final level of education (Kena et al., 2016) and of first marriage (Goldstein & Kenney, 2001) during the waves of the health module's administration (1998–2002). In other words, ages 18–25 represent a more unstable time, during which respondents are completing education, entering the workforce for the first time, and transitioning into marriage. Ages 26–40, by contrast, represent a time when most respondents have concluded their education and have already entered the workforce. I look at events by number of occurrences; for employment events, this means looking at outcomes for one, two, or three events, and for marital events, this means looking at events happening once or twice. Thus, I create a set of event variables, for each event considered, as events ever occurring during the specified age range and as a count of events occurring during that age range; this includes events I can observe in the time period covered by the NLSY79. I separately analyse events ever occurring and event counts.

There are a number of important covariates to include; health outcomes have been shown to vary by factors such as social class (Adler et al., 1994; Blane, 2006; Marmot, 2005), education (Cutler & Lleras-Muney, 2010; Schnittker, 2004), and race (Krieger, Rowley, Hermann, Avery, & Phillips, 1993; Nazroo & Williams, 2006). I control for gender, race/ethnicity (Black, Hispanic), high school completion (12 years of education) at age 40, any college attendance (between 12 and 16 years of education) at age 40, college completion (16 or more years of schooling) at age 40, ability (measured by the ASVAB in 1981), number of children at age 40, count of years on welfare between ages 18 and 40, health limitations on work at age 18, Rotter Locus of Control score (measured in 1979), and the Rosenberg Self-Esteem Scale (10-item; measured in 1980). The last three variables provide a baseline of physical and mental health; though it would be preferable and advantageous to have the same health measures at the start of the survey as the outcome variables, these are not available. (This choice is further discussed in a later section.) Given research that shows differences by gender in terms of event experience (Nathanson, 1980) and health outcomes (Bird & Rieker, 1999; Crimmins & Saito, 2001; Mirowsky, 1996; Moen & Chermack, 2005), I stratify analyses by gender to examine differences.

Logistic regression models are used to determine effects of events on health outcomes at age 40. Regressions are unweighted, following recommendation from the NLSY79 (Bureau of Labor Statistics, 2016) and email correspondence with NLS staff (S. McClaskie, personal communication, April 6, 2016). Since models are not weighted, I decided not to weight the descriptive statistics I report, since these are simple reports and not analyses. For some models, I report both results for the full sample and results stratified by gender.

Results

Descriptive results

I start with an overview of descriptive variables and outcomes. Table 1 shows the means for covariates, relevant health variables, and event variables (all events occurring before age 40). For both depression and fair/poor health scores, more respondents are female and Black. Those in poor mental or physical health have fewer years of education, on average, and especially lower rates of college completion. They have lower average ability scores, higher Locus of Control scores (i.e. they feel less control over their lives), lower self-esteem scores, and slightly more children. They have spent significantly more years receiving welfare of some kind (AFDC, food stamps, SSI, or any other public assistance). Those in poor mental and physical health show lower physical and mental scores, as expected, and a slightly higher rate of health limitations on work at age 18.

Descriptive statistics show differences in event variables as well. Those who rate as depressed and those in fair or poor health experience more employment events, on average, particularly losses. They experience fewer marriages and more divorces. This illustrates that those in poorer physical and mental health seem to experience more disruptive events (and fewer positive events such as marriages), over the course of their lives. In this sample, 24% of respondents experience all four events before age 40, while 31% experience any three, illustrating that experiencing one event in the context of other events is common.

Logistic regression model results

In tables 2 and 3, I report the effect (log-odds) for each event separately (that is, a model looking at employment loss does not control for employment gain, marriage, or divorce) on depression and self-rated health, respectively. These models include events occurring between ages 18 and 40 (the variable is coded as 1 if events ever occurred, so this does not control for the number of times an event may have occurred in this time period), and depression and self-rated health are measured at age 40. Once controls for gender, race, education, ability, number of children, welfare receipt, health limitations on work, Locus of Control, and self-esteem are included, employment losses and divorces decrease self-rated health and increase depression. For self-rated health, these effects are stronger for men than for women. Interestingly, the reverse is true for the effect of employment loss on depression. However, these models do not account for events' co-incidence in the life course, nor do they address timing.

To analyse co-incidence, in tables 2 and 3, I report the effect for each event from models that include all four events on depression and self-rated health (listed as 'one model'). With controls, coefficient values for employment losses and divorces increase, and marriage is shown to be protective of depression, significantly so for women. Employment gains show counter-effects to losses, and marriages to divorces as well. Given this more nuanced view of a variety of events occurring in the life course, in the rest of the models in this paper I include all events in the same model.

As shown in descriptive statistics, it is not unlikely for people to experience events more than once, particularly for employment events. To address this variation, I create dummy variables for counts of one, two, three employment losses and gains, and counts of one or two marriages and divorces. I combine this count distinction with variation in timing in adulthood. Thus, I create the aforementioned indicators for ages 18–25 and 26–40 separately. I then run models containing all events, at each count (e.g. 1, 2, 3 occurrences), for each age group, with controls. The results from these models, for depression and self-rated health, are shown in tables 4 and 5. As the model constant, R-squared, n, and other information shows in the bottom rows of the tables, each set of two columns is one model; in other words, each of those tables

Table 1: Descriptive Statistics (means), by Outcome Variables (CES-D Score and Self-Rated Health Score)

	CES-D score < 16 (not depressed)	CES-D score ≥ 16 (depression)	Self-rated health (excellent, very good, good)	Self-rated health (fair, poor)
Covariates				
Male (0/1)	0.494	0.360	0.499	0.429
Black (0/1)	0.304	0.405	0.295	0.382
Hispanic (0/1)	0.194	0.194	0.187	0.241
Years of education attained (by age 40)	13.212	12.036	13.354	12.004
High school only (0/1; by age 40)	0.434	0.482	0.423	0.517
High school (0/1; by age 40)	0.892	0.761	0.906	0.765
College attendance (0/1; by age 40)	0.214	0.207	0.253	0.172
College completion (0/1; by age 40)	0.214	0.072	0.230	0.076
Ability (-3 - 3)	0.039	-0.432	0.084	-0.364
Rotter Locus of Control	8.708	9.468	8.635	9.362
Self-esteem	478.287	449.172	481.164	452.341
Number of children (by age 40)	1.924	2.176	1.907	2.083
Welfare (between ages 18 and 40)	0.361	0.712	0.329	0.648
Welfare (count, between ages 18 and 40)	1.852	5.311	1.578	4.266
Health variables				
Physical score (at age 40)	5226.390	4359.054	5379.249	4013.518
Mental score (at age 40)	5357.614	3230.703	5395.106	4660.735
Health limits (in 1979)	0.046	0.077	0.042	0.076
Event variables, before age 40				
Employment loss (ever)	0.709	0.820	0.700	0.791
Employment gain (ever)	0.810	0.860	0.807	0.838
Employment event (ever)	0.830	0.887	0.826	0.873
Marriage (ever)	0.727	0.613	0.736	0.643
Divorce (ever)	0.356	0.504	0.349	0.430
Marital event (ever)	0.747	0.667	0.754	0.679
Employment loss (count)	1.360	1.896	1.325	1.710
Employment gain (count)	1.648	2.018	1.630	1.851
Employment event (count)	3.008	3.914	2.954	3.560
Marriage (count)	0.913	0.829	0.920	0.851
Divorce (count)	0.445	0.667	0.436	0.552
Marital event (count)	1.358	1.495	1.356	1.403
N (out of 8,247)	8,025	222	7,183	1,064
Proportion of total sample	0.97	0.03	0.87	0.13

Notes: Both count and ever variables refer to events or coverage between the ages of 18 and 40 years. Sample includes all respondents who were included in the additional health questionnaire at age 40 and completed the CES-D and self-rated health questions and were non-missing on all other covariates.

Table 2: The Effects of Disruptive Events (at ages 18–40) on Depression at age 40

Event	All		Men		Women	
Separate models						
Employment losses	0.383	*	-0.153		0.619	*
Employment gains	0.193		-0.131		0.295	
Marriage	-0.091		-0.012		-0.182	
Divorce	0.462	**	0.595	*	0.346	†
One model						
Employment losses	0.434	†	-0.183		0.730	*
Employment gains	-0.166		0.026		-0.239	
Marriage	-0.487	†	-0.500		-0.532	*
Divorce	0.696	***	0.850	**	0.586	**

Notes: Results are log-odds. All models include controls for gender, race, education, ability, number of children, welfare receipt, health limitations on work, Rotter Locus of Control score, and Rosenberg Self-Esteem Scale score.

† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

Table 3: The Effects of Disruptive Events (at ages 18–40) on Self-Rated Health at age 40

Event	All		Men		Women	
Separate models						
Employment losses	-0.366	***	-0.507	***	-0.225	†
Employment gains	-0.144		-0.197		-0.064	
Marriage	-0.014		0.028		-0.001	
Divorce	-0.197	**	-0.305	**	-0.086	
One model						
Employment losses	-0.474	***	-0.679	***	-0.301	*
Employment gains	0.218		0.321		0.154	
Marriage	0.105		0.18		0.069	
Divorce	-0.234	**	-0.375	**	-0.111	

Notes: Results are log-odds. All models include controls for gender, race, education, ability, number of children, welfare receipt, health limitations on work, Rotter Locus of Control score, and Rosenberg Self-Esteem Scale score.

† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

Table 4: The Effects of Disruptive Events (at ages 18–25 & 26–40) on Depression at age 40, by timing and incidence

Event		All respondents 18-25	All respondents 26-40	Men 18-25	Men 26-40	Women 18-25	Women 26-40
Employment losses							
	1x	0.212	0.942 ***	0.178	0.535	0.199	1.120 ***
	2x	0.617 †	1.492 ***	0.880	0.424	0.553	1.947 ***
	3x	1.892 ***	1.918 ***	2.242 **	1.058	1.851 **	2.342 ***
Employment gains							
	1x	-0.092	-0.910 ***	-0.273	-0.738 *	-0.054	-1.053 ***
	2x	-0.595 *	-1.095 ***	-1.206 *	-0.639	-0.396	-1.309 ***
	3x	-1.053 †	-1.556 ***	1.188	-0.722	-1.259 †	-1.909 ***
Marriage							
	1x	-0.399 *	-0.729 ***	-0.364	-0.426	-0.562 *	-0.955 ***
	2x	-0.302	-0.487	-1.167	-1.019	-0.045	-0.320
Divorce							
	1x	-0.189	0.977 ***	0.587	1.225 ***	-0.001	0.873 ***
	2x	0.685	0.951 *	---	-0.035	0.839	1.164 *
Constant		-4.001 ***		-4.275 ***		-3.738 ***	
LR χ^2		235.90		123.74		139.62	
P > χ^2		0.000		0.000		0.000	
Pseudo R^2		0.124		0.171		0.121	
n		7,712		3,711		3,978	

Notes: Results are log-odds. All models include controls for gender, race, education, ability, number of children, welfare receipt, health limitations on work, Rotter Locus of Control score, and Rosenberg Self-Esteem Scale score.
† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

Table 5: The Effects of Disruptive Events (at ages 18–25 & 26–40) on Self-Rated Health at age 40, by timing and incidence

Event		All respondents 18-25	All respondents 26-40	Men 18-25	Men 26-40	Women 18-25	Women 26-40
Employment losses							
	1x	-0.169 †	-0.928 ***	-0.122	-1.264 ***	-0.199	-0.708 ***
	2x	-0.429 **	-1.578 ***	-0.293	-1.631 ***	-0.536 *	-1.528 ***
	3x	-1.188 ***	-1.956 ***	-1.283 **	-2.106 ***	-1.121 *	-1.876 ***
Employment gains							
	1x	0.066	0.592 ***	0.081	0.778 ***	0.062	0.512 ***
	2x	0.394 **	1.302 ***	0.379 †	1.188 ***	0.420 *	1.357 ***
	3x	0.907 **	1.582 ***	1.050 *	1.578 ***	0.812 *	1.547 ***
Marriage							
	1x	0.113	0.082	0.138	0.170	0.142	0.016
	2x	0.209	0.214	0.175	0.343	0.275	0.094
Divorce							
	1x	-0.226 †	-0.198 *	-0.383 †	0.363 *	-0.119	-0.061
	2x	-0.696 †	-0.259	-0.588	-0.209	-0.763	-0.259
Constant		2.024 ***		2.060 ***		1.831 ***	
LR χ^2		737.75		344.72		412.64	
P > χ^2		0.000		0.000		0.000	
Pseudo R^2		0.126		0.132		0.128	
n		7,712		3,734		3,978	

Notes: Results are log-odds. All models include controls for gender, race, education, ability, number of children, welfare receipt, health limitations on work, Rotter Locus of Control score, and Rosenberg Self-Esteem Scale score.
† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

contains results from three models in total. For employment losses and gains, there is a clear upward gradient for effects on both outcomes as event counts increase and age at event increases. Marriage is associated with decreases in depression, particularly when it occurs in the later age range, but not for those who experience more than one marriage. Divorces are associated with decreases in self-rated health regardless of their timing, but they are only associated with increases in depression when they occur in the later age range.

It is again clear that employment disruptions affect men's self-rated health but not mental health, whereas women's physical and mental health are affected by these disruptions. Marriage is associated with decreases in depression in women but has no significant effects for men. Divorce is associated with increases in depression in women only.

These results suggest that disruptive events do affect both self-rated health and depression, in slightly different ways for men and women. Co-incidence of events matters, as does the number of occurrences and timing in adulthood. Analyses using physical and mental health scores in lieu of self-rated health and depression scores produced similar results. However, there are several issues that arise with this analysis: reverse causality, temporal distance from event to outcome, and age ranges and time-variant characteristics. I address these concerns in the next sections.

Addressing reverse causality

Reverse causality is a fundamental issue in this analysis: those who are in poorer mental or physical health could be disproportionately likely to experience events (Adler & Ostrove, 1999; Lerner et al., 2004). The results reported include a control for health limitations at age 18 which somewhat addresses this concern. In this section, I use other pre-treatment variables to further examine the possibility of selection. I also run models to examine pre-treatment heterogeneity, where I treat selection into events by health as a form of heterogeneity.

The NLSY79 health module for age 40 is repeated at age 50, though the sample size is quite a bit smaller than the age 40 module, with the total sample for the age 50 module being 1,603 respondents. However, this allows for linkages from age 40 to age 50 for both self-rated health and depression variables. Event variables that pre-date age 40 cannot be included when using health scores at age 40 as a control in models, so I create new event variables that only include events between ages 40 and 50. Due to data limitations, I code these variables as having occurred any number of times versus never occurring; I lose the ability to examine variation by number of incidences in this analysis.

Employment losses and divorces are associated with an increase in depression while employment gains are associated with a decrease in depression (appendix A). Marriage shows no effects. Stratifying the analysis by gender suggests that the effect due to employment losses is driven by men whereas the effect due to divorce is driven by women. There is a strong negative association between employment losses and self-rated health for both men and women (appendix B). Results for women suggest an increase in self-rated health from employment gains. Results for marriage and divorce are not significant, which is likely a reflection of the smaller sample size. Taken as a whole, these results indeed reflect the same associations that the previous logistic models for the full sample showed.

The NLSY79 asks the same CES-D battery of questions in 1992 (7-item and 20-item versions) and 1994 (7-item version only). I use these questions to construct a depression score at age 30 (using those who were age 27–30 in 1992 and 1994). If the 20-item score is available in 1992, I use that value first, and then I fill in missing values with the other scores. Since events that occurred prior to age 30 pre-date this CES-D score, I create new event variables that only include events between ages 30 and 40 in the same way as explained previously for events between ages 40 and 50; again, I face the limitation of not being able to examine counts of events but rather focus on a dichotomous indicator of event occurrence only. Again, the direction of each event's effects on depression are the same as previous results, though only results for employment loss and divorce (for men only) are significant (appendix C).

There are innovative methodological techniques by which to address reverse causality, calculating an average treatment effect (ATE) for the population, which is the "mean causal effect for a unit whose characteristics are represented by [a set of covariates]" (Ho, Imai, King, & Stuart, 2007:204), or an average treatment effect on the treated (ATT), which is the mean causal effect for individuals who did experience the treatment/event. These techniques do not solve reverse causality outright, but they allow for more robust estimates and sensitivity analyses. To use one such pathway, I define potential selection by pre-existing health measures as a source of pre-treatment heterogeneity, since this is defined as "the propensity of selection into treatment" (Xie, Brand & Jann, 2012:2). This allows me to look at potential variation in effects due to pre-treatment selection (Brand & Simon Thomas, 2013; Xie, Brand & Jann, 2012). If there is systematic variation in effects by strata, created using pre-treatment controls, this provides evidence of reverse causality. As pre-treatment controls, I use the same control variables used in previous models, including event variables for the event not being addressed in each model, to calculate propensity scores for each event. I use logistic models to calculate propensity scores, at the default 0.01 significance level for balancing, using kernel matching for least bias (Morgan & Winship, 2007) and most sample inclusion (Caliendo & Kopeinig 2005; Garrido et al., 2014). Strata with few values are combined with the nearest neighboring stratum, following Harder, Stuart, and Anthony (2010). These propensity scores can be broken into strata by their values, and then results can be examined for each stratum and compared to each other. I do this for both the full sample, the sample of events between ages 30 and 40, and the sample of events between ages 40 and 50; in the first set of models, I again use health limitations at age 18 as a pre-treatment control variable, in the second set of models, I use the CES-D score at age 30 as this control, and in the third set I use the CES-D or self-rated health score at age 40 as this control.

Graphs of effects across strata (available upon request) do not show any specific patterns, for self-rated health or depression outcomes. For all three samples, effects do not clearly show a pattern across strata. I conclude that this sensitivity analysis does not show evidence of reverse causality. Events are more likely for those who experience other events (i.e. they have a higher propensity score for each event) more so than for those who have a history of poorer mental or physical health, which again speaks to the importance of the context of other events.

I use propensity scores to create weights, using the same control variables and logistic regression models to calculate weights. I 'weigh' each observation by the inverse of the propensity to be selected (Inverse Probability of Treatment Weights, or IPTW). In other words, those respondents who are highly likely to experience an event (i.e. receive the treatment) are 'down-weighted' and those who are less likely to experience an event are 'up-weighted' (Sampson, Sharkey & Raudenbush, 2008).

Since treatment assignment indeed appears to be 'ignorable', meaning that "there are no unobserved covariates related to the outcome that are also predictive of treatment group assignment once the observed covariates are controlled" (Sampson, Sharkey & Raudenbush, 2007:846), it is appropriate to continue with the IPTW analysis (see also Statacorp, 2013). Again, this means that logistic regression models are run similarly to the previous models, with the important inclusion of survey weights, which are inversely related to the probability of being treated, or experiencing an event. Thus, these survey weights adjust for the probability of experiencing events. Further information on covariate balance in the treatment weights is included in appendix F (for events during ages 18 and 40; similar statistics are available upon request for models for other age ranges). Overall results from the IPTW model are remarkably similar to prior results (appendix D), as shown in tables 2 and 3 (for the full sample), appendices A and B (for the ages 40–50 sample), and C (for the ages 30–40 sample). Again, employment losses and divorces are, on the whole, detrimental to one's physical and mental health, while employment gains show significant benefits for self-rated health (effects on depression are insignificant but in the direction of improvement), and marriages show mainly advantageous results.

These analyses by no means entirely disprove reverse causality. Indeed, unobserved covariates (e.g. concurrent health changes or underlying health issues exacerbated by the stress of disruptive events) could still cause biased estimates (see e.g. Sampson, Sharkey & Raudenbush (2008) for further discussion). However, the various analyses presented in this section do support estimates from prior models, providing consistency across model specifications and methodology.

Exploring temporal distance from events

Events that occurred during the earlier (18–25) age range are necessarily more temporally distant from the outcome at age 40 than events in the later age range (26–40). This means that effects from earlier events could have waned over time, whereas later events could be more salient at the time of the outcome. To test this possibility, I look at events occurring in the earlier age range, using variables for each number of incidences as in previous models, and use the CES-D score at age 30 as the outcome (self-rated health questions are not available at other ages). This makes the outcome more proximate to earlier events' occurrences.

Indeed, employment losses and divorces increase chances of depression, whereas marriage decreases these chances (appendix E). Employment gains are insignificant, though suggestive of providing a decrease in depression. Comparing these results to the 18–25 column in table 4, this suggests that over time, the effects of employment losses and gains from this earlier age range are exacerbated rather than waning. On the other hand, the effects of divorce and higher counts of marriages show some evidence of waning over time, further underscoring the increased effects from events occurring in the later age range. Thus, temporal proximity does matter to a certain extent, but generally effects are persistent over time. This corroborates the idea that the later age is a more sensitive period for event occurrence. Again, this sensitivity analysis is only possible for CES-D scores, not self-rated health outcomes, so these conclusions are drawn with caution.

Effect trends across ages

Though the distinction between early and later adulthood as being 18–25 versus 26–40 is established in literature (Kasl & Jones, 2000), it is nonetheless useful for comparison to look at trends across the full age range (18–40 years old). Thus, I run separate logistic regression models for event variables at each age. I include the same control variables as in previous models, as well as controls for having ever experienced other events between ages 18 and 40. I graph the coefficients across ages and force a linear trend line to see potential patterns.

Though standard errors are large for these estimates, given the small sample sizes of those who experience events, and therefore the values of estimates are independently not particularly informative, the pattern of effects can easily be seen in figures 1a and 1b. Results indicate that the pattern of effects from the previous analysis is upheld. For self-rated health (figure 1a), the effects of employment losses become more negative across age. Employment gains become more positive. Effects of marriage and divorce both improve slightly over time. For depression (figure 1b), both employment losses and divorce lead to increased effects on depression across age, whereas employment gains and marriages show slight decreases.

Discussion

The analyses presented in this paper look at employment losses and gains, marriage, and divorce to see how these disruptive transitions affect physical and mental health at age 40. I find that employment losses decrease self-rated health and increase depression. This is true for both men and women, though effects on depression are stronger for women (which complements research showing higher levels of depression among women (Mirowsky, 1996) but counters research showing stronger effects of unemployment on mental health for men (Paul & Moser, 2009)) and effects on self-rated health are stronger for men (following evidence that employment loss increases mortality among men (Morris et al., 1994)). Employment gains, conversely, improve health by decreasing depression and increasing self-rated health (which follows work that shows reversal of the health-damaging effects of job loss by reemployment (Kessler, Turner & House, 1989)). Again, effects on depression are stronger for women and effects on self-rated health are stronger for men. Marriages lead to some decreases in depression for women specifically (which is not surprising; e.g. Kiecolt-Glaser & Newton, 2001); divorces lead to increases in depression in women and decreases in self-rated health for men and women. Thus, employment and partner losses appear to have similar effects as do employment and partner gains.

Beyond gender differences, there are other nuances. First, timing matters: employment transitions lead to worse health outcomes when they occur later in adulthood, whereas marital transitions affect self-rated health more strongly when they occur in early adulthood and depression more strongly when they occur in later adulthood. Second, the numbers of incidences matter: those who experience a larger number of employment transitions have worse health outcomes. For numbers of marital transitions, evidence is more mixed, with respondents who experienced one marriage in the same time span showing greater

Figure 1a.

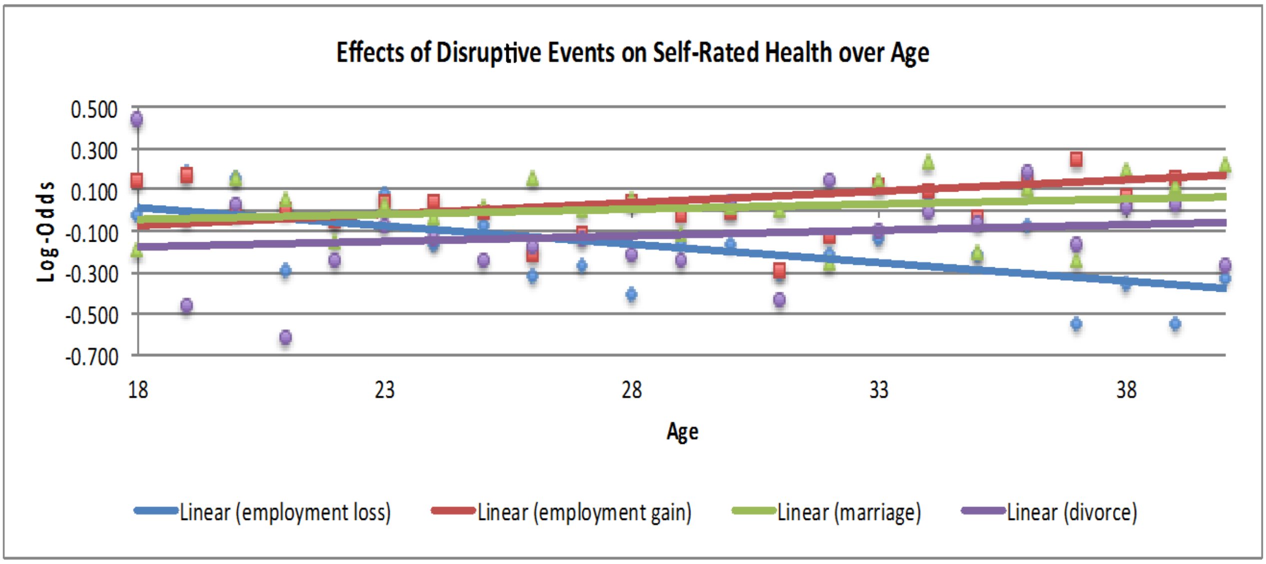
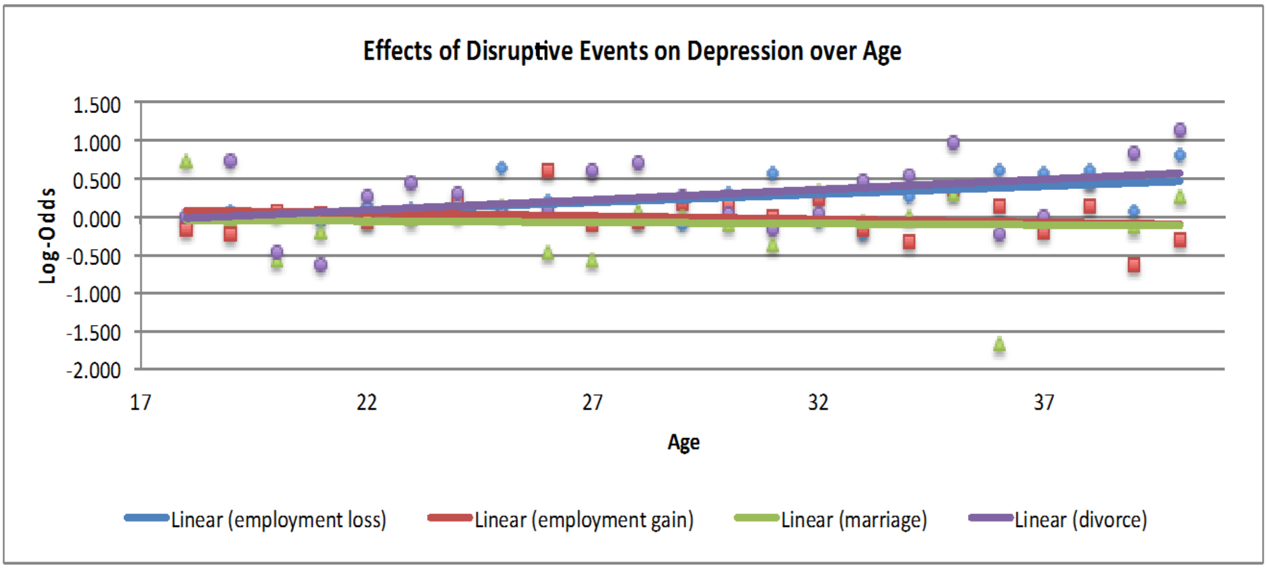


Figure 1b.



health benefits as compared to those who experienced two marriages, while those who experienced two or more divorces in the same time span show lower self-rated health but no differences in depression when comparing those same groups.

This paper aims to bring together findings from two domains of disruptive life events – work and family – as these are transitions that are relatively commonly experienced by many people. Allowing these transitions to co-exist shows the effects of each event within the context of other transitions that might be happening in the life course. Though seemingly positive transitions such as marriage and employment gains show improvements in health, overall, they do not fully 'offset' transitions such as divorce and employment losses, which could go hand-in-hand with the more positive events. Indeed, looking at a simple event count of all four events combined into one variable shows a 0.052 decrease in the log-odds of self-rated health for each additional event and a 0.078 increase in the log-odds of depression for each additional event (full results available upon request).This speaks to literature on the negative effects of instability (e.g. Osborne & McLanahan, 2007): a greater count of events means more instability and worse health outcomes, even when some included events seem like positive transitions. I speculate that that cumulative disadvantage (DiPrete & Eirich, 2006), biological wear and tear (Geronimus, 1996; Hayward et al., 2000; Seeman et al., 2001), and role history (Wheaton, 1990) could be explanations for increased effects with an increased number of transitions, though I do not test these theories directly. Since people often do not experience one disruptive transition in isolation of other transitions, this context is important to understanding how events matter in the life course. A relevant next step might be to characterise combinations of events into pathways (e.g. Eliason, Mortimer & Vuolo, 2015) to see how different pathways alter health outcomes.

Importantly, the results show that there are similarities across work and family domains in terms of the effects that transitions evoke on health outcomes; moreover, experiencing several transitions, within the same or different domains, can lead to important overlaps in effects that might lessen or worsen health outcomes overall. Thus, considering experiences of transitions across work and family domains matters for health outcomes.

Acknowledgements

I thank Jason Beckfield, Lisa Berkman, and Subu Subramanian for helpful comments and suggestions.

References

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist*, 49(1), 15–24. https://doi.org/10.1037/0003-066X.49.1.15

Adler, N. E. & Ostrove, J. M. (1999). Socioeconomic Status and Health: What We Know and What We Don't. *Annals of the New York Academy of Sciences*, *896*, 3–15. https://doi.org/10.1111/j.1749-6632.1999.tb08101.x

Ali, J. & Avison, W. R. (1997). Employment Transitions and Psychological Distress: The Contrasting Experiences of Single and Married Mothers. *Journal of Health and Social Behavior*, *38*, 345–362. https://doi.org/10.2307/2955430

Backhans, M. C. & Hemmingsson. T. (2011). Unemployment and Mental Health – Who Is (Not) Affected? *The European Journal of Public Health*, *22*(3), 429–33. https://doi.org/10.1093/eurpub/ckr059

Bambra, C. (2011). *Work, Worklessness, and the Political Economy of Health*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199588299.001.0001

Bambra, C. (2014). Health Inequalities, Work, and Welfare. *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society*. Pp. 989–992.

Bartley, M., Ferrie, J., & Montgomery, S. M. (2006). Health and Labour Market Disadvantage: Unemployment, Non-Employment, and Job Insecurity. In M. Marmot and R. G. Wilkinson (Eds.), *Social Determinants of Health* (78–96). Oxford: Oxford University Press.

Ben-Shlomo, Y. & Kuh, D. (2002). A Life Course Approach to Chronic Disease Epidemiology: Conceptual Models, Empirical Challenges and Interdisciplinary Perspectives. *International Journal of Epidemiology*, *31*(2), 285–93. https://doi.org/10.1093/intjepid/31.2.285

Bird, C. E. & Rieker, P. P. (1999). Gender Matters: An Integrated Model for Understanding Men's and Women's Health. *Social Science and Medicine*, *48*, 745–755. https://doi.org/10.1016/S0277-9536(98)00402-X

Blane, D. (2006). The Life Course, the Social Gradient, and Health. In M. Marmot and R. G. Wilkinson (Eds.), *Social Determinants of Health* (54–77). Oxford: Oxford University Press.

Blekesaune, M. (2008). Partnership Transitions and Mental Distress: Investigating Temporal Order. *Journal of Marriage and Family*, *70*, 879–890. https://doi.org/10.1111/j.1741-3737.2008.00533.x

Booth, A. & Amato, P. R. (2001). Parental Predivorce Relations and Offspring Postdivorce Well-Being. *Journal of Marriage and Family*, *63*,197–212. https://doi.org/10.1111/j.1741-3737.2001.00197.x

Brand, J. E. & Simon Thomas, J. (2013). Causal Effect Heterogeneity. In S. L. Morgan (Ed.), *Handbook of Causal Analysis for Social Research* (189–213). Dordrecht: Springer Publishing.

Bronfenbrenner, U. (2004). *Making Human Beings Human: Bioecological Perspectives on Human Development*. Thousand Oaks, CA: SAGE Publications, Inc.

Bureau of Labor Statistics, U.S. Department of Labor. *National Longitudinal Survey of Youth 1979 cohort, 1979–2012 (rounds 1–25)*. (2012). Produced and distributed by the Center for Human Resource Research, Ohio State University, Columbus, OH.

Bureau of Labor Statistics, U.S. Department of Labor. *Sample Weights and Clustering Adjustments*. 2016. Retrieved from https://www.nlsinfo.org/ content/cohorts/nlsy79/using- and-understanding-the-data/ sample-weights-clustering- adjustments on May 5, 2016.

Burgard, S. A., Brand, J. E. & House, J. S. (2009). Perceived Job Insecurity and Worker Health in the United States. *Social Science & Medicine*, *69*(5), 777–85. https://doi.org/10.1016/j.socscimed.2009.06.029

Caliendo, M. & Kopeinig, S. (2005). Some Practical Guidance for the Implementation of Propensity Score Matching. IZA (Institute for the Study of Labor) Discussion Paper No. 1588.

Case, A. & Paxson, C. (2005). Sex Differences in Morbidity and Mortality. *Demography*, *42*(2), 189–214. https://doi.org/10.1353/dem.2005.0011

Charles, K. K., & Stephens Jr., M. (2004). Job Displacement, Disability, and Divorce*. Journal of Labor Economics*, *22*, 489–522. https://doi.org/10.1086/381258

Cheung, Y. B. & Sloggett, A. (1998). Health and Adverse Selection Into Marriage. *Public Health*, *112*(5), 309–11. https://doi.org/10.1038/sj.ph.1900491

Coyne, J. C. & DeLongis, A. (1986). Going Beyond Social Support: The Role of Social Relationships in Adaptation. *Journal of Consulting and Clinical Psychology*, *54*(4), 454–460. https://doi.org/10.1037/0022-006X.54.4.454

Crimmins, E. M. & Saito, Y. (2001). Trends in Healthy Life Expectancy in the United States, 1970–1990: Gender, Racial, and Educational Differences. *Social Science & Medicine*, *52*, 1629–1641. https://doi.org/10.1016/S0277-9536(00)00273-2

Cutler, D. & Lleras-Muney, A. (2010). Understanding Differences in Health Behaviors by Education. *Journal of Health Economics*, *29*, 1–28. https://doi.org/10.1016/j.jhealeco.2009.10.003

DiPrete, T. A. & Eirich, G. M. (2006). Cumulative Advantage as a Mechanism for Inequality: a Review of Theoretical and Empirical Developments. *Annual Review of Sociology*, *32*(1), 271–97. https://doi.org/10.1146/annurev.soc.32.061604.123127

Elder, G. H. (1998). The Life Course as Developmental Theory. *Child Development*, *69*(1), 1–12. https://doi.org/10.1111/j.1467-8624.1998.tb06128.x

Elder, G. H., Johnson, M. K., & Crosnoe, R. (2003). The Emergence and Development of Life Course Theory. In J.T. Mortimer & M. Shanahan (Eds.), *Handbook of the Life Course*, (pp. 3–19). New York: Springer Science and Business Media. https://doi.org/10.1007/978-0-306-48247-2_1

Eliason, S. R., Mortimer, J. T. & Vuolo, M. (2015). The Transition to Adulthood: Life Course Structures and Subjective Perceptions. *Social Psychology Quarterly*, *78*(3), 205–27.

Evans, G. W., Li, D. & Whipple, S. S. (2013). Cumulative Risk and Child Development. *Psychological Bulletin*, *139*(6), 1342–96. https://doi.org/10.1037/a0031808
Ferrie, J. (2001). Is Job Insecurity Harmful to Health? *Journal of the Royal Society of Medicine*, *94*, 71–76. https://doi.org/10.1177/014107680109400206
Ferrie, J., Shipley, M. J., Marmot, M., Stansfeld, S. & Davey Smith, G. (1998). The Health Effects of Major Organisational Change and Job Insecurity. *Social Science & Medicine*, *46*(2), 243–54. https://doi.org/10.1016/S0277-9536(97)00158-5
Frech, A. & Damaske, S. (2012). The Relationship between Mothers' Work Pathways and Physical and Mental Health. *Journal of Health and Social Behavior*, *53*(4), 396–412. https://doi.org/10.1177/0022146512453929
Frech, A. & Williams, K. (2007). Depression and the Psychological Benefits of Entering Marriage. *Journal of Health and Social Behavior*, *48*(2), 149–63. https://doi.org/10.1177/002214650704800204
Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S. & Aldridge, M. D. (2014). Methods for Constructing and Assessing Propensity Scores. *Health Services Research*, *49*(5), 1701–1720. https://doi.org/10.1111/1475-6773.12182
Geronimus, A. T. (1996). Black/White Differences in the Relationship of Maternal Age to Birthweight: A Population-Based Test of the Weathering Hypothesis. *Social Science and Medicine*, *42*(4), 589–597. https://doi.org/10.1016/0277-9536(95)00159-X
Goldstein, J. R. & Kenney, C. T. (2001). Marriage Delayed or Marriage Forgone? New Cohort Forecasts of First Marriage for U.S. Women. *American Sociological Review*, *66*(4), 506–519. https://doi.org/10.2307/3088920
Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity Score Techniques and the Assessment of Measured Covariate Balance to Test Causal Associations in Psychological Research. *Psychological Methods*, *15*(3), 234–249. https://doi.org/10.1037/a0019623
Hawkins, D. N. & Booth, A. (2005). Unhappily Ever After: Effects of Long-Term, Low-Quality Marriages on Well-Being. *Social Forces*, *84*(1), 451–71. https://doi.org/10.1353/sof.2005.0103
Hayward, M. D., Gorman, B. K. (2004). The Long Arm of Childhood: the Influence of Early-Life Social Conditions on Men's Mortality. *Demography*, *41*(1), 87–107. https://doi.org/10.1353/dem.2004.0005
Hayward, M. D., Miles, T. P., Crimmins, E. M. & Yang, Y. (2000). The Significance of Socioeconomic Status in Explaining the Racial Gap in Chronic Health Conditions. *American Sociological Review*, *65*(6), 910–930. https://doi.org/10.2307/2657519
Ho, D. E., Imai, K., King, G., & Stuart, E. (2007). Matching As Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*, *15*(3), 199–236. https://doi.org/10.1093/pan/mpl013
Hughes, M. E. & Waite, L. J. (2009). Marital Biography and Health at Mid-Life. *Journal of Health and Social Behavior*, *50*(3), 344–58. https://doi.org/10.1177/002214650905000307
Idler, E. L. & Benyamini, Y. (1997). Self-Rated Health and Mortality: a Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior*, *38*(1), 21–37. https://doi.org/10.2307/2955359
Karsten, P. & Moser, K. (2009). Unemployment Impairs Mental Health: Meta-Analyses. *Journal of Vocational Behavior*, *74*(3), 264–82. https://doi.org/10.1016/j.jvb.2009.01.001
Kasl, S. V. & Jones, B. A. (2000). The Impact of Job Loss and Retirement on Health. In L. F. Berkman & I. Kawachi (Eds.), *Social Epidemiology* (118–36).
Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., Zhang, J., Rathbun, A., Wilkinson-Flicker, S., Diliberti, M., Barmer, A., Bullock Mann, F. & Dunlop Velez, E. (2016). *The Condition of Education 2016 (NCES 2016–144).* U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved December 6, 2016 from http://nces.ed.gov/pubsearch
Kessler, R. C., House, J. S. & Turner, J. B. (1987). Unemployment and Health in a Community Sample. *Journal of Health and Social Behavior*, *28*(1), 51–59.

Kessler, R. C., Turner, J. B. & House, J. S. (1989). Unemployment, Reemployment, and Emotional Functioning in a Community Sample. *American Sociological Review*, *54*(4), 648–57. https://doi.org/10.2307/2137140

Kiecolt-Glaser, J. K. & Newton, T. L. (2001). Marriage and Health: His and Hers. *Psychological Bulletin*, *127*(4), 472–503. https://doi.org/10.1037/0033-2909.127.4.472

Koball, H. L., Moiduddin, E., Henderson, J., Goesling, B. & Besculides, M. (2010). What Do We Know About the Link Between Marriage and Health? *Journal of Family Issues*, *31*(8), 1019–40. https://doi.org/10.1177/0192513X10365834

Krieger, N., Rowley, D. L., Hermann, A. A., Avery, B. & Phillips, M. T. (1993). Racism, sexism, and social class: implications for studies of health, disease, and well-being. *American Journal of Preventative Medicine*, *9*(suppl), 82–122. https://doi.org/10.1016/S0749-3797(18)30666-4

Lerner, D., Adler, D. A., Chang, H., Lapitsky, L., Hood, M. Y., Perissinotto, C., Reed, J., McLaughlin, T. J., Berndt, E. R., Rogers, W. H. (2004). Unemployment, Job Retention, and Productivity Loss Among Employees with Depression. *Psychiatric Services*, *55*(12), 1371–78. https://doi.org/10.1176/appi.ps.55.12.1371

Link, B. G. and Phelan, J. (1995). Social Conditions as Fundamental Causes of Disease. *Journal of Health and Social Behavior* (The State of the Art and Directions for the Future), 80–94. https://doi.org/10.2307/2626958

Liu, H. & Umberson, D.J. (2008). The Times They Are a Changin': Marital Status and Health Differentials From 1972 to 2003. *Journal of Health and Social Behavior*, *49*(3), 239–53. https://doi.org/10.1177/002214650804900301

Lowe, S. R., Rhodes, J. E. & Waters, M. C. (2015). Understanding Resilience and Other Trajectories of Psychological Distress: a Mixed-Methods Study of Low-Income Mothers Who Survived Hurricane Katrina. *Current Psychology*, *34*(3), 537–50. https://doi.org/10.1007/s12144-015-9362-6

Marmot, M. (2005). *The Status Syndrome: How Social Standing Affects Our Health and Longevity*. New York, NY: Owl Books.

Masten, A. S. & Monn, A. R. (2015). Child and Family Resilience: a Call for Integrated Science, Practice, and Professional Training. *Family Relations*, *64*(1), 5–21. https://doi.org/10.1111/fare.12103

Masten, A. S. (2013). Risk and Resilience in Development. In P. D. Zelazo (Eds.), *The Oxford Handbook of Developmental Psychology* (pp. 1–31) New York, NY: Oxford University Press.

Masten, A. S. (2014). Global Perspectives on Resilience in Children and Youth. *Child Development*, *85*(1), 6–20. https://doi.org/10.1111/cdev.12205

Mayer, K. U. (2009). New Directions in Life Course Research. *Annual Review of Sociology*, *35*, 413–433. https://doi.org/10.1146/annurev.soc.34.040507.134619

McLeod, J. D. & Almazan. E. P. (2004). Connections Between Childhood and Adulthood. In J.T. Mortimer & M. Shanahan (Eds.), *Handbook of the Life Course* (pp. 391–411). New York, NY: Springer Publishing.

Mirowsky, J. (1996). Age and the Gender Gap in Depression. *Journal of Health and Social Behavior*, *37*, 362–380. https://doi.org/10.2307/2137263

Moen, P. & Chermack, K. (2005). Gender Disparities in Health: Strategic Selection, Careers, and Cycles of Control. *Journal of Gerontology*, *60B*, 99–108. https://doi.org/10.1093/geronb/60.Special_Issue_2.S99

Morgan, S. & Winship, C. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research.* Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511804564

Morris, J. K., Cook, D. G. & Shaper, A. G. (1994). Loss of Employment and Mortality. *British Medical Journal*, *308*, 1135–39. https://doi.org/10.1136/bmj.308.6937.1135

Nathanson, C. A. (1980). Social Roles and Health Status Among Women: The Significance of Employment. *Social Science and Medicine*, *14A*, 463–471.

Nazroo, J. Y. & Williams, D. R. (2006). The Social Determination of Ethnic/Racial Inequalities in Health. In M. Marmot & R. G. Wilkinson (Eds.), *Social Determinants of Health* (pp. 238–266) Oxford: Oxford University Press.

Noelke, C. & Beckfield, J. (2014). Recessions, Job Loss, and Mortality Among Older US Adults. *American Journal of Public Health*, *104*(11), 126–34. https://doi.org/10.2105/AJPH.2014.302210
Osborne, C. & McLanahan, S. (2007). Partnership Instability and Child Well-Being. *Journal of Marriage and Family*, *69*,1065–83. https://doi.org/10.1111/j.1741-3737.2007.00431.x
Palloni, A. (2006). Reproducing Inequalities: Luck, Wallets, and the Enduring Effects of Childhood Health. *Demography*, *43*(4), 587–615. https://doi.org/10.1353/dem.2006.0036
Paul, K., & Moser, K. (2009). "Unemployment Impairs Mental Health: Meta-Analyses." *Journal of Vocational Behavior*, 74(3), 264-282. https://doi.org/10.1016/j.jvb.2009.01.001
Prigerson, H. G., Meciejewski, P. K., & Rosenheck, R. A. (1999). The Effects of Marital Dissolution and Marital Quality on Health and Health Service Use Among Women. *Medical Care*, *37*(9), 858–73. https://doi.org/10.1097/00005650-199909000-00003
Repetti, R. L., Taylor, S. E. & Seema T. E. (2002). Risky Families: Family Social Environments and the Mental and Physical Health of Offspring. *Psychological Bulletin*, *128*(2), 330–66. https://doi.org/10.1037/0033-2909.128.2.330
Riley, M. W., Riley, J. W. Jr., (1994). Stuctural lag: Past and future. In M. W. Riley, R. L. Kahn, & A. Foner, (Eds.), *Age and structural lag: Society's failure to provide meaningful opportunities in work, family and leisure* (pp. 15–36). New York: Wiley.
Sampson, R. J., Sharkey, P., & Raudenbush, S. W. 2008. Durable Effects of Concentrated Disadvantage on Verbal Ability Among African-American Children. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(3), 845–52. https://doi.org/10.1073/pnas.0710189104
Sayer, L. C., England, P., Allison, P., & Kangas, N. (2011). She Left, He Left: How Employment and Satisfaction Affect Men's and Women's Decisions to Leave Marriages. *American Journal of Sociology*, *116*(6), 1982–2018. https://doi.org/10.1086/658173
Schnittker, J. (2004). Education and the Changing Shape of the Income Gradient in Health. *Journal of Health and Social Behavior*, *45*(3), 286–305. https://doi.org/10.1177/002214650404500304
Seeman, T. E. & Crimmins, E. (2001). Social Environment Effects on Health and Aging. *Annals of the New York Academy of Sciences*, 88–117.
Seeman, T. E., McEwen, B. S., Rowe, J. W., & Singer, B. H. (2001). Allostatic Load as a Marker of Cumulative Biological Risk: MacArthur Studies of Successful Aging. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(8), 4770–75. https://doi.org/10.1073/pnas.081072698
Sleskova, M., Salonna, F., Geckova, A.M., Nagyova, I., Stewart, R.E., van Dijk, J.P., & Groothoff, J.W. (2006). Does Parental Unemployment Affect Adolescents' Health? *Journal of Adolescent Health*, *38*(5), 527–35. https://doi.org/10.1016/j.jadohealth.2005.03.021
StataCorp (2013). *Stata: Release 13.* Statistical Software. College Station, TX: StataCorp LP.
Turner, R. J., Thomas, C. S., & Brown, T. H. (2016). Childhood Adversity and Adult Health: Evaluating Intervening Mechanisms. *Social Science and Medicine*, *156*, 114–124. https://doi.org/10.1016/j.socscimed.2016.02.026
Uecker, J. E. 2012. Marriage and Mental Health Among Young Adults. *Journal of Health and Social Behavior*, *53*(1), 67–83. https://doi.org/10.1177/0022146511419206
Uhlenberg, P. & Mueller, M. M. (2004). Family Context and Individual Well-Being. In J.T. Mortimer & M. Shanahan (Eds.), *Handbook of the Life Course* (pp 391–411). New York, NY: Springer Publishing.
Waldron, I., Hughes, M. E., & Brooks, T. L. (1996). Marriage Protection and Marriage Selection – Prospective Evidence for Reciprocal Effects of Marital Status and Health. *Social Science & Medicine, 43*(1), 113–23. https://doi.org/10.1016/0277-9536(95)00347-9
Wheaton, B. (1990). Life Transitions, Role Histories, and Mental Health. *American Sociological Review*, *55*(2), 209–23. https://doi.org/10.2307/2095627
Wheaton, B. (1997). The Nature of Chronic Stress. In B. H. Gottlieb (Ed.), *Coping with Chronic Stress* (pp. 43–73). New York, NY: Springer Publishing. https://doi.org/10.1007/978-1-4757-9862-3_2
Wheaton, B. & Gotlib, I. H. (1997). Trajectories and Turning Points Over the Life Course: Concepts and Themes. In I.H. Gotlib & B. Wheaton (Eds.), *Stress and Adversity over the Life Course* (pp. 1–28). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511527623.001

Wooldridge, J. M. (2006). *Introductory Econometrics: A Modern Approach.* The Thomson Corporation: Thomson South-Western.

Xie, Y., Brand, J. E., & Jann, B. (2012). Estimating Heterogeneous Treatment Effects with Observational Data. *Sociological Methodology*, *42*(1), 314–47. https://doi.org/10.1177/0081175012452652

Endnotes

1. Center for Epidemiologic Studies Depression Scale
2. Short Form Health Survey scale

Appendix

Appendix A: The Effects of Disruptive Events (at ages 40-50) on Depression at age 50, logistic models

Event	All respondents	Men	Women
Employment losses	0.671 †	1.103 †	0.475
Employment gains	-0.706 †	-1.218	-0.704
Marriage	-0.121	0.139	-0.292
Divorce	0.673 †	-0.056	1.106 *

Notes: Results are log-odds. All models include controls for gender, race, education, ability, number of children, welfare receipt, health limitations on work, Rotter Locus of Control score, and Rosenberg Self-Esteem Scale score.

† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

Appendix B: The Effects of Disruptive Events (at ages 40-50) on Self-Rated Health at age 50, logistic models

Event	All respondents	Men	Women
Employment losses	-0.794 ***	-0.924 **	-0.697 **
Employment gains	0.312	0.078	0.501 †
Marriage	0.161	0.318	0.063
Divorce	0.109	0.580	-0.241

Notes: Results are log-odds. All models include controls for gender, race, education, ability, number of children, welfare receipt, health limitations on work, Rotter Locus of Control score, and Rosenberg Self-Esteem Scale score.

† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

Appendix C: The Effects of Disruptive Events (at ages 30-40) on Depression at age 40, logistic models

Event	All respondents	Men	Women
Employment losses	0.758 **	0.752 †	0.761 *
Employment gains	-0.153	-0.457	0.130
Marriage	-0.196	0.062	-0.323
Divorce	0.273	0.784 †	-0.004

Notes: Results are log-odds. All models include controls for gender, race, education, ability, number of children, welfare receipt, health limitations on work, Rotter Locus of Control score, and Rosenberg Self-Esteem Scale score.

† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

Appendix D: The Effects of Disruptive Events on Health Outcomes, Using IPTW Models

	Full sample (ages 18-40)		Ages 40-50		Ages 30-40
Event	Effects on Depression	Effects on Self-Rated Health	Effects on Depression	Effects on Self-Rated Health	Effects on Depression
Employment loss	0.406	-0.370 **	0.775 *	-0.747 ***	0.628 *
Employment gain	-0.171	0.437 *	-0.212	0.099	-0.199
Marriage	-0.839 *	0.165	-0.275	0.211	-0.228
Divorce	0.792 **	-0.174 †	0.544	0.144	0.257

Notes: Results are log-odds. All models include controls for other events (ever during the corresponding time frame) gender, race, education, ability, number of children, welfare receipt, health limitations on work, and Rotter Locus of Control score. Models are weighted using IPTW.
† p<.10 * p <.05 ** p < .01 *** p < .001 (two-tailed tests)

Appendix E: The Effects of Disruptive Events (at ages 18-25) on Depression at age 30

Event		**All respondents** ***18-25***
Employment losses		
	1x	0.121
	2x	0.285 †
	3x	0.782 **
Employment gains		
	1x	-0.110
	2x	0.125
	3x	-0.438
Marriage		
	1x	-0.120 *
	2x	-0.515 †
Divorce		
	1x	0.402 **
	2x	0.918 *
Constant		-0.201
LR χ^2		288.00
P > χ^2		0.000
Pseudo R^2		0.076
n		3,696

Notes: Results are log-odds. All models include controls for gender, race, education, ability, number of children, welfare receipt, health limitations on work, Rotter Locus of Control score, and Rosenberg Self-Esteem Scale score.
† p<.10 * p <.05 ** p < .01
*** p < .001 (two-tailed tests)

Appendix F: Covariate Balance for Inverse Probability of Treatment Weights

Covariates	Employment Loss			Employment Gain			Marriage			Divorce		
	Mean in Treated	Mean in Untreated	Stand. Diff.	Mean in Treated	Mean in Untreated	Stand. Diff.	Mean in Treated	Mean in Untreated	Stand. Diff.	Mean in Treated	Mean in Untreated	Stand. Diff.
Employment Loss (ever, by age 40)	-	-	-	0.72	0.74	-0.051	0.72	0.70	0.040	0.71	0.72	-0.009
Employment Gain (ever, by age 40)	0.83	0.82	0.024	-	-	-	0.82	0.76	0.151	0.79	0.81	-0.060
Marriage (ever, by age 40)	0.73	0.72	0.028	0.73	0.73	-0.007	-	-	-	0.70	0.73	-0.062
Divorce (ever, by age 40)	0.36	0.36	0.011	0.36	0.34	0.040	0.36	0.57	-0.525	-	-	-
Male (0/1)	0.46	0.46	-0.001	0.48	0.43	0.103	0.48	0.46	0.052	0.46	0.49	-0.045
Black (0/1)	0.31	0.32	-0.023	0.31	0.34	-0.072	0.30	0.21	0.206	0.25	0.31	-0.121
Hispanic (0/1)	0.19	0.19	0.014	0.19	0.19	0.009	0.19	0.33	-0.339	0.22	0.19	0.064
High school only (0/1; by age 40)	0.43	0.43	0.006	0.43	0.40	0.054	0.43	0.47	-0.072	0.43	0.43	-0.012
College attendance (0/1; by age 40)	0.24	0.23	0.005	0.24	0.25	-0.024	0.24	0.24	-0.003	0.25	0.24	0.020
College completion (0/1; by age 40)	0.22	0.22	0.017	0.22	0.24	-0.067	0.21	0.18	0.080	0.22	0.21	0.006
Ability (-3 - 3)	0.00	-0.01	0.019	0.03	-0.04	0.091	0.02	0.09	-0.092	0.07	0.03	0.055
Number of children (by age 40)	1.92	1.95	-0.019	1.93	1.91	0.019	1.95	3.03	-0.724	2.15	1.92	0.161
Welfare (between ages 18 and 40)	0.38	0.39	-0.016	0.37	0.44	-0.140	0.38	0.37	0.008	0.35	0.37	-0.038
Health limits (in 1979)	0.05	0.05	-0.002	0.05	0.05	-0.003	0.05	0.06	-0.043	0.05	0.05	0.019
Rotter Locus of Control	8.76	8.78	-0.007	8.75	8.93	-0.075	8.73	8.52	0.087	8.57	8.73	-0.064

Notes: Covariates reported here are matched on CESD outcomes. Matches for self-rated health outcomes are closely similar. Rosenberg self-esteem measures are removed from the matching analysis for better balance; model results are not affected by this choice. These are matches generated for models with events during ages 18 to 40.

Longitudinal and Life Course Studies 2018 Volume 9 Issue 4 Pp 433–446 ISSN 1757-9597

A software package for the application of probabilistic anonymisation to sensitive individual-level data: a proof of principle with an example from the ALSPAC birth cohort study

Demetris Avraam University of Cyprus, Cyprus and Newcastle University, UK
Andy Boyd University of Bristol, UK
a.w.boyd@bristol.ac.uk
Harvey Goldstein University of Bristol, UK and University College London, UK
Paul Burton Newcastle University, UK and University of Bristol, UK

(Received September 2017 Revised July 2018) http://dx.doi.org/10.14301/llcs.v9i4.478

Abstract

Individual-level data require protection from unauthorised access to safeguard confidentiality and security of sensitive information. Risks of disclosure are evaluated through privacy risk assessments and are controlled or minimised before data sharing and integration. The evolution from 'Micro Data Laboratory' traditions (i.e. access in controlled physical locations) to 'Open Data' (i.e. sharing individual-level data) drives the development of efficient anonymisation methods and protection controls. Effective anonymisation techniques should increase the uncertainty surrounding re-identification while retaining data utility, allowing informative data analysis. 'Probabilistic anonymisation' is one such technique, which alters the data by addition of random noise. In this paper, we describe the implementation of one probabilistic anonymisation technique into an operational software written in R and we demonstrate its applicability through application to analysis of asthma-related data from the ALSPAC cohort study. The software is designed to be used by data managers and users without the requirement of advanced statistical knowledge.

Keywords

Probabilistic anonymisation; disclosure control; measurement error; *h*-rank index; ALSPAC

Introduction

Data custodians managing longitudinal study data resources use a variety of policies and processes to manage risks to participant confidentiality and data security when sharing data. This can form a means to help meet legal requirements and also a component of wider strategies to retain participant trust and the public acceptability of research (Carter, Laurie, & Dixon-Woods, 2015). Approaches range from: 1) removing directly identifiable information (see panel 1 for term definitions); 2) only providing access to accredited users; 3) allocating (project-specific) pseudo IDs to each subject; 4) making adjustments to outlying values and small cell counts; 5) sub-setting datasets to only include data required for specific investigations; 6) transforming data through complex statistical processes that mask or block access to the underlying individual-level data; and, 7) sharing and using data within secure policy and procedural frameworks (Elliot, Mackey, O'Hara, &

Tudor, 2016), such as Data Safe Havens (Burton et al., 2015).

The EU General Data Protection Regulation (GDPR) (European Parliament, 2018), through national implementations such as the Data Protection Act 2018 (DPA) (UK Parliament, 2018), distinguishes between personal data and anonymous data. Personal data is defined as "information relating to natural persons who: a) can be identified or who are identifiable, directly from the information in question; or b) who can be indirectly identified from that information in combination with other information". Therefore, personal information, includes data with direct identifier variables or data where identity can be determined through linking to other readily available information. This classification is important as the safeguards required for the use of personal information are far more stringent than the safeguards required for the use of anonymous data. The DPA – even when research exemptions apply – requires that individuals are informed of the use of their personal information, and that the security of the data is maintained through the research process. Furthermore, even when these safeguards are in place, the DPA requires that data are de-identified as soon in the research process as possible – ideally prior to the point when the data are provided to researchers. In contrast, anonymous data do not fall under the scope of the DPA (or GDPR) and are therefore exempt from these requirements.

Under the new DPA 2018 legislation, longitudinal research studies are required (Article 35 of the GDPR) to consider the risks associated with data processing and use. Through conducting 'Data Protection Impact Assessments', data custodians will assess risks (e.g. loss of control of data when sharing with external research users) and will have to implement controls to mitigate these risks (e.g. effectively anonymising the data). Given the pressures to share data, it seems inevitable that DPA 2018 will provide a new impetus for data guardians to explore options for effective disclosure control. As a community, the data guardians of longitudinal studies should work together to understand the options available, the impact these may have on research utility and how to implement anonymisation strategies effectively. The risk of not doing this is that poorly executed anonymisation strategies reveal sensitive information about participants and bring the research community into disrepute. While we are fortunate that there are no known examples of this within the longitudinal research community, we should take note of parallel examples of poor practice (e.g. in 2014, the New York Taxi & Limousine Commission released data on 173 million individual journeys, yet a poor anonymisation strategy meant that individuals could easily be re-identified and their sensitive information breached (Pandurangan, 2014)).

Panel 1: Disclosure Control Terminology

Data Custodians: authorised individuals/entities who manage and share study data. While (typically) authorised to view identifiable data, there is a risk they can accidentally disclose data through data breaches or accidentally and spontaneously identify a participant.

Accredited User: a bona-fide professional working for a bona-fide institution for a bona-fide purpose who can be expected to operate professionally and to not deliberately disclose information. The potential for accidental disclosure remains. Similar to the 'Safe Researcher' concept.

External Attacker: an individual who will attempt to deliberately disclose participant information for malicious means.

Individual of Interest: the participant(s) targeted by an external attacker.

Direct Identifier(s): a data item that on its own could identify an individual (e.g. name, full date of birth, full address, health or other service ID number). The GDPR/DPA 2018 has expanded the legal definition of personal identifiers to include genetic sequence information (when used for linkage) as well as digital network identifiers (such as internet 'IP' addresses).

Indirect-Identifier(s) (aka Quasi-Identifiers): social or health variables with (context-specific) potential to disclose an individual's identity (i.e. they are likely to be known or discoverable to an external attacker or spontaneously recognisable to someone who knows the individual), for example: parity, height, weight, disease status, occupation categories.

Non-Identifier(s): variables with exceptionally limited potential to disclose an individual's identity. These will tend to be transient values (e.g. blood pressure readings).

Achieving anonymity in a dataset is challenging and is complicated by the fact that much population discovery science, particularly that informed by longitudinal studies, relies on broad datasets of granular detailed individual-level data. These data are ideal for assessing life-course associations and controlling for socially mediated status, yet are also ideally suited – given their rich and typically unique patterns of values – for identifying participants' real-world identities. This situation is further complicated by the fact that some indirect identifiers have research value (e.g. age, gender), so that the different classes of identifiers (direct, indirect, non) often cannot be viewed in isolation and that identification risk is context specific. Existing approaches to controlling for this risk, such as *k*-anonymisation (El Emam & Dankar, 2008; Sweeney, 2002), attempt to mask these patterns of uniqueness through supressing and aggregating data values. While this technique offers some protection to disclosure risk (Domingo-Ferrer, Sebé, & Castellà-Roca, 2004), it also has the potential to impact the utility of the data to inform the research question.

Goldstein and Shlomo (Goldstein & Shlomo, 2018) suggest the use of a probabilistic anonymisation approach to perturb the data through the addition of random noise to some or all variables in the dataset. In this approach, the risks posed by an external 'attacker' who wished to re-identify an individual of interest from a dataset, are assessed. In this risk scenario, it is assumed that the attacker independently knows the individual's data values for some or all the identifying variables within the dataset. Using this information, the attacker could 'link' to the target individual's record using the unique patterns in their data, and therefore learn new information about that person from their associated attribute variables. To avoid such identification, Goldstein and Shlomo propose that sufficient noise is generated and added to the identifying variables to disguise their values as they appear to any attacker. From the research perspective, the accredited user is provided with sufficient information to remove the effects of the noise during the analysis stage to recover the underlying data structure and therefore to produce consistent parameter estimates. This is done through the use of statistical techniques for fitting models with measurement errors (see Goldstein, Browne, & Charlton, 2017).

This paper, in contrast to Goldstein and Shlomo's methodological manuscript, presents a pragmatic perspective with worked examples. To apply Goldstein and Shlomo's methodology, we have written an operational software package using the open-source statistical programming language R. We use data from participants of the Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort study (Boyd et al., 2013) to demonstrate the feasibility and practicality of the approach. For illustration of the method, we anonymise asthma-related data by adding differing degrees of noise. We then perform three exemplar analyses on the differing versions of anonymised data, treating the noise as measurement error. Finally, we assess how well the true model parameters are retrieved and we compare the differing risks of residual disclosure in the different datasets.

Software package

We have developed two functions in R (Avraam, 2018); the function *probAnon(),* which adds noise to an input dataset, and the function *hRanks(),* which generates a re-identification risk measure. Software to carry out the data modelling has been written in MATLAB (Mathworks, 2016).

The *probAnon()* function

The function *probAnon()* applies probabilistic anonymisation to an input dataset. The algorithm first separates the input data into two subset data frames, one for the continuous (numerical) and one for the categorical (integer or factor) variables. Then, normally distributed random noise with user-specified variances is added, independently, to the continuous and categorical variables. For continuous variables, the variance of noise is specified as a percentage of each variable's observed variance in the argument *weights*, which is a vector ($w_1, \ldots, w_s$) of length s, where s is the number of continuous variables in the input dataset. If the user does not specify the vector of weights, each weight is set to 0.1 by default, which means that the variance of the added noise is equal to the 10% of the observed variance of the variable. The random noise added to each binary variable follows a normal distribution with zero mean and variance specified by the user. The added noise therefore, converts binary data to continuous. For the 'noisy' continuous form of 0–1 binary variables, the algorithm then truncates any negative values to

0 and any values greater than 1 to 1. This step is not strictly necessary, particularly since it tends to increase identifiability risk, but may be convenient for presentational purposes and serves in the present context to present a worst case scenario. The output is then the input dataset plus the added noise. In addition, the argument *seed*, allows the user to set a certain random number generator. If this argument is not specified, the function bases the *seed* parameter on the local time (as determined by the computer's internal clock).

The *hRanks()* function

The function *hRanks()* calculates a re-identification risk measure (the *h*-rank index) of anonymised data using the method proposed by Goldstein and Shlomo (Goldstein & Shlomo, 2018). The function takes as input arguments the original dataset and the anonymised dataset (both having the same dimensions). The conceptual basis of the *h*-rank index is to estimate the probability of an attacker being successful in identifying their individual of interest within the anonymised dataset. We assume that a potential attacker will have access to some information about an individual they are targeting (we note that this assumption is also explicit within Data Protection legislation and represents a data guardian's worst case scenario).

We describe the logic of this function here and illustrate this in panel 2. Initially (step 1), the algorithm calculates the Euclidean distances (defined as the square root of the sum of the squares of the differences between the corresponding coordinates of two vectors) between each row in the true dataset and all rows in the noisy dataset (i.e. a pair-wise comparison that ultimately assesses all possible pairs). It then (step 2), ranks the distances to determine how close each true record is to every record in the noisy dataset (i.e. a 1 to *n* comparison where *n* is the total number of records), and identifies the position of the closest record (i.e. the record that corresponds to rank equal to one). We use the standard competition ranking method (where ties are allocated the same rank, and the next allocated rank is offset by the number of ties, e.g. "1224"), which is performed by the R function *rank()* with the argument *ties.method='min'*. In step 3, the algorithm generates a duplicate copy of the true dataset and computes the Euclidean distances between each row in the true dataset with all rows in the copy of the true dataset and ranks them in order of distance (i.e. a 1 to *n* comparison similar to step 2). In step 4, the algorithm identifies the ranks of the distances calculated in step 3 at the locations specified in step 2. Finally (step 5), the algorithm calculates the difference (*h*-rank index) between the ranks located in step 4 and the ranks located in step 2, and returns a vector with those differences. Note that the critical observations to identify in step 2 are all ranked 1 (or a tied equivalent) as the role of step 2 is to search for the closest noisy record to each true record. If *h*=0 for any one record that an attacker has available (and belongs to the dataset), then this implies that the noisy record identified by the attacker as the closest one in terms of the distance metric, is in fact the true one. The average value of *h*-rank indices provides a metric of disclosiveness. The larger the average value of *h*, the greater the level of unreliability in any attempt to disclose identity through exploiting a given individual's known pattern of data values. Where the average of *h* is small (i.e. lower than an acceptable threshold pre-specified by the study data custodian), the *probAnon()* function can be re-used to alter the data with a higher level of noise in order to increase the uncertainty of re-identification.

Some care is needed where there are more than a negligible number of tied distances. This will be a particular issue with categorical, including binary, data. For example, where a dataset consists of only four binary indicators, there are only 16 possible patterns; meaning that for any given record where noise has been added there will be many tied rank distances. For an attacker, when estimating *h*, this will result in additional uncertainty. Thus, for example, if there are *p* tied ranks and the correct true record is among these, the attacker will be confronted with *p* records with *h*=0, and will be able to choose the correct one only with probability 1/*p*. To reflect this so that we can consistently report our risk measure on the scale of *h*, a very small amount of noise is added to each of the identifiers in order to break the ties and so that the true record will therefore be identified as the closest with probability 1/*p.* In the present implementation (for the case that we have a dataset with only categorical variables) we have added, for all categorical (binary) variables, noise following a normal distribution with variance 10^{-8}.

Panel 2: Schematic illustration of *probAnon()* and *hRanks()* algorithms

In this schematic illustration we use a hypothetical dataset containing three variables (var1 – var3) – two are continuous and one binary - relating to six data subjects (1-6). Our aim is to use the *probAnon()* function to generate 'noisy' version of the true records, and then to use *hRanks()* function to determine the strength of the disclosure control through quantifying the similarity of the noisy data to the true data via h.

probAnon()

In this example, the *probAnon()* function adds normal noise with zero mean and variance equal to 0.1% of the true variability to each continuous variable and normal noise with zero mean and variance equal to 0.5 to the binary variable. The binary variable after the addition of noise is returned as continuous where negative values are truncated to 0 and values greater than 1 are truncated to 1.

TRUE RECORDS

	var1	var2	var3
1	9.63	3.84	0
2	10.39	3.69	0
3	9.76	3.95	1
4	9.77	4.21	0
5	9.5	3.61	0
6	9.93	3.35	1

NOISY RECORDS

	var1	var2	var3
1	9.58	3.88	0.28
2	10.37	3.57	0.08
3	9.91	3.89	0.61
4	9.78	4.17	1
5	9.51	3.72	0.35
6	10.1	3.38	0

hRanks()

Step 1: The Euclidean distances between each row in the true dataset and each row in the noisy dataset are calculated (i.e. a pair-wise comparison that ultimately assesses all possible pairs). It then ranks the distances true-record-wise (in Step 2) in order to determine how close each individual's true record is to every noisy record. It is assumed that the potential attacker has access to one or more recorded true values.

Step 2: Ranks (lowest to highest) of distances row-wise. The position of the lowest distance (circled number) indicates which of the noisy records is the closest to each true record.

Distances

TRUE RECORD (rows) / NOISY RECORD (columns)

	1	2	3	4	5	6	→						
1	0.2872	0.7918	0.6731	1.0637	0.3890	0.6576	→	1	5	4	6	2	3
2	0.8778	0.1456	0.8016	1.2659	0.9475	0.4245	→	4	1	3	6	5	2
3	0.7455	1.1674	0.4221	0.2209	0.7334	1.2002	→	4	5	2	1	3	6
4	0.4727	0.8809	0.7029	1.0008	0.6559	0.8932	→	1	4	3	6	2	5
5	0.3971	0.8746	0.7865	1.1798	0.3670	0.6426	→	2	5	4	6	1	3
6	0.9601	1.0433	0.6664	0.8336	0.8578	1.0148	→	4	6	1	2	3	5

Step 3: The Euclidean distances between each row in the true dataset (any one of which may be in the possession of the attacker) and each row in its duplicated copy dataset are calculated (i.e. a pair-wise comparison for all possible pairs). It then ranks the distances row-wise (Step 4).

Step 4: Enclosed numbers are set to the positions identified in Step 2.

TRUE RECORD (rows) / TRUE RECORD (columns)

	1	2	3	4	5	6	→						
1	0	0.7747	1.0144	0.3956	0.2642	1.1533	→	1	4	5	3	2	6
2	0.7747	0	1.2102	0.8092	0.8936	1.1520	→	2	1	6	3	4	5
3	1.0144	1.2102	0	1.0333	1.0877	0.6236	→	3	6	1	4	5	2
4	0.3956	0.8092	1.0333	0	0.6580	1.3286	→	2	4	5	1	3	6
5	0.2642	0.8936	1.0877	0.6580	0	1.1192	→	2	4	5	3	1	6
6	1.1533	1.1520	0.6236	1.3286	1.1192	0	→	5	4	2	6	3	1

Step 5: The h-rank index is calculated as the difference of each circled rank of each individual record, selected in Step 4 and its corresponding (i.e. having the same position) circled rank located in Step 2. The vector of the h-rank indices for these six individual-level records is then $h = (0,0,3,1,0,1)$.

By carrying out the random noise addition many times, we obtain a distribution of h, so that we can estimate the cumulative distributions as given in the examples. We can further, evaluate these distributions for records at different percentiles of the distribution of distances from the centroid of the joint distribution of the identifiers.

Examples

To demonstrate the feasibility of the software, we show its applicability to childhood asthma data from participants in ALSPAC; a longitudinal birth cohort study collecting information of participants' life-course exposures, and health, social and wellbeing outcomes (Boyd et al., 2013). ALSPAC recruited pregnant women living in, and around, the city of Bristol (south-west UK) – who were due to deliver between 01/04/91 and 31/12/92. An initial total of 14,062 live-born children were enrolled. By age 18, the enrolled sample had extended to include 14,775 live-born individuals from 15,247 pregnancies. The assessment in this paper was conducted on a sample of 15,211 participants. Data is collected via questionnaires, study assessment visits, biological and 'omic characterisations and linkage to routine records (see: www.bristol.ac.uk/alspac/researchers/access/ for ALSPAC data dictionary).

Ethical approval for ALSPAC was obtained from the ALSPAC Law and Ethics Committee and the NHS Research Ethics Committees. The variables used in this exemplar application (see table 1) were selected and then reviewed by an ALSPAC data custodian (author AB) using the ALSPAC privacy impact/risk assessment template. This assessment (based on an assumption that a potential attacker had some access to real information about their target) noted that the dataset contained Direct Identifiers (study ID), Indirect Identifiers, Non-Identifiers and Outcome variables (see table 1).

Table 1: Asthma-related variables from the ALSPAC birth cohort study.

Variable identification	Type	Identifier/ Outcome	Missing values*	Explanation
b650	binary	Indirect	2009	ever smoked (completed by mother at 18 weeks of gestation)
kz021	binary	Indirect	517	child's sex
kc362	binary	Indirect	4144	never exposed to passive smoke (completed by mother at 15 months)
kc401	multi-categorical	Indirect	4231	ever breast fed (completed by mother at 15 months)
m2110	binary	Non	7036	there is damp/condensation/ mould in home (completed by mother at 7 years 1 month)
dda_91	binary	Outcome	7053	doctor ever diagnosed asthma (completed by mother at 91 months)
kv1059	multi-categorical	Outcome	7426	child had asthma in past 12 months (completed by mother at 128 months)
height_f8	continuous	Indirect	8028	child's height (cm), (measured by fieldworker at 'focus@8' clinical assessment visit at mean age 103.8 months)
weight_f8	continuous	Indirect	8249	child's weight (kg), (measured by fieldworker at 'focus@8' clinical assessment visit at mean age 103.8 months)
raw_fev1_f8	continuous	Outcome	8301	forced expiratory volume in 1 second, (measured by fieldworker at 'focus@8' clinical assessment visit at mean age 103.8 months)

*The number of missing values includes also the 'not completed', 'don't know' and 'no response' answers.

We conducted a complete case analysis that was restricted to participants with non-missing data on all relevant variables. We calculated the children's body mass index (BMI) using the relationship $\mathrm{BMI} = \mathrm{weight} / (\mathrm{height}/100)^2$. We then created three separate datasets: dataset A with the variables $\mathrm{ASTHMA} = \mathrm{dda\_91}$, $\mathrm{SMOKE} = \mathrm{kc362}$, $\mathrm{BREAST\ FED} = \mathrm{kc401}$ and $\mathrm{MOULD} = \mathrm{m2110}$; dataset B with the variables $\mathrm{ASTHMA} = \mathrm{dda\_91}$, $\mathrm{BMI} = \mathrm{weight\_f8}/(\mathrm{height\_f8})^2$ and $\mathrm{BREAST\ FED} = \mathrm{kc401}$; dataset C with the variables $\mathrm{FEV1} = \mathrm{raw\_fev1\_f8}$, $\mathrm{BMI} = \mathrm{weight\_f8} / (\mathrm{height\_f8})^2$, $\mathrm{SEX} = \mathrm{kz021}$ and $\mathrm{SMOKE} = \mathrm{b650}$. From each of the three datasets we removed any rows with missing values. This results in datasets with 6837, 4975 and 5942 complete records respectively. We then converted the multi-categorical variables (see type of each variable in table 1) to binary data. For the variable kc401 (ever breast fed), we combined together the categories "Yes, no longer" and "Yes, still" and replaced their values with ones while we replaced the values in the category "No, never" with zeros. For the variable kv1059 (child had asthma in past 12 months, completed by mother at 128 months), we combined together the categories "Yes, but did not see a doctor" and "Yes, saw a doctor" and replaced their values with ones while we replaced the values in the category "No, did not have" with zeros. We finally generated 'noisy' datasets for each true dataset (A–C) using the *probAnon()* function. For datasets A and B we did not add noise to the ASTHMA variable, which is used as the response variable in a probit regression model. For all the other variables we added normally distributed noise with zero mean and variance equal to the value shown in table 2. We did not consider the case where noise is added to the response variable where this is binary. This feature is described in Goldstein and Shlomo (Goldstein & Shlomo, 2018) but has not yet been implemented in the analysis software.

Table 2: Variances of the added noise. Note that for binary variables we add different levels of noise.

Dataset	Variable	Type	True variance of variable	Variance of added noise
A	ASTHMA	binary response	0.160	-
	SMOKE	binary covariate	0.231	0.05, 0.1, 0.2, 0.5
	BREAST FED	binary covariate	0.171	0.05, 0.1, 0.2, 0.5
	MOULD	binary covariate	0.244	0.05, 0.1, 0.2, 0.5
B	ASTHMA	binary response	0.109	-
	BMI	continuous covariate	5.512	0.55
	BREAST FED	binary covariate	0.151	0.05, 0.1, 0.2, 0.5
C	FEV1	continuous response	0.069	0.0069
	BMI	continuous covariate	5.828	0.58
	SEX	binary covariate	0.250	0.05, 0.1, 0.2, 0.5
	SMOKE	binary covariate	0.248	0.05, 0.1, 0.2, 0.5

Results

We apply regression models to the true and noisy data of each dataset (A–C) and compare the estimated coefficients. Each regression model is applied to the true data using common functions for generalised linear models (e.g. *glm()* function in R) and to the noisy data using a Bayesian Markov Chain Monte Carlo (MCMC) algorithm that allows the recovery of the original data structure (see description of this procedure in (Goldstein et al., 2017). We have not run full simulations of the data. We note that a true simulation to derive population estimates will require both the generation of a model using assumed population parameters and for each of these generated datasets the further generation of a set of models where the noise is sampled from the assumed noise distribution. Goldstein and Shlomo (Goldstein & Shlomo, 2018) ran simulations with both continuous and binary covariates that pointed to negligible bias for the general procedure.

Dataset A

Dataset A consists only of binary data and without special care will give many tied rank distances. To illustrate this, we present the frequencies of all possible combinations for the true values of dataset A in table 3.

We see from table 3 that the smallest set of identical combinations of identifiers confronting an attacker is 61, and the largest 1691. We have therefore used the procedure of adding additional 'tie-breaking' noise and see that the probabilities for a successful attack are still acceptably small. Table 4 gives the estimates of disclosiveness as expressed in terms of h based on 100 simulations. The big number of simulations is used to demonstrate stable estimates.

To analyse the data from dataset A, we apply a probit regression model where the asthma indicator is regressed on smoking, breastfeeding and presence of mould

$$\text{probit}(ASTHMA) = \beta_0 + \beta_1(SMOKE) + \beta_2(BREAST\ FED) + \beta_3(MOULD). \qquad (1)$$

The estimated coefficients from the analysis, and their standard errors, are shown in table 5. We observe that the estimates of the model applied to noisy data using the procedure that removes the noise are close to the estimates of the model applied to the true data (i.e. an overlap between their confidence intervals).

Table 3. Frequencies for all possible combinations of values for dataset A.

ASTHMA	SMOKE	BREAST FED	MOULD	Frequency
0	0	0	0	380
0	0	0	1	181
0	0	1	0	1691
0	0	1	1	1308
0	1	0	0	357
0	1	0	1	227
0	1	1	0	758
0	1	1	1	567
1	0	0	0	91
1	0	0	1	61
1	0	1	0	348
1	0	1	1	310
1	1	0	0	133

Table 4: Cumulative probabilities of *h* for noisy dataset A, at 10th, 50th and 90th percentiles. No noise is added to the response variable. Noise with variance 0.05 (scenario 1), 0.1 (scenario 2), 0.2 (scenario 3) and 0.5 (scenario 4) is added to all predictors. 100 simulated noise additions used.

Scenario	Percentile	$\mathrm{P}(h=0)$	$\mathrm{P}(h \leq 1)$	$\mathrm{P}(h \leq 2)$	$\mathrm{P}(h \leq 3)$	$\mathrm{P}(h \leq 4)$	$\mathrm{P}(h \leq 5)$
	10%	0.0265	0.0286	0.0307	0.0327	0.0343	0.0369
1	50%	0.0105	0.0128	0.0155	0.0177	0.0200	0.0219
	90%	0.0088	0.0112	0.0139	0.0162	0.0185	0.0208
	10%	0.0230	0.0250	0.0275	0.0301	0.0326	0.0340
2	50%	0.0091	0.0116	0.0141	0.0167	0.0191	0.0213
	90%	0.0069	0.0094	0.0119	0.0143	0.0167	0.0191
	10%	0.0221	0.0239	0.0253	0.0273	0.0281	0.0300
3	50%	0.0084	0.0109	0.0129	0.0150	0.0169	0.0193
	90%	0.0063	0.0087	0.0110	0.0132	0.0151	0.0174
	10%	0.0143	0.0156	0.0165	0.0175	0.0182	0.0189
4	50%	0.0048	0.0063	0.0076	0.0088	0.0102	0.0113
	90%	0.0034	0.0049	0.0063	0.0078	0.0092	0.0104

Table 5: Estimated parameters and their standard errors for dataset A. The first row shows the estimated coefficients of the model applied to the true records and scenarios 1–4 show the estimated coefficients of the model applied to noisy data using procedures to recover the data structure. Note that the response variable was without noise and noise with variance 0.05 (scenario 1), 0.1 (scenario 2), 0.2 (scenario 3) and 0.5 (scenario 4) was added to all predictors. The results in scenarios 1–4 show the means of 50 MCMC simulations.

Scenario	Data	β_0 (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)
	True data	-0.809 (0.043)	0.124 (0.036)	-0.131 (0.042)	0.053 (0.035)
1	Noisy data	-0.817 (0.018)	0.115 (0.032)	-0.114 (0.038)	0.048 (0.031)
2	Noisy data	-0.777 (0.042)	0.115 (0.043)	-0.167 (0.044)	0.048 (0.036)
3	Noisy data	-0.828 (0.031)	0.062 (0.027)	-0.063 (0.026)	0.027 (0.026)
4	Noisy data	-0.842 (0.024)	0.040 (0.019)	-0.027 (0.020)	0.016 (0.019)

Dataset B

Dataset B includes both binary and continuous covariates. We add noise with variance equal to 10% of the true variance to BMI variable and noise with variance 0.05, 0.1, 0.2 and 0.5 to the breastfeeding variable. The cumulative probabilities of *h* based on 100 simulations are shown in table 6. We observe that probabilities of *h* to be less than a certain value are increasing with the increase of noise added to the binary variable (i.e. comparing the values between scenario 1, which refers to noise with variance 0.05, and scenario 4, which refers to noise with variance 0.5 added to the binary breastfeeding variable). We also observe higher values of probabilities at the 10th and 90th percentiles in contrast with the lower values at the median (50th percentile). In addition, the probabilities at the 10th percentile are systematically higher than the probabilities at the 90th percentile, which is related to the slightly right-skewed actual distribution of the continuous BMI.

For dataset B, we apply a probit regression model where the asthma indicator is regressed on BMI and breastfeeding

$$\text{probit}(ASTHMA) = \beta_0 + \beta_1(BMI) + \beta_2(BREAST\ FED). \qquad (2)$$

A comparison of the results derived from the model applied to the true and the noisy data are shown in table 7. Similarly to the results obtained from dataset A, we observe a highly accurate estimation of the model parameters by fitting the model to the noisy data and removing the noise using MCMC procedures.

Table 6: Cumulative probabilities of *h* for noisy dataset B, at 10^{th}, 50^{th} and 90^{th} percentiles. No noise is added to the response variable. Noise with variance 0.55 was added to BMI and noise with variance 0.05 (scenario 1), 0.1 (scenario 2), 0.2 (scenario 3) and 0.5 (scenario 4) was added to breastfeeding variable. 100 simulated noise additions used in computations.

Scenario	Percentile	$\mathrm{P}(h=0)$	$\mathrm{P}(h\leq 1)$	$\mathrm{P}(h\leq 2)$	$\mathrm{P}(h\leq 3)$	$\mathrm{P}(h\leq 4)$	$\mathrm{P}(h\leq 5)$
	10%	0.0071	0.0136	0.0191	0.0246	0.0299	0.0356
1	50%	0.0060	0.0115	0.0168	0.0219	0.0269	0.0319
	90%	0.0070	0.0128	0.0183	0.0234	0.0285	0.0336
	10%	0.0071	0.0136	0.0191	0.0247	0.0300	0.0358
2	50%	0.0059	0.0113	0.0167	0.0217	0.0267	0.0316
	90%	0.0069	0.0126	0.0181	0.0232	0.0283	0.0333
	10%	0.0066	0.0127	0.0178	0.0232	0.0282	0.0335
3	50%	0.0055	0.0106	0.0156	0.0203	0.0249	0.0296
	90%	0.0065	0.0118	0.0170	0.0218	0.0266	0.0314
	10%	0.0056	0.0106	0.0149	0.0196	0.0237	0.0283
4	50%	0.0047	0.0089	0.0131	0.0171	0.0210	0.0250
	90%	0.0055	0.0101	0.0145	0.0186	0.0227	0.0268

Table 7: Estimated parameters and their standard errors for dataset B. The first row shows the estimated coefficients of the model applied to the true records and scenarios 1–4 show the estimated coefficients of the model applied to noisy data where noise with variance 0.55 was added to BMI and noise with variance 0.05 (scenario 1), 0.1 (scenario 2), 0.2 (scenario 3) and 0.5 (scenario 4) was added to the breastfeeding variable. Note that the response variable is without noise. The results in scenarios 1–4 show the means of 50 MCMC simulations.

Scenario	Data	β_0 (SE)	β_1 (SE)	β_2 (SE)
	True data	-1.261 (0.174)	0.010 (0.010)	-0.078 (0.058)
1	Noisy data	-1.256 (0.174)	0.009 (0.010)	-0.072 (0.054)
2	Noisy data	-1.250 (0.170)	0.008 (0.011)	-0.065 (0.070)
3	Noisy data	-1.325 (0.176)	0.011 (0.010)	-0.027 (0.036)
4	Noisy data	-1.330 (0.167)	0.011 (0.010)	-0.014 (0.026)

Dataset C

Probabilistic anonymisation has been also applied to dataset C and the cumulative probabilities for *h* at different percentiles based on 100 simulations are shown in table 8. The difference in dataset C (in contrast with datasets A and B) is the outcome variable, which is continuous instead of binary, and therefore noise is added to the outcome in the same way as the noise is added to any continuous explanatory variables.

For dataset C, the force expiratory volume in 1 second is regressed on BMI, sex and smoking

$$FEV1 = \beta_0 + \beta_1\,(BMI) + \beta_2\,(SEX) + \beta_3(SMOKE) \qquad (3)$$

The estimated coefficients with their standard errors are shown in table 9.

We see from these example analyses that the disclosure risk increases with the number of identifying variables used, but remains acceptable. These results suggest that for similar datasets the amount of noise added could safely be reduced. Nevertheless, when a large number of variables is involved in a dataset, the values of *h*-rank index will be expected to increase and this is clearly an area for further exploration. We also note that, especially for binary variables, estimates derived from the noisy data can have large standard errors and the true estimates from the real data can be very different. The amount of noise added to the binary variables in scenario 4 has a standard deviation $\sqrt{0.5} = 0.71$, which is very large compared to the range (0,1) of the true data and therefore we get a lot of instability (as 25% of observed zeros get wrongly defined to their true category). This suggests that further work is needed for such cases and table 8 suggests that smaller values should produce acceptable values for *h*. We conclude that for even low levels of noise the method is sufficient to effectively anonymise the records, but we show the example of noise added to the binary with variance 0.5 as a warning to data managers on the increase in the loss of utility.

Table 8: Cumulative probabilities of *h* for noisy dataset C, at 10th, 50th and 90th percentiles. Noise with variance 0.0069 was added to the outcome FEV1 variable, noise with variance 0.58 was added to BMI and noise with variance 0.05 (scenario 1), 0.1 (scenario 2), 0.2 (scenario 3) and 0.5 (scenario 4) was added to sex and smoke indicators. 100 simulations used.

Scenario	Percentile	$\mathrm{P}(h=0)$	$\mathrm{P}(h \leq 1)$	$\mathrm{P}(h \leq 2)$	$\mathrm{P}(h \leq 3)$	$\mathrm{P}(h \leq 4)$	$\mathrm{P}(h \leq 5)$
	10%	0.0204	0.0377	0.0522	0.0661	0.0785	0.0896
1	50%	0.0206	0.0371	0.0514	0.0644	0.0761	0.0866
	90%	0.0209	0.0375	0.0522	0.0650	0.0768	0.0874
	10%	0.0194	0.0365	0.0505	0.0648	0.0772	0.0878
2	50%	0.0196	0.0356	0.0496	0.0625	0.0741	0.0845
	90%	0.0200	0.0361	0.0505	0.0631	0.0748	0.0852
	10%	0.0179	0.0332	0.0463	0.0591	0.0708	0.0807
3	50%	0.0180	0.0327	0.0457	0.0575	0.0685	0.0782
	90%	0.0183	0.0332	0.0465	0.0582	0.0691	0.0789
	10%	0.0140	0.0264	0.0365	0.0465	0.0561	0.0641
4	50%	0.0143	0.0258	0.0359	0.0453	0.0540	0.0615
	90%	0.0146	0.0264	0.0369	0.0461	0.0548	0.0625

Table 9: Estimated parameters and their standard errors for dataset C. The first row shows the estimated coefficients of the model applied to the true records and scenarios 1–4 show the estimated coefficients of the model applied to noisy data. Noise with variance 0.0069 was added to the outcome FEV1 variable, noise with variance 0.58 was added to BMI and noise with variance 0.05 (scenario 1), 0.1 (scenario 2), 0.2 (scenario 3) and 0.5 (scenario 4) was added to sex and smoking indicators. The results in scenarios 1–4 show the means of 50 MCMC simulations.

Scenario	Data	β_0 (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)
	True data	1.139 (0.024)	0.029 (0.001)	0.105 (0.007)	-0.005 (0.007)
1	Noisy data	1.145 (0.025)	0.029 (0.001)	0.096 (0.006)	-0.004 (0.006)
2	Noisy data	1.140 (0.026)	0.029 (0.001)	0.051 (0.032)	-0.005 (0.006)
3	Noisy data	1.182 (0.025)	0.028 (0.001)	0.051 (0.005)	-0.004 (0.005)
4	Noisy data	1.194 (0.025)	0.028 (0.001)	0.031 (0.004)	-0.001 (0.004)

Discussion

We have shown how a probabilistic anonymisation procedure can be applied to data management procedures in such a way that disclosure risk is reduced to acceptable levels while retaining the ability to carry out statistical analysis. The analysis conducted on the noisy, anonymous, data suffered some loss of statistical efficiency when compared with analysis on the true data; a consequence of which is enlarged confidence intervals and fewer significant inferences. Where the variance of noise added to the binary covariates is large (i.e. > 0.2), there is likely to be unacceptably high loss of statistical efficiency and for binary data biases may also be introduced. This example illustrates the challenge in balancing disclosure control with retaining data utility.

When considering disclosure risk, data custodians should consider the risk of motivated external attackers, accidental disclosure to authorised users and also the possible consequences of human error. In the first scenario (external attacker) the attacker may be motivated to identify a given individual due to their notoriety (for example an investigative journalist following a story or a researcher illustrating the fallacy of supposed 'anonymity' (Sweeney, 2002)) or out of personal interest. In the second, an accredited user or data custodian may recognise an individual during their legitimate work, and in the third an authorised user may inadvertently release a dataset to a wider than authorised audience through a data breach. In all these scenarios identification of a given data subject is likely to occur through matching known 'real world' information about an individual to equivalent information about the same individual within a dataset. Probabilistic anonymisation helps control for these risks by removing certainty about whether the values being considered in the noisy data are true 'real world' values. The *h*-rank index disclosure measure proposed by Goldstein and Shlomo (Goldstein & Shlomo, 2018), adopts this perspective by seeking to establish how well any single individual is 'hidden amongst the crowd' of the other individuals in the dataset. While this approach seems conceptually appropriate – it has some limitations. We found that, in its current state of development, the *h*-rank index was unable to adequately account for disclosure risk of outliers (this was acknowledged by Goldstein and Shlomo (Goldstein & Shlomo, 2018)) but this can be addressed through suitable pre-processing techniques such as truncating them. It was also unable to account for the disclosure risk of clusters of individuals who all have the same outcome value, i.e. that it is not necessary to identify the Individual of Interest within the cluster if they all have the same outcome of interest. This phenomenon is known elsewhere in the privacy literature (Machanavajjhala, Gehrke, Kifer, & Venkitasubramaniam, 2006) and could be quantified by including *l*-diversity metrics to assess outcome value diversity. Finally we found that the *h*-rank index was also unable to account for the protective benefits of the sample being selected from a wider population (again a point noted by Goldstein and Shlomo (Goldstein & Shlomo, 2018)). While this last point would be difficult to accommodate in a metric, it could be incorporated into the Data Custodians risk assessment process.

In practice, a data custodian, in conjunction with potential accredited users would need to evaluate the risks associated with applying any given amount of noise related to the potential loss of analysis efficiency. In some cases, it may not be desirable to release data into the public domain. We suggest that there are few, if any, situations where some variables of interest would need to be excluded, though this remains an area for further study. However, our findings that large amounts of noise impact model estimates suggest this may limit the application for datasets treated with large amounts of noise, e.g. they may be suitable only for training or data exploration rather than applied research. A more realistic application would be the use of probabilistic anonymisation to applying limited amounts of noise (to protect against spontaneous recognition or contained (i.e. not public) data breaches) and to supply accredited users with these noisy data under controlled 'safe haven' conditions. As such, probabilistic anonymisation will add to the range of tools available to data managers that include manual data transformations (such as outlier suppression), statistical approaches (e.g. synthetic data and *k*-anonymisation) and distributed 'black box' computing approaches (e.g. DataSHIELD (Wilson et al., 2017)). All such approaches involve trade-offs between disclosure control, impact on utility and impact on usability. Probabilistic anonymisation has one clear advantage over some of these approaches (e.g. *k*-anonymisation and synthetic data) in that it allows efficient and accurate data linkage to additional datasets, given that the noisy data can contain ID numbers and the noise can be applied over multiple datasets independently. Further work is needed to assess the extent to which these trade-offs apply in order to help inform the Data Custodian community as to which approach may best fit any given situation.

The software functions developed here are proof of principle rather than fully developed 'commercial grade' software. We have identified that improvements would be needed in the following areas before wider adoption: 1) the code needs to accommodate multi-category categorical variables; 2) missing values are not currently supported, we need to allow for these or to develop alternative approaches (e.g. imputation); 3) the *h*-rank index needs developing (as described above) and further consideration given to accommodating outlying values in a flexible manner.

We have demonstrated that probabilistic anonymisation can be effectively deployed to help control for disclosure risk while producing accurate estimates. We have assumed that the data to be used is for bona-fide research scenarios (i.e. not for releasing data into the public domain) where responsible and verifiable data security measures are in place. Additionally, this concept would be novel to many data custodians who may not have advanced statistical expertise, so that determining the appropriate balance between disclosure risk control and retaining data utility would require training. With the enhancements we have identified the software assessed here could be developed into a fully functional tool for Data Custodians. This software would be a useful tool to help longitudinal studies maintain participant trust and to share data securely and effectively while meeting ever more stringent data protection requirements.

Acknowledgements

We are extremely grateful to the families who took part in ALSPAC, the midwives who helped recruit them, and the whole ALSPAC team (including interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses). The UK Medical Research Council and the Wellcome Trust (102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors (DA developed the software and drafted the methods and results; AB drafted the introduction and discussion; all authors contributed to the final manuscript) who serve as guarantors. This research is part of a collaborative research programme entitled 'Cohorts and Longitudinal Studies Enhancement Resources' (CLOSER). This programme is funded by the UK Economic and Social Research Council and Medical Research Council ES/K000357/1. The funders took no role in the design, execution, analysis or interpretation of the data or in the writing up of the findings. This work also formed part of the methods program addressing the securing of sensitive data under the Connected Health Cities project funded by the Department of Health and Social Care.

References

Avraam, D. (2018). R functionality for application of probabilistic anonymisation. Retrieved from https://github.com/davraam/Probabilistic_Anonymisation/releases/tag/v1.0.0

Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., . . . Davey Smith, G. (2013). Cohort Profile: the 'children of the 90s' - the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology, 42*(1), 111–127. https://doi.org/10.1093/ije/dys064

Burton, P. R., Murtagh, M. J., Boyd, A., Williams, J. B., Dove, E. S., Wallace, S. E., . . . Knoppers, B. M. (2015). Data Safe Havens in health research and healthcare. *Bioinformatics, 31*(20), 3241–3248. https://doi.org/10.1093/bioinformatics/btv279

Carter, P., Laurie, G. T., & Dixon-Woods, M. (2015). The social licence for research: why care.data ran into trouble. *Journal of Medical Ethics, 41*(5), 404–409. https://doi.org/10.1136/medethics-2014-102374

Domingo-Ferrer, J., Sebé, F., & Castellà-Roca, J. (2004). On the Security of Noise Addition for Privacy in Statistical Databases. In J. Domingo-Ferrer & V. Torra (Eds.), *Privacy in Statistical Databases: CASC Project Final Conference, PSD 2004, Barcelona, Spain, June 9–11, 2004. Proceedings* (pp. 149–161). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/b97945

El Emam, K., & Dankar, F. K. (2008). Protecting Privacy Using k-Anonymity. *Journal of the American Medical Informatics Association, 15*(5), 627–637. https://doi.org/10.1197/jamia.M2716

Elliot, M., Mackey, E., O'Hara, K., & Tudor, C. (2016). *The Anonymisation Decision-Making Framework*. Manchester, UK: UKAN.

European Parliament. (2018). General Data Protection Regulation. Retrieved from https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en

Goldstein, H., Browne, W. J., & Charlton, C. (2017). A Bayesian model for measurement and misclassification errors alongside missing data, with an application to higher education participation in Australia. *Journal of Applied Statistics*, 1–14.

Goldstein, H., & Shlomo, N. (2018). *A Probabilistic Procedure for Anonymisation and Analysis of Perturbed Datasets*. Under Review.

Machanavajjhala, A., Gehrke, J., Kifer, D., & Venkitasubramaniam, M. (2006). *L-diversity: privacy beyond k-anonymity.* Paper presented at the 22nd International Conference on Data Engineering (ICDE'06). https://doi.org/10.1109/ICDE.2006.1

Mathworks. (2016). MATLAB. Retrieved from http://uk.mathworks.com/products/matlab/

Pandurangan, V. (2014). On Taxis and Rainbow Tables: Lessons for researchers and governments from NYC's improperly anonymized taxi logs. Retrieved from http://blogs.lse.ac.uk/impactofsocialsciences/2014/07/16/nyc-improperly-anonymized-taxi-logs-pandurangan/

Sweeney, L. (2002). k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10*(5), 557–570. https://doi.org/10.1142/S0218488502001648

UK Parliament. (2018). Data Protection Act. Retrieved from http://www.legislation.gov.uk/ukpga/2018/12/pdfs/ukpga_20180012_en.pdf

Wilson, R. C., Butters, O. W., Avraam, D., Baker, J., Tedds, J. A., Turner, A., . . . Burton, P. R. (2017). DataSHIELD – New Directions and Dimensions. *Data Science Journal, 16*, 21. https://doi.org/10.5334/dsj-2017-021

Longitudinal and Life Course Studies 2018 Volume 9 Issue 4 Pp 447–455 *ISSN 1757-9597*

Sex of older siblings and stress resilience

Scott Montgomery Örebro University, Sweden, Karolinska Institutet, Sweden and University College London, UK

scott.montgomery@oru.se

Cecilia Bergh Örebro University, Sweden

Ruzan Udumyan Örebro University, Sweden

Mats Eriksson Örebro University, Sweden

Katja Fall Örebro University, Sweden

Ayako Hiyoshi Örebro University, Sweden

Received November 2017 Revised September 2018) http://dx.doi.org/10.14301/llcs.v9i4.486

Abstract

The aim was to investigate whether older siblings are associated with development of stress resilience in adolescence and if there are differences by sex of siblings. The study used a Swedish register-based cohort of men (n=664 603) born between 1970 and 1992 who undertook military conscription assessments in adolescence that included a measure of stress resilience: associations were assessed using multinomial logistic regression. Adjusted relative risk ratios (95% confidence intervals) for low stress resilience (n=136 746) compared with high (n=142 581) are 1.33 (1.30, 1.35), 1.65 (1.59, 1.71) and 2.36 (2.18, 2.54) for one, two and three or more male older siblings, compared with none. Equivalent values for female older siblings do not have overlapping confidence intervals with males and are 1.19 (1.17, 1.21), 1.46 (1.40, 1.51) and 1.87 (1.73, 2.03). When the individual male and female siblings are compared directly (one male sibling compared with one female sibling, etc.) and after adjustment, including for cognitive function, there is a statistically significant (p<0.005) greater risk for low stress resilience associated with male siblings. Older male siblings may have greater adverse implications for psychological development, perhaps due to greater demands on familial resources or inter-sibling interactions.

Key words

Siblings; sex; psychological functioning; stress resilience; adolescence

Funding

This study was supported by *Nyckelfonden* and by grants from the UK Economic and Social Research Council (ESRC) to the International Centre for Life Course Studies (grants RES-596-28-0001 and ES/JO19119/1).

Introduction

The ability to cope with stress has consequences for disease risk, as demonstrated by associations of a measure of stress resilience, which was designed to assess suitability for military service in Sweden. To produce the score, the men underwent a psychological assessment of their potential ability to cope with stress, based on whether they could control and channel nervousness, their tolerance of stress and disposition to anxiety (Bergh et al., 2014; Bergh, Udumyan, Fall, Almroth and Montgomery, 2015; Crump, Sundquist, Winkleby and Sundquist, 2016; Hiyoshi et al., 2015; Kennedy et al., 2017). A low value for this measure of stress resilience in adolescence, which is often categorised in three groups, has been associated with a raised risk for a variety of diseases in subsequent adulthood, including type 2 diabetes, cancer, cardiovascular disease, anxiety and depression (Bergh et al., 2014, Bergh, et al. 2015; Crump et al., 2016; Hiyoshi et al., 2015; Kennedy et al., 2017). These studies indicate potentially lifelong health implications associated with stress resilience in adolescence, so it is important to identify precursors of low stress resilience to determine if preventative measures are desirable or feasible. Some types of stressful exposures in childhood have been linked with lower stress resilience (Kennedy et al., 2018) and this paper is concerned with identifying further familial factors in childhood that may influence development of stress resilience.

It has been argued that variation in psychologically relevant exposures makes siblings in the same family notably different from each other (Plomin and Daniels, 2011). Birth order has been linked with development of personality and intelligence, albeit inconsistently (Damian & Roberts, 2015; Rohrer, Egloff & Schmukle, 2015). The direction of association with birth order for aspects of mental health and development is not always consistent and older siblings do not always represent an adverse exposure: presence of older siblings has been reported as being associated with relatively better mental health than having younger siblings (Lawson & Mace, 2010). However, associations with intelligence consistently demonstrate an inverse association with presence of older siblings (Kristensen & Bjerkedal, 2007). While this may be in part due to confounding by socioeconomic circumstances, there does appear to be a genuine effect of social rank (social hierarchy) defined by birth order within the family, such that having older siblings, and thus being lower in the hierarchy, represents a risk for lower intelligence (Kristensen & Bjerkedal, 2007). We suggest that if social rank associated with having older siblings is relevant, then some characteristics of the older siblings may play a role in influencing the psychological development of their younger siblings. What is certain is that siblings have an important influence on development (Sulloway, 1996) and can potentially be a source of stressful exposures through inter-sibling aggression, which can be influenced by characteristics such as the sex of siblings (Tippett & Wolke 2015), therefore with possible implications for development of stress resilience.

The sex of older siblings may be relevant to stressful exposures in childhood, as there is evidence that the sex of children in families influences other outcomes, as one study found mothers were more likely to have heart disease if they had sons rather than daughters, possibly because of a greater domestic burden for the mother associated with having sons as they may help less with domestic tasks (D'Ovidio, d'Errico, Scarinzi & Costa, 2015). Therefore, if the sex of children has implications for maternal health, then the sex of older siblings may also be relevant to the development of their younger brothers and sisters, including for development of stress resilience. A component of the association between stress resilience and subsequent psychiatric disease is explained by cognitive function, indicating some shared risks for low cognitive function and low stress resilience (Hiyoshi et al., 2015). We therefore examined whether associations with stress resilience are independent of assessed cognitive function (to indicate intelligence), particularly as birth order has been linked with intelligence (Kristensen & Bjerkedal, 2007). This was undertaken as a separate step in the analysis as stress resilience and cognitive function may theoretically influence each other.

This study used longitudinal Swedish register data for a large number of males to examine the association of number of older male and female siblings with a measure of stress resilience in late adolescence.

Methods

Swedish registers identified all men born between 1970 and 1992 who were assessed for military conscription (1987 to 2010): conscription was compulsory for the majority of males during this period. A total of 843,291 men were identified and 664,403 (79%) remained after exclusions for missing values, mainly due to exemption from conscription, non-participation in the psychological functioning test or other tests by those deemed unsuitable for military service, and some linkage failures to identify parents. The assessment of stress resilience, is described in greater detail elsewhere (Bergh et al., 2015; Hiyoshi et al., 2015) and was designed to quantify suitability in terms of ability to cope with the stress of military action. It involved a 25–30 minute semi-structured interview performed by licensed psychologists with access to additional background information, including a self-completion questionnaire. The test produced a nine-point score covering four dimensions: mental energy, emotional control, social maturity and active/passive interests. A score of nine indicates high stress resilience, thus a greater ability to cope with stress. We previously identified that the functional form of this measure indicates categorisation into three groups rather than use as a continuous measure, as the least stress resilient category has a disproportionally higher magnitude association with a number of adverse outcomes and therefore, as in previous studies, it is categorised as low (1–3), intermediate (4–6) and high (7–9) (Bergh et al., 2015; Hiyoshi et al., 2015; Kennedy et al., 2017; Kennedy et al., 2018). Cognitive function (intelligence) was assessed by written tests covering inductive ability, linguistic understanding and spatial recognition (Carlstedt, 2000). A nine-point normally distributed score was produced, with a score of nine indicating high cognitive function.

The Multi-Generation Register (Ekbom, 2011) was used to provide information on older siblings and mothers. Parental socioeconomic circumstances nearest in time to the cohort member's birth were characterised using census data (1970–1985) and LISA (LISA, 2017), the *Longitudinal Database of Health, Insurance and Labour Market Studies*, after 1990. The highest-level parental occupation was used to produce a three-category version of the European Socioeconomic Classification (ESeC). Information on dates of birth, death and migration of the cohort members was provided by the Total Population Register (SCB, 2017).

The Uppsala regional ethics committee approved this study.

Statistical analysis

Multinomial regression was used to examine associations with the three-category stress resilience measure as the dependent variable. High stress resilience was used as the reference category, so that relative risk ratios (with 95% confidence intervals) are produced for both intermediate and low stress resilience. In model 1, the associations with stress resilience were assessed separately for number of older male and female siblings; multiple birth (a particular sibling type); mother's age at delivery and ESeC (as markers of cultural and material circumstances). Each of these analyses was adjusted for age at conscription assessment and year of assessment (to tackle potential variation in assessment scores in adolescence by age and period). All of the measures were modelled as categorical, including mother's age at delivery as this measure has a non-linear association with stress resilience so cannot be modelled simply as a continuous variable. In model 2, all of the above measures are included in the same model simultaneously to assess the consequences of mutual adjustment. In model 3, the cognitive function measure was added to a model also adjusted for all of the above measures. Cognitive function was modelled as a continuous variable to provide the most effective adjustment. The inclusion of cognitive function was undertaken in a separate model, as including cognitive function may represent an over-adjustment due to its positive association with stress resilience and because it is hypothesised that stress resilience may in turn influence cognitive function. The cluster function was used to account for multiple cohort members coming from the same family, but did not influence the results at the level of precision presented.

A sensitivity analysis compared male and female older siblings directly with each other in category of number of siblings in separate multinomial regression models with stress resilience as the dependent variable. Having one older male sibling was compared with one older female sibling; two older male siblings was compared with two older female siblings; and three or more older male siblings was compared with three or more older female siblings. These models were adjusted for multiple birth, mother's age at delivery, ESeC, age at conscription assessment, year of assessment and cognitive function.

The analysis was performed using Stata MP version 14.2.

Results

Table 1 shows that male adolescents with lower stress resilience had a larger number of older siblings, had a lower average cognitive function score, lower parental ESeC, mothers who gave birth before age 18 years or mothers that were older than average. There is no notable or consistent association with being part of a multiple birth. Table 2 presents relative risk ratios and 95% confidence intervals for medium and low stress resilience compared with high stress resilience. Having a larger number of older male siblings has a higher magnitude association with low stress resilience than having female older siblings. These results showed a gradient of risk across intermediate and low stress resilience compared with high. There is some attenuation of magnitude of associations after adjustment in model 2, mostly due to inclusion of the ESeC variable. Further adjustment for cognitive function in model 3 resulted in attenuation of magnitude for the estimates, but did not eliminate statistical significance. The confidence intervals for older male and female siblings do not overlap, including after adjustment for the potential confounding factors and cognitive function. The higher magnitude association with low stress resilience for male siblings was further assessed in a sensitivity analysis comparing individual male and female siblings directly (having one male sibling compared with one female sibling; two male siblings compared with two female; and three or more male compared with three or more female). After adjustment for all of the potential confounding factors, including cognitive function, there is a statistically significant ($p<0.005$) greater risk for low stress resilience associated with older male siblings compared with female siblings, for each of the three comparisons (data not shown).

Low parental ESeC and having either mothers who were older or younger than average were statistically significantly associated with low stress resilience in all of the models. Higher cognitive function is associated with a statistically significant reduced risk of having contemporaneous low stress resilience. The Pearson correlation coefficient for these two measures is 0.368 ($p<0.001$) with covariance of 1.179. The magnitude of the association between stress resilience and cognitive function is only slightly reduced by adjustment for the other measures in model 3.

Discussion

Having a larger number of older male siblings was associated with lower stress resilience than having the equivalent number of female siblings, independent of measures of socioeconomic circumstances of the family and cognitive function. Having parents in the low ESeC category of the family of origin, indicating more adverse socioeconomic characteristics, is also a risk for low stress resilience.

To the best of our knowledge, this is the first study investigating associations of sex of older siblings with stress resilience, which was measured systematically in adolescence among a large and generally representative population of males. While there is good evidence that siblings can influence personality and other aspects of mental development, (Plomin and Daniels, 2011; Sulloway, 1996) the existing literature does not clearly predict the pattern of association with stress resilience observed here. We believe the association with siblings is due in part to exposure to psychosocial and other forms of stress. It has been suggested that first-born children may be more fearful, even if they have a greater tendency to be 'intellectually oriented' (Eisenman, 1992) and being a younger sibling may result in relatively better mental health (Lawson & Mace, 2010). Both of these studies would imply a greater risk of low stress resilience for first-borns, who may be different in several ways from other siblings. They spend more time alone with parents than younger siblings (Eisenman, 1992) and are more likely to accept the authority of parents than subsequent children (Sulloway, 1996). Other potentially contradictory aspects of having siblings are that, on one hand, sibling relationships offer protection from the effects of stressful life events (Gass, Jenkins & Dunn, 2007) but, on the other, there can be aggression and bullying between siblings that can be a significant source of adverse exposures in childhood (Tippett & Wolke, 2015).

The association of older siblings with development of stress resilience is in part explained by socioeconomic circumstances as signalled by the influence of adjustment for parents' ESeC and by

Table 1. Characteristics of the cohort by a measure of psychological functioning (stress resilience) in adolescence

	High (7–9) stress resilience n=142 581	**Intermediate (4–6) stress resilience n=385 276**	**Low (1–3) stress resilience n=136 746**
Number of older male siblings, N (%)			
0	97 438 (68.3)	254 383 (66.0)	84 989 (62.2)
1	37 312 (26.2)	105 054 (27.3)	39 962 (29.2)
2	6 802 (4.8)	21 845 (5.7)	9 419 (6.9)
3 or more	1 029 (0.7)	3 994 (1.0)	2 376 (1.7)
Number of older female siblings, N (%)			
0	97 994 (68.7)	258 290 (67.0)	89 176 (65.2)
1	37 002 (26.0)	102 959 (26.7)	37 110 (27.1)
2	6 569 (4.6)	20 364 (5.3)	8 438 (6.2)
3 or more	1 016 (0.7)	3 663(1.0)	2 022 (1.5)
Multiple birth, N (%)			
Singleton	140 112 (98.3)	378 238 (98.2)	134 423 (98.3)
Multiple birth	2 469 (1.7)	7 038 (1.8)	2 323 (1.7)
Mother's age at delivery (years), N (%)			
Under 18	1 048 (0.7)	4 829 (1.3)	2 680 (2.0)
18–24	43 555 (30.6)	133 567 (34.7)	50 655 (37.0)
25–29	58 498 (41.0)	144 061 (37.4)	46 996 (34.4)
30–34	29 864 (21.0)	75 639 (19.6)	25 933 (19.0)
35–39	8 439 (5.9)	23 509 (6.1)	8 916 (6.5)
40–44	1 145 (0.8)	3 535 (0.9)	1 507 (1.1)
45+	32 (0.0)	136 (0.0)	59 (0.0)
ESeC, N (%)			
High	62 789 (44.0)	119 413 (31.0)	33 137 (24.2)
Intermediate	25 406 (17.8)	62 972 (16.3)	18 978 (13.9)
Low	54 386 (38.1)	202 891 (52.7)	84 631 (61.9)
Cognitive function[a]	6.1 (1.6)	5.3 (1.7)	4.2 (1.9)

Notes: [a]Mean (SD).
ESeC, European Socioeconomic Classification

Table 2. Risks of intermediate and low stress resilience in adolescence compared with high stress resilience

	Intermediate stress resilience			Low stress resilience		
	Model 1 RRR (95% CI)	Model 2 RRR (95% CI)	Model 3 RRR (95% CI)	Model 1 RRR (95% CI)	Model 2 RRR (95% CI)	Model 3 RRR (95% CI)
Number of older male siblings						
0	Reference	Reference	Reference	Reference	Reference	Reference
1	1.08 (1.06 to 1.09)	1.10 (1.09 to 1.12)	1.03 (1.01 to 1.04	1.23 (1.21 to 1.25)	1.33 (1.30 to 1.35)	1.14 (1.22 to 1.17)
2	1.23 (1.19 to 1.26)	1.21 (1.18 to 1.25)	1.10 (1.07 to 1.14	1.58 (1.53 to 1.63)	1.65 (1.59 to 1.71)	1.33 (1.28 to 1.38)
3 or more	1.47 (1.37 to 1.58)	1.32 (1.23 to 1.41)	1.14 (1.06 to 1.23)	2.67 (2.48 to 2.87)	2.36 (2.18 to 2.54)	1.69 (1.56 to 1.83)
Number of older female siblings						
0	Reference	Reference	Reference	Reference	Reference	Reference
1	1.05 (1.04 to 1.07)	1.08 (1.06 to 1.10)	1.01 (1.00 to 1.03)	1.10 (1.08 to 1.12)	1.19 (1.17 to 1.21)	1.04 (1.02 to 1.06)
2	1.17 (1.14 to 1,21)	1.15 (1.12 to 1.19)	1.04 (1.01 to 1.08)	1.41 (1.36 to 1.45)	1.46 (1.40 to 1.51)	1.17 (1.13 to 1.22)
3 or more	1.35 (1.27 to 1.46)	1.20 (1.12 to 1.29)	1.05 (0.98 to 1.13)	2.19 (2.03 to 2.36)	1.87 (1.73 to 2.03)	1.37 (1.26 to 1.49)
Multiple birth						
Singleton	Reference	Reference	Reference	Reference	Reference	Reference
Multiple	1.06 (1.01 to 1.11)	1.07 (1.02 to 1.12)	1.03 (0.98 to 1.08)	0.98 (0.92 to 1.04)	0.99 (0.93 to 1.05)	0.91 (0.86 to 0.97)
Mother's age at delivery (years)						
Under 18	1.85 (1.73 to 1.98)	1.46 (1.36 to 1.56)	1.21 (1.13 to 1.29)	3.41 (3.17 to 3.67	2.62 (2.43 to 2.82)	1.77 (1.63 to 1.91)
18–24	1.24 (1.22 to 1.26)	1.09 (1.07 to 1.10)	1.02 (1.00 to 1.03)	1.47 (1.45 to 1.50	1.26 (1.24 to 1.29)	1.09 (1.07 to 1.12)
25–29	Reference	Reference	Reference	Reference	Reference	Reference
30–34	1.03 (1.01 to 1.04)	1.05 (1.03 to 1.06)	1.08 (1.06 to 1.10)	1.05 (1.03 to 1.07	1.03 (1.03 to 1.06)	1.10 (1.08 to 1.13)
35–39	1.12 (1.09 to 1.15)	1.10 (1.07 to 1.13)	1.16 (1.12 to 1.19)	1.27 (1.23 to 1.32	1.12 (1.09 to 1.16)	1.25 (1.21 to 1.30)
40–44	1.24 (1.16 to 1.32)	1.14 (1.07 to 1.22)	1.23 (1.15 to 1.32)	1.63 (1.51 to 1.76	1.22 (1.13 to 1.33)	1.43 (1.32 to 1.56)
45+	1.68 (1.14 t0 2.47)	1.39 (0.94 to 2.05)	1.56 (1.05 to 2.31)	2.38 (1.54 to 3.66)	1.41 (0.91 to 2.19)	1.85 (1.17 to 2.93)
ESeC						
High	Reference	Reference	Reference	Reference	Reference	Reference
Intermediate	1.31 (1.29 to 1.33)	1.32 (1.29 to 1.34)	1.14 (1.12 to 1.16)	1.41 (1.38 to 1.45)	1.41 (1.38 to 1.45)	1.04 (1.01 to 1.06)
Low	1.98 (1.95 to 2.00)	1.93 (1.90 to 1.96)	1.52 (1.50 to 1.55)	3.06 (3.00 to 4.11)	2.81 (2.76 to 2.86)	1.64 (1.61 to 1.68)
Cognitive function						
	0.74 (0.74 to 0.75)	-	0.77 (0.76 to 0.77)	0.52 (0.53 to 0.53)	-	0.55 (0.54 to 0.55)

Notes: Model 1: adjusted for year of conscription assessment and age at assessment. Model 2: further adjusted for all measures in the table, except cognitive function. Model 3: further adjusted for cognitive function. All measures are categorical, except for the continuous cognitive function variable. N=664 603 for all models; RRR: relative risk ratio; 95% CI: 95% confidence interval associations with maternal age at delivery. However, socioeconomic characteristics do not fully explain sibling associations with stress resilience and this is consistent with other studies indicating that birth order has an influence on intelligence not fully explained by socioeconomic factors (Kristensen & Bjerkedal, 2007). Further, familial socioeconomic characteristics are unlikely to explain the higher magnitude associations with stress resilience for having male, rather than female, siblings. It is unlikely that there is a larger proportion of male children due to a larger number of male births in more socioeconomically disadvantaged families. Disadvantage has been linked to a small excess of *female* births (Magnuson, Bodin & Montgomery, 2007) possibly because adversity leads to loss of male foetuses early in pregnancy. More plausible explanations for the association of older male siblings with low stress resilience include the possibility that male siblings make greater demands on available resources – both in terms of time with parents and material factors – within the family (D'Ovidio et al., 2015). As the first child receives more parental attention before the birth of younger children (Eisenman, 1992), and possibly subsequently, this may influence resource availability for younger siblings. It is also conceivable that previous births (older siblings) have an influence on the *in utero* environment, as foetal characteristics have been linked with this measure of stress resilience (Nilsson, Nyberg & Ostergren, 2001), but a study of influences on intelligence found that *living* siblings (rather than those who did not survive beyond pregnancy) and the resulting sibling hierarchy arising from birth order were more relevant to cognitive development than *in utero* effects (Kristensen & Bjerkedal, 2007).

We hypothesise that the presence of older male siblings increases the risk of low stress resilience by a combination of influences such as greater use of resources, including parental attention. Also, as aggression between siblings has stressful sequelae (Tippett & Wolke, 2015), it seems plausible that older male siblings may be in the most dominant position to bully their younger siblings, creating a more threatening environment, with possible implications for stress resilience. These exposures could influence aspects of development relevant to how individuals cope with stress, though psychological and possibly neuroendocrine pathways. One, but not necessarily the most important, aspect of this could be the physiological stress response, which has been most comprehensively described in animal models, such that early exposure to stress is more likely to lead to a tendency to greater stress reactivity and lower stress resilience, thus greater risk of prolonged stress arousal (Liu et al., 1997; Sapolsky 1997). Biological pathways include exposure to stress reducing the expression of glucocorticoid receptors in areas of the brain such as the hippocampus and thus limiting the effectiveness of the negative feedback mechanism to downregulate physiological stress responses. Such processes could also have implications for cognitive function: higher levels of circulating glucocorticoids can have a neurotoxic effect, but this can occur at all ages, as trauma and psychosocial stress is linked with lower hippocampal volume (Woon, Sood & Hedges, 2010), which is relevant to learning and memory. Lower hippocampal volume may result from stressful exposures, but in turn low volume is associated with greater *susceptibility* to exposures increasing the risk of some psychiatric outcomes (van Rooij et al., 2015). This study indicates that associations of older siblings with stress resilience are not explained entirely by a pathway acting though cognitive function, even though higher cognitive function is associated with a reduced risk of low stress resilience (Hiyoshi et al., 2015). Due to the association of stress resilience with cognitive function, we chose to add cognitive function to our models in a separate step to help estimate the extent to which cognitive function may be involved in the association of older siblings with stress resilience. We can only speculate on the mechanisms, but poor stress resilience may have damaged cognitive function, as described above.

Stress resilience indicates the ability to cope with stress: while the measure used here was designed to assess how well someone will cope with combat and other situations faced by the military, it was based on experiences in normal daily life and thus relevant to stress in the general population. This helps to explain why it is associated with a variety of mental and physical health outcomes in adulthood (Bergh et al., 2014; Bergh et al., 2015; Crump et al., 2016; Hiyoshi et al., 2015; Kennedy et al., 2017) and we believe that low resilience is likely to signal a greater tendency to chronic stress arousal with both behavioural and metabolic consequences that are harmful to health.

Childhood experiences signalled by the presence of older siblings and other sources of potentially stressful exposures may have important implications for both stress resilience and cognitive function, and thus for adult health. The differences by sex of siblings likely demonstrates that the older siblings themselves play a role in influencing development, rather that signalling other characteristics of the family. While older male siblings appear to represent a greater risk for low stress resilience, we hypothesise that the risk is not raised by all male siblings, but that it is due to the greater likely of aggressive, bullying or domineering behaviour exhibited by males. Therefore, the stressful aspects of family life, including inter-sibling interactions should be examined and, if possible, reduced for children, as they may have lifelong consequences.

Potential limitations are that the study is limited to males, but the cohort is broadly representative of the male general population. The cohort also comprises ostensibly healthy men, and this is likely to exclude more of those with low stress resilience, perhaps blunting our estimates. There is only a single measurement of stress resilience in late adolescence but, as it has been linked with outcomes in middle age (Hiyoshi et al., 2015), there is evidence of persistence and thus indicates relevance of this measure in adolescence for health in later life. The measure will reflect a combination of inherited characteristics and childhood exposures, but we cannot identify inherited susceptibility factors that may modify the consequences of childhood exposures.

In conclusion, older siblings, particularly males, appear to influence development of stress resilience, highlighting the importance of familial conditions in childhood. This may be due to stressful inter-sibling interactions and unequal allocation of familial resources between siblings.

References

Bergh, C., Udumyan, R., Fall, K., Almroth, H. & Montgomery, S. (2015). Stress resilience and physical fitness in adolescence and risk of coronary heart disease in middle age. *Heart*, *101*(8), 623–629. https://doi.org/10.1136/heartjnl-2014-306703

Bergh, C., Udumyan, R., Fall, K., Nilsagard, Y., Appelros, P., & S. Montgomery (2014). Stress resilience in male adolescents and subsequent stroke risk: cohort study. *J Neurol Neurosurg Psychiatry*, *85*(12), 1331–1336. https://doi.org/10.1136/jnnp-2013-307485

Carlstedt, B. (2000). *Cognitive Abilities*. Göteborgs Universitet.

Crump, C., Sundquist, J., Winkleby, M. A., & Sundquist, K. (2016). Stress resilience and subsequent risk of type 2 diabetes in 1.5 million young men. *Diabetologia*, *59*(4), 728–733. https://doi.org/10.1007/s00125-015-3846-7

D'Ovidio, F., d'Errico, A., Scarinzi, C., & Costa, G. (2015). Increased incidence of coronary heart disease associated with "double burden" in a cohort of Italian women. *Soc Sci Med*, *135*, 40–46. https://doi.org/10.1016/j.socscimed.2015.04.015

Damian, R. I. & Roberts, B. W. (2015). Settling the debate on birth order and personality. *Proc Natl Acad Sci USA*, *112*(46), 14119–14120. https://doi.org/10.1073/pnas.1519064112

ESeC. SOC2010 volume 3: the National Statistics Socio-economic classification (NS-SEC rebased on SOC2010) https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010/soc2010volume3thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010 Retrieved 15 September 2017.

Eisenman, R. (1992). Birth order, development and personality. *Acta Paedopsychiatr*, *55*(1), 25–27.

Ekbom, A. (2011). The Swedish Multi-generation Register. *Methods Mol Biol*, *675*, 215–220. https://doi.org/10.1007/978-1-59745-423-0_10

Gass, K., Jenkins, J., & Dunn, J. (2007). Are sibling relationships protective? A longitudinal study. *J Child Psychol Psychiatry*, *48*(2), 167–175. https://doi.org/10.1111/j.1469-7610.2006.01699.x

Hiyoshi, A., Udumyan, R., Osika, W., Bihagen, E., Fall, K., & Montgomery, S. (2015). Stress resilience in adolescence and subsequent antidepressant and anxiolytic medication in middle aged men: Swedish cohort study. *Soc Sci Med*, *134*, 43–49. https://doi.org/10.1016/j.socscimed.2015.03.057

Kennedy, B., Chen, R., Valdimarsdottir, U., Montgomery, S., Fang, F., & Fall, K. (2018). Childhood Bereavement and Lower Stress Resilience in Late Adolescence. *J Adolesc Health*, *63*(1),108-114. https://doi.org/10.1016/j.jadohealth.2018.02.002

Kennedy, B., Fang, F., Valdimarsdottir, U., Udumyan, R., Montgomery, S., & Fall, K. (2017). Stress resilience and cancer risk: a nationwide cohort study. *J Epidemiol Community Health*, *71*(10), 947–953. https://doi.org/10.1136/jech-2016-208706

Kristensen, P. and Bjerkedal, T. (2007). Explaining the relation between birth order and intelligence. *Science*, *316*(5832), 1717. https://doi.org/10.1126/science.1141493

Lawson, D. W. & Mace, R. (2010). Siblings and childhood mental health: evidence for a later-born advantage. *Soc Sci Med*, *70*(12), 2061–2069. https://doi.org/10.1016/j.socscimed.2010.03.009

Liu, D., Diorio, J., Tannenbaum, B., Caldji, C., Francis, D., Freedman, A., Sharma, S., Pearson, D., Plotsky P. M., & Meaney, M. J. (1997). Maternal care, hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal responses to stress. *Science*, 277(5332), 1659–1662. https://doi.org/10.1126/science.277.5332.1659

Magnuson, A., L. Bodin and S. M. Montgomery (2007). Father's occupation and sex ratio of offspring. *Scand J Public Health*, *35*(5), 454–459. https://doi.org/10.1080/14034940701246066

Nilsson, P. M., Nyberg, P., & Ostergren, P. O. (2001). Increased susceptibility to stress at a psychological assessment of stress tolerance is associated with impaired fetal growth. *Int J Epidemiol*, *30*(1), 75–80. https://doi.org/10.1093/ije/30.1.75

Plomin, R. & Daniels, D. (2011). Why are children in the same family so different from one another? *Int J Epidemiol*, *40*(3), 563–582. https://doi.org/10.1093/ije/dyq148

Rohrer, J. M., Egloff, B., & Schmukle, S. C. (2015). Examining the effects of birth order on personality. *Proc Natl Acad Sci USA*, *112*(46), 14224–14229. https://doi.org/10.1073/pnas.1506451112

Sapolsky, R. M. (1997). The importance of a well-groomed child. *Science*, *277*(5332), 1620–1621. https://doi.org/10.1126/science.277.5332.1620

SCB: The Swedish Total Population Register http://www.scb.se/Statistik/_Publikationer/OV9999_2015A01_BR_BE96BR1501.pdf Retrieved 20 March 2017.

Sulloway, F. J. (1996). *Born to rebel: birth order, family dynamics, and creative lives*. New York: Pantheon Books.

Tippett, N. & Wolke, D. (2015). Aggression between siblings: Associations with the home environment and peer bullying. *Aggress Behav*, *41*(1), 14–24. https://doi.org/10.1002/ab.21557

van Rooij, S. J., Kennis, M., Sjouwerman, R., van den Heuvel, M. P., Kahn, R. S. & Geuze, E. (2015). Smaller hippocampal volume as a vulnerability factor for the persistence of post-traumatic stress disorder. *Psychol Med*, *45*(13), 2737–2746. https://doi.org/10.1017/S0033291715000707

Woon, F. L., Sood, S. & Hedges, D. W. (2010). Hippocampal volume deficits associated with exposure to psychological trauma and posttraumatic stress disorder in adults: a meta-analysis. *Prog Neuropsychopharmacol Biol Psychiatry*, *34*(7), 1181–1188.

Longitudinal and Life Course Studies 2018 Volume 9 Issue 4 Pp 456–465 *ISSN 1757-9597*

Generations and Gender Survey study profile

Anne Gauthier Netherlands Interdisciplinary Demographic Institute and University of Groningen, the Netherlands

Susana Cabaço Netherlands Interdisciplinary Demographic Institute, the Netherlands
cabaco@nidi.nl

Tom Emery Netherlands Interdisciplinary Demographic Institute and Erasmus University Rotterdam, the Netherlands

(Received January 2018 Revised September 2018) http://dx.doi.org/10.14301/llcs.v9i4.500

Abstract

The Generations and Gender Survey (GGS) is a panel study on families, life course trajectories and gender relations. The GGS is part of the Generations and Gender Programme (GGP), a unique research infrastructure providing open access data to registered researchers. We will be focusing on the GGS waves that were already collected. With large samples per country, the GGS microdata provides researchers with a key resource to examine changes in family life, inter-generational and gender relations. The analysis of these trends is at the core of the research produced in several social science disciplines and the GGS data users have extensively used it to better understand topics such as the transition to adulthood, partnership formation and dissolution, fertility, gender roles and caring responsibilities. In the first part of this study profile, we focus on the design features of the GGS (data collection and adjustment, panel maintenance, and coverage) and subsequently we provide an overview of the organisational setup and outputs of the GGP. In the last part we reflect on the opportunities and challenges ahead of the next round of data collection.

Keywords

Panel microdata; cross-national survey; family dynamics; life course trajectories; inter-generational relations

Introduction

The GGS is a cross-national longitudinal survey that provides open access data for researchers on a broad array of topics including partnerships, fertility, work-life balance, gender relations, transition to adulthood, intergenerational exchanges, care and later life. The GGS is a central part of the Generations and Gender Programme, a social science research infrastructure initiated in 2001. The Fertility and Families Survey (FFS) is the predecessor of the GGS and was conducted in the 1990s in 23 member states of the United Nations Economic Commission for Europe (UNECE).

The GGS is an individual-level fixed panel – it collects data from the same persons on multiple occasions. It is not a household or multiple generation within families panel. As of June 2018, it covers 25 countries. Out of these, the following 21 countries carried out a full GGS or survey that is closely comparable: Australia[1], Austria, Belarus[2], Belgium, Bulgaria, Czech Republic, Estonia, France, Georgia, Germany, Hungary, Italy[3], Japan[4], Kazakhstan[5], Lithuania, the Netherlands[6], Norway, Poland, Romania, Russian Federation, Sweden. In addition, data from Spain, United Kingdom, United States and Uruguay have contributed to the Harmonized Histories collection.

The Harmonized Histories data file was created by the Non-Marital Childbearing Network (http://www.nonmarital.org/). It harmonises childbearing and marital histories from GGP countries with data from Spain, United Kingdom, United States and Uruguay – ensuring the data is ready for use in event history analysis.

An important issue to note is that in the first round of data collection the fieldwork implementation was co-ordinated nationally, which contributed to some discrepancies in the start dates of fieldwork, with wave 1 fieldwork extending from 2002–03 until 2012 (see table 1 below). The new round of data collection, starting in early 2020, will be based on a fresh sample of respondents in each country and, to a large extent, the fieldwork implementation will be co-ordinated centrally, allowing for a more aligned (and narrower) chronological fieldwork window.

The GGP was developed to allow researchers to investigate and understand the changing family relations and intergenerational dynamics that have been part of deep demographic transformations in recent years. In addition, the survey includes a longitudinal component that enables the examination of the causes and consequences of behaviours and life events. The GGP is unique in its significant coverage of Central and East European countries and is also the only comparative panel study that covers the core adult age range (from age 18 to 79) (Gauthier and Emery, 2014).

In recent decades, studies based on the GGS have contributed to advancing knowledge about the factors that have been affecting family and life course dynamics and has generated key insights for policymakers. Our main aim in this paper is to provide an overview of the GGS, presenting its main features, from the design to data dissemination, including the background and characteristics of this unique data resource.

We start by focusing on aspects of the design of the GGS, including its main sampling characteristics and data collection methods. Next, we turn to the procedures used to minimise attrition, as well as data adjustment measures and scope of the thematic coverage of the GGS. The last sections will concentrate on organisational aspects (funding, management), outputs and evaluation of the main strengths and weaknesses of the GGS.

Design

The GGS data has been used to investigate partnership dynamics, transition to adulthood, fertility, care and support networks, division of household tasks, and contraception, among other topics. These data are an essential resource in the understanding of fundamental societal challenges across Europe and beyond and form a substantial basis for the formulation of evidence-based policies. The GGS is designed as a three-wave panel with three-year intervals between waves. This longitudinal design is of crucial importance for researchers who are interested in exploring the causes and consequences behind key societal questions. In addition, the longitudinal design enables researchers to explore the timing, frequency and duration of events or circumstances.

The sampling guidelines, summarised by Simard and Franklin (2005), specify three central elements: the target population is the resident non-institutionalised population aged 18–79 (at the time of the first wave); the sample size of wave 1 should be sufficiently high to achieve at least 8,000 interviews in the event of a 3rd wave; and the use of probability sampling is required. It should be noted that the large sample sizes in each wave are one of the distinguishing features of the GGS and this allows data users to study, for example, specific social groups (low-income families) or particular family types. In addition, the broad age range further contributes to opening up new research possibilities in terms of the analysis of intergenerational relations and support.

Given that the GGS is a cross-national survey, the sampling frames were drawn at the national level, in line with country-specific characteristics and adopting the best available resources to define the sampling framework. Three main types of frames were used so far: population registers with names as samplings elements (Austria, Belgium, Italy, Norway and Sweden); area sampling with addresses or dwellings as sampling elements[7] (Austria, Czech Republic, Germany, Hungary, Lithuania, the Netherlands, Romania, Russian Federation); and a combination of area and census information with names or dwellings as sampling elements (Bulgaria, Estonia, France, Georgia and Poland) (Fokkema, Kveder, Hiekel, Emery, & Liefbroer, 2016). A detailed description of the sampling procedures and further documentation can be found via the GGP NESSTAR webpage (under 'Metadata' for each country). Also, the GGS metadata complies with the Data Documentation Initiative (DDI) standard, including information on sampling, questionnaire and codebooks (in the 'Study description' field, users can find information about the distributors, keywords, abstract, and guidelines on bibliographic citation).

In Germany, the study included a Turkish sub-sample. For data users, the over-sampling of sub-populations is very important to facilitate quantitative research on minorities, which enables the comparative analysis of groups with different backgrounds. For example, Wolf (2014) used this data to study the fertility behaviour of Turkish migrants in Germany and concluded that it is strongly associated with migration history (age at migration and duration of stay).

In terms of response rates in GGS wave 1, the average response rate was 56%. However, four countries had relatively low response rates – Belgium, Lithuania, the Netherlands, and the Russian Federation – due to the inability to contact the individuals and their unwillingness to cooperate (Fokkema et al., 2016).

Some national teams – Australia, Austria, Germany, Hungary, the Netherlands, Norway, Russian Federation – also opted to use incentives to stimulate respondent participation in the survey. The incentives were either provided as cash, voucher or lottery ticket. In the Panel Maintenance section, we will explore attrition and measures adopted try to reduce nonresponse and to stimulate survey participation.

Data collection

In the first wave of the GGP, data collection was conducted by national teams, usually composed by national statistical offices and/or national research institutes. To facilitate the alignment of the fieldwork procedures, data collection guidelines were provided to each national team. One good example of this is the implementation of probability sampling across all participating countries. Nonetheless, it is also important to take into consideration that the timing of fieldwork in wave 1 differed considerably between countries – this was related to constraints in each specific country. This information is relevant for data users as the differences between countries might, to some extent, be related to time-specific contextual elements. However, the survey is designed to examine retrospective and within person life-course dynamics, which reduces the need for strict comparability between countries in the timing of fieldwork. It is however an important point for researchers using the data to take note of.

On what concerns the modes of data collection, there is also some diversity in wave 1: Austria, Belgium, France and Germany opted for computer-assisted personal interviewing (CAPI), while eastern and southern European countries implemented PAPI (paper and pencil interviewing). In five other countries a mixed-methods strategy was used – Australia (PAPI, self-administered paper questionnaire (SAPQ)), Estonia (PAPI, SAPQ), the Netherlands (CAPI, SAPQ), Norway (CATI, SAPQ), Sweden (CATI, SAPQ). In the majority of the countries – with the exception of Austria[8] and Italy – a pilot survey was used to test fieldwork procedures and the questionnaire.

Fokkema et al. (2016) studied the average interview length in wave 1 and found large variations across countries: Sweden with 26 minutes[9] to the Russian Federation with 72 minutes – these differences across countries seem to be associated with the survey mode(s) used (countries using CAPI had shorter average interview length); additionally, some countries included some optional sub-modules or added country-specific questions. In wave 2 and 3, most countries continued to use similar survey modes, with PAPI and CAPI prevailing as the most used survey modes.

Regarding the number of contact attempts, in general at least three attempts were made to contact respondents. The minimum number of contact attempts varied according to the contact method used (with more attempts done via telephone than visits to the addresses).

Furthermore, in some countries it was possible to link administrative records with each survey respondent, dependent on the respondent's consent. Benefiting from the fact that Sweden has a central population register, the full sample of the Swedish GGS was linked to a wide range of administrative records before the fieldwork process (carried out by Stockholm University in collaboration with Statistics Sweden) – participation in the survey was dependent on respondent's consent to record linkage. In fact, this strategy allowed also for administrative data validation "this basis of linkage consent enabled the fieldwork to pre-load administrative records […] enabling respondents the opportunity to correct the data where they deemed necessary" (Emery, 2016, p. 6) – as an example, 18.3% of the respondents corrected the educational level information interview.

Table 1. Main fieldwork characteristics of the Generations and Gender Survey round 1

	Gross sample size	Response rate	Mode	Data collection			
		%		Wave 1		Wave 2	
				Start	End	Start	End
Australia	13,571	52.5	PAPI or Phone, SAPQ	Aug-05	Mar-06	Aug-08	Feb-09
Austria	9,006	61.3	CAPI	Sep-08	Feb-09	Sep-08	Feb-09
Belgium	17,836	41.8	CAPI	Feb-08	May-10	-	-
Bulgaria	18,591	74.8	PAPI	Nov-04	Jan-05	Apr-07	Jun-06
Czech Republic	23,824	49.1	PAPI	Feb-05	Sep-05	Jan-08	Mar-09
Estonia	11,192	70.2	PAPI, SAPQ	Sep-04	Dec-05	Jan-08	Mar-09
France	18,009	65.2	CAPI	Sep-05	Dec-05	Oct-08	Dec-08
Georgia	14,000	71.5	PAPI	Mar-06	May-06	Apr-09	Jun-09
Germany	20,623	55.4	CAPI	Feb-05	May-05	Sep-08	Mar-09
Hungary	24,138	83.7	PAPI	Nov-04	Jan-05	Nov-08	Feb-09
Italy	20,787	19.1	PAPI	Nov-03	Jan-04	Feb-07	Mar-07
Lithuania	29,884	35.6	PAPI	Apr-06	Dec-06	Jun-09	Dec-09
The Netherlands	24,425	44.6	CAPI, SAPQ	Oct-02	Jan-04	Sep-06	Jun-07
Norway	25,848	60.2	CATI, SAPQ, Register	Jan-07	Sep-08	-	-
Poland	20,000	33.3	PAPI	Oct-10	Feb-11	Sep-14	Jan-15
Romania	14,280	83.9	PAPI	Nov-05	Dec-05	-	-
Russian Fed.	27,089	44.8	PAPI	Jun-04	Aug-04	2007	2007
Sweden	18,000	54.7	CATI, SAPQ, Register	Apr-12	Apr-13	-	-
Germany – Turkish subsample	13,890	34.5	CAPI	May-06	Nov-16	Sep-09	Feb-10

Note: Adapted from Fokkema et al. (2016). In this table, wave 3 data is not documented because the data is not yet harmonised and not publicly available. Wave 3 was carried out in five countries: Australia, France, the Netherlands, Hungary and Russian Federation. Information about GGS Belarus and GGS Kazhakstan is beyond the scope of this work as these are part of the new round of data collection (GGP 2020).

In Norway, the personal interviews were linked with administrative register data that was used as a data source, with the aim to reduce interview length, fieldwork costs and improve data accuracy. Administrative records were also used in the preparation of the sampling frame and for data quality control purposes. Before the interview, the personal identification number was collected and, for several of the survey questions, the information was collected via the administrative records (Lappegård & Veenstra, 2010). This data linkage allowed for the continuous update of the information after the interview on a number of key topics: births, marital history, migration history, parental leave, income and wealth, educational activity and attainment, and social benefits. An additional advantage was the use of the administrative records information to add rigorous information about the distances between the addresses of family members.

Panel maintenance

The GGP has, since the first wave, devoted efforts to minimise attrition and has put forward some recommendations for fieldwork practices: continuous and close co-operation between the research institute, the fieldwork agency, the interviewers and the respondents; incentives for respondents that can vary across countries; regular contact with respondents through letters, information brochures, requests for updated contact information, and where feasible, the collection of annual information via a short questionnaire; where possible, interviewer continuity is recommended to help establish a rapport between the respondents and fieldwork staff; and specialised interviewer training and supervision is essential (UNECE, 2005).

In the GGS, the challenge is to be able to trace and contact all respondents after three years (the time between GGS waves), which requires measures to trace and motivate respondents in the period in-between waves. Achieving high response rates and low levels of attrition is a big challenge for social science research infrastructures. In fact, "high nonresponse rates pose a major threat to survey quality as they can cause unwanted systematic deviations from the true outcome of a survey" (Stoop, 2005, p. 5).

Focusing now on specific countries, in Austria the attrition level between the first two waves was 22%. Bubber-Ennser (2014) studied the causes of attrition and concluded that it was affected by "a small bias towards family-oriented persons as well as less-educated respondents and persons with migration background" (p. 460), however this deviation does not affect the reliability of the data.

Among the procedures used for tracing respondents between waves in Austria, central register data was used to track any residential moves. Given that Austrian legislation requires that individuals notify the authorities about any residential move, the central register is continually updated. This way, if respondents moved between wave 1 and 2, the contact address in the register was updated (attrition due to unknown address was expected to be comparatively low) – it was crucial for panel maintenance that Statistics Austria had access to the central register. In order to maintain contact and motivate respondents, postcards were sent to respondents with details about the study and findings. Before the second wave, respondents received an invitation letter with general information about the study and were referred to the Austrian GGS webpage for more information (including the results of the first wave). The matching of respondents and interviewers by gender was another measure adopted to facilitate communication and survey participation.

The attrition levels between the first two waves varied between countries with different factors affecting panel maintenance: in Lithuania the level of attrition was 77.1%; in Czech Republic 68%; in Germany 67%; in France 35%; in Bulgaria 27.3%; in the Netherlands 25.4%; and in Georgia 17%.

In France, the cumulative attrition after three waves of the survey was 43% (the highest attrition was registered between wave 1 and 2: 35%; and 17% between wave 2 and 3). According to Régnier-Loilier and Guisse, this decline in attrition can be attributed to the fact that "the people uninterested in the study or who found the questions too intrusive left at the end of the first wave" (2012, p. 12). In terms of socio-demographic characteristics, these authors found a significant effect of gender, age, education level and nationality on attrition. In the next section, we turn to the data adjustments available in the GGS.

Data adjustment

In the context of data management, a number of data adjustment procedures are necessary to produce comparable, representative data. This is because there are a number of issues that usually bias representation, namely, unequal probabilities of selection, coverage rates and nonresponse. A commonly used solution is the construction of weighting variables, which are used to compensate for factors that can make the data collected unrepresentative of the population.

In wave 1 and 2, most GGS countries designed and provided their own country-specific post-stratification weights, with the exception of Bulgaria, Czech Republic, Poland, Romania, and Italy. The weighting factors used for the construction of the weights varied between countries, but the most consistently used were age, gender and region or urbanisation (these data is available in most countries and these are key indicators for researchers).

Fokkema et al. (2016) analysed the representativeness of the wave 1 data – taking into account age, gender, region, marital status, household size and educational level – and concluded that the unweighted data included some bias. However, "when the data were weighted (...) biases for age, gender, region, and household size were substantially lower" (p. 521). Therefore, given the recurrent issues with the representativeness of survey microdata, data users are advised to apply the weights provided ('aweight' in wave 1 and 'bweight' in wave 2).

Data quality is a priority for any GGP outputs and, ahead of any GGS data release, the pre-harmonised data submitted by national teams is prepared and processed. The main goal is to achieve a clear and comparable format for microdata files that is suitable for cross-national comparison.

The harmonisation procedure involves a number of checks and edits, including i) label checks: ensuring that variable names and values labels are consistent across GGS datasets; ii) table harmonisation: tables need to be harmonised according to the ordering criteria for table-rows; iii) routing: in order to ensure that the data matches the underlying structure of the questionnaire; iv) consolidation: the goal is to compile the information scattered across several variables (this is carried out in the Children, Partnership, Parent and Parental Home sections); v) calculation of derived variables: in order to organise the information available, a number of variables are derived from the grid variables (household, children and partnership), month and year, and frequency and unit variables.

Coverage

The GGS data covers a wide range of topics and focus on fertility and partnership histories, gender relations, division of housework, work–family balance, transition to adulthood, intergenerational exchanges, economic activity, retirement, health and well-being. The GGS adopts a life course approach and collects both retrospective information (fertility, family formation and dissolution) and intentions (intentions to have children, intentions of union formation, are examples of prospective questions)[10]. The selection of the themes included in the questionnaire follows theoretically grounded criteria, described by Vikat et al. (2007).

Until now, the GGS has been used extensively both in the population studies community and by users across multiple scientific disciplines. The data has been used in a number of international research projects, and master and doctoral theses.

Furthermore, a recent study by Zimmermann and Konietzka (2018) illustrates the potential of GGS data for cross-national research. This work investigates how family life course patterns vary (and become more destandardised) across individuals according to their levels of educational attainment in seven European countries. The authors find that across cohorts, "family life courses have become more destandardized among the lower than the higher educated in all countries except Germany" (p. 71).

Table 2. Information collected in the GGS

Questionnaire modules	Examples
Household	Household roster; nationality and ethnicity; dwelling unit; building, occupancy; problems and satisfaction with the accommodation; education
Children	Childcare; non-resident children; step children; grandchildren; consolidated children information; complete childbearing history by month; total number of children
Partnership	Current co-resident partner or spouse; current non-resident partner or spouse; intentions of union formation; complete partnership history by month; child alimony/ maintenance; partner alimony
Household organisation and partnership quality	Household organisation; decision-making; partnership quality
Parent and parental home	Co-residence with biological parents; questions about biological parents; brothers, sisters, grandparents; parental home during childhood; intentions to start living separately from parents
Fertility	Ever had sexual intercourse; current pregnancy; fecundity; intentions to have children
Health and well-being	Health in general; personal care; emotional support; locus of control; well-being
Respondent's activity and income	Current activity status; additional job or business; working conditions and availability of reconciliation policies; income from work, benefits and other sources
Household possessions, income and transfers	Household possessions and economic deprivation; income from other sources than employment; total household income; monetary transfers and inheritance
Value orientations and attitudes	Religiosity; attitudes about interpersonal trust; attitudes about marriage
Additions in Wave 2	Complete activity and education history (working status by month; highest level of education reached; full- or part-time employment)

Note: Additional information about each module and variables is available in the Data Documentation section of the GGP webpage: http://www.ggp-i.org/data/methodology/

Organisation

A number of national and international institutions participate and support the GGP Research Infrastructure, including research institutes in population studies, governmental demographic research units and universities. Currently, this governance and management structure is being evaluated in the context of the GGP-EPI (Evaluate, Plan, Initiate) project and work is being done to prepare the Research Infrastructure for inclusion on the ESFRI European Strategy Forum on Research Infrastructures (ESFRI) Roadmap.

In terms of data collection, the national teams – usually composed by central statistical offices, research institutes and governmental agencies – manage the fieldwork and provide the anonymised microdata along with documentation for all survey instruments. They also enter into a legally binding agreement with the GGP that allows dissemination of the national data by the GGP. Data release agreements have been signed between the UNECE and all countries submitting data.

In terms of data access, through the GGP webpage, it is possible to browse and explore the data (via http://www.ggp-i.org/data/browse-the-data/) and perform basic descriptive analyses. Researchers interested in downloading the data need to sign and submit a statement of affiliation, confidentiality and acceptable usage. The UNECE revises and makes the final decision on user accreditation: when the user is granted data access, the data can be downloaded directly from the user account.

Outputs

On what concerns outputs, the GGP has a set of online products available, including data documentation – core and national questionnaires, guidelines – codebooks (for wave 1 and wave 2), sampling information and country-specific documentation. An important online resource is the NESSTAR data interface that provides direct access to GGS data and metadata, allowing also for simple descriptive analysis (the User Guide is available online). The codebooks contain particularly useful information on variable coding and country specificities. This way, all those interested in the GGS data can easily find all the relevant details about data collection, preparation and harmonisation procedures. The GGP also provides aggregate level data – demographic, economic and policy indicators – through the Contextual Database, which can be linked with the microdata. The Contextual Database includes information for 60 countries on more than 100 indicators.

A selection of technical papers is also available covering technical aspects of the GGP: attrition, sampling, fieldwork methods, response rates, and non-response, among other issues. Moreover, some of these papers explore specific national contexts: for example, Régnier-Loilier and Guisse (2012) studied sample attrition and distortion on the French GGS, while Vanderschrick and Sanderson (2012) analysed item non-response in the Belgium GGS. In addition, the GGP website includes a list of hundreds of scientific publications available, based on GGS data.

Evaluation and concluding remarks

Evaluating the Generations and Gender Survey is not an easy exercise. The GGS has unique characteristics and has faced a number of challenges during its existence. In this final section, we start with an overview of the main strengths and weaknesses of the GGS and conclude with a brief outline of the strategy and objectives for the future of the infrastructure.

The GGS has achieved impressive results and made important contributions as a unique longitudinal data resource on families and life course trajectories, shaping to large extent the research agenda in the field. This is in large part due to the theory-driven and multidisciplinary questionnaire, panel design and large sample sizes that are key features of the GGS. Due to its complex structure and numerous constraints, the GGS has faced a number of challenges and problems in the past. First, given the decentralised structure of the GGP and difficulty in enforcing centralised guidelines, there were issues with the modifications introduced by national teams in the questionnaires fielded, which in turn affected negatively the comparability of the data collected. For example, the reference periods for support in Italy were adjusted from one year to four weeks for reasons of national level comparability. In addition, these deviations also contributed to the slow data cleaning and processing that involved time-consuming post-hoc data harmonisation activities. All these deviations are however clearly documented in the online codebook available via NESSTAR.

One other issue concerns the levels of attrition – that also affect a vast number of longitudinal studies – and the difficulties encountered in tracing and motivating panel members. Furthermore, the GGS is not immune to the difficulties that most social science research infrastructures face, namely problems in accessing sustainable funding and raising fieldwork costs. All these issues demand close attention and an effective strategy for tackling these challenges.

One area in which the GGP has been able to advance the incorporation of innovations in surveys concerns the use of administrative records in the survey process. Administrative data can contribute to reducing the burden on survey respondents, which represents an important goal in social science research, not only due to the rising costs of survey data collection, but also because administrative records can potentially provide a more comprehensive picture in certain domains (for example, tax records or social security payments data). Nonetheless, statistical disclosure concerns and data privacy legislation have been at the forefront of the arguments advanced by the data owners/ custodians – typically national statistical offices – to restrict and control the usage of administrative data.

In preparation for the next round of data collection, planned for early 2020, an experimental study on fieldwork strategies is being designed to test the implementation of a mixed-mode strategy of data collection ('push to web' approach) in three national contexts – Croatia, Germany and Portugal – the idea is that these countries represent much of the diversity of European countries, including also in terms of the availability of central register data for drawing samples and fieldwork costs. This work is being conducted in the context of the GGP-EPI project, which is funded by a €2 million grant (n. 739511) by the European Commission. The project has three main objectives: to evaluate the current executive and operational structures of the GGP and assess alternative models; to identify the best model of operations for the GGP's future; and to develop the required legal, technical and financial arrangements necessary for the next phase.

This experimental study will also allow for the testing of a more centralised and standardised model of data collection and management. In each country an additional experiment will be conducted to evaluate a number of fieldwork specifications: in Germany, different strategies for providing incentives will be tested; in Croatia, different strategies for sending out reminders will be evaluated; and in Portugal, different strategies for delivering the incentive to the selected respondent will be assessed. Part of the vision and ambition for the next round will involve improvements in all stages of the survey in order to achieve higher efficiency, lower fieldwork costs, quicker data processing and better microdata.

Acknowledgements

The GGP is partly funded by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 739511 for the project Generations and Gender Programme: Evaluate, Plan, Initiate.

References

Azjen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Bubber-Ennser, I. (2014). Attrition in the Austrian Generations and Gender Survey: Is there a bias by fertility-relevant aspects? *Demographic Research*, *31*, 459–496. https://doi.org/10.4054/DemRes.2014.31.16

Emery, T. (2016). Administrative data linking: enriching administrative data with surveys. Paper presented at the European Conference on Quality in Official Statistics, 31 May–3 June 2016, Madrid.

Fokkema, T., Kveder, A., Hiekel, N., Emery, T., & Liefbroer, A. (2016). Generations and Gender Programme Wave 1 data collection: an overview and assessment of sampling and fieldwork methods, weighting procedures, and cross-sectional representativeness. *Demographic Research*, 34, 499–524. https://doi.org/10.4054/DemRes.2016.34.18

Gauthier, A., & Emery, E. (2014). Generations and Gender Programme: A Research Infrastructure for Analysing Relationships over the Life-Course. In A. Duşa, D. Nelle, G. Stock, & G. Wagner (Eds.), *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences* (pp. 99–108). Berlin: SCIVERO Verlag.

Lappegård, T., & Veenstra, M. (2010). Life-course, generation and gender. LOGG 2007. Field Report of the Norwegian Generations and Gender Survey. Oslo: Statistics Norway. Accessed 29th November 2017 from ggpsurvey.ined.fr/documents/Questionnaires/Norway_FieldReport&Questionnaire_W1_en.pdf

Régnier-Loilier, A., & Guisse, N. (2012). Sample attrition and distortion over the waves of the French Generations and Gender Survey, GGP Technical Paper. Accessed 30th October 2017 from http://www.ggp-i.org/sites/default/files/technicalpapers/Regnier-Lollier_and_Guisse_-_Sample_attrition_and_distortion_over_the_waves.pdf

Simard, M., & Franklin, S. (2005). Sample design guidelines. In United Nations Economic Commission for Europe (ed.). *Generations and Gender Programme. Survey Instruments* (pp. 3–14). New York and Geneva: United Nations.

Stoop, I. (2005). *The Hunt for the Last Respondent*. The Hague: Social and Cultural Planning Office of the Netherlands.

UNECE (2005). *Generations and Gender Programme. Survey Instruments*. New York and Geneva: United Nations.

Vandeschrick, C., & Sanderson, J. (2012). GGS Wave 1 Belgium: Item non-response. GGP Belgium Paper Series n. 5. Accessed 7th October 2017 from http://www.ggps.be/doc/GGP_Belgium_Paper_Series_5.pdf

Vikat, A., Spéder, Z., Beets, G., Billari, F., Bühler, C., Désesquelles, A., Fokkema, T., Hoem, J., MacDonald, A., Neyer, G., Pailhé, A., Pinnelli, A., & Solaz, A. (2007). Generations and Gender Survey (GGS): towards a better understanding of relationships and processes in the life course. *Demographic Research*, *17*, 389–440. https://doi.org/10.4054/DemRes.2007.17.14

Wolf, K. (2014). Fertility of Turkish migrants in Germany: duration of stay matters. MPIDR Working paper WP2014-001. Accessed 20th October 2017 from http://www.demogr.mpg.de/papers/working/wp-2014-001.pdf.

Zimmermann, O. & Konietzka, D. (2018). Social disparities in destandardization – changing family life course patterns in seven European countries. *European Sociological Review*, *34*(1), 64–78. https://doi.org/10.1093/esr/jcx083

Endnotes

1. Data for Australia originates from the 'Household, Income and Labour Dynamics in Australia Survey'.
2. The GGS Belarus data was collected in 2017 and this is the first country to take part in the new round of data collection (GGP 2020).
3. Data for Italy originates from 'Famiglia e soggetti sociali (FSS)' in wave 1 and 'Criticità dei percorsi lavorativi in un'ottica di genere' in wave 2.
4. Data for Japan originates from the 'International Comparative Survey on Marriage and the Family'.
5. The GGS Kazakhstan data was collected recently in 2018 and this is the second country to take part in the new round of data collection (GGP 2020).
6. Data for the Netherlands originates from the 'Netherlands Kinship Panel Study' and more recently from the 'Onderzoek Gezinsvorming'.
7. Details on unit selection available via 'Metadata' in the GGP NESSTAR webpage. The methods used were random number generator (Austria and Estonia), last birthday method (Lithuania and Poland), next birthday method (the Netherlands), first-name method – among those eligible in the household, the person whose first name begins with the letter closest to the beginning of the alphabet was selected – (France), Kish tables (Germany, Romania and Russian Federation).
8. In the case of Austria, a pilot study was not conducted because the same questionnaire had already been used in Germany.
9. The interview length in Sweden was shorter because participation in the survey required respondents' consent to linkage with administrative data records (since the information was pre-loaded for some answers, this reduced the amount of questions asked).
10. Parts of the questionnaire were inspired in the Theory of Planned Behaviour (Azjen, 1991).

AUTHOR GUIDELINES SUMMARY

Submission of Papers

All papers, written in the English language, should be submitted via the *LLCS* website as a Microsoft Word file. If there is a good reason why this is not possible, authors are requested to contact sjeal@slls.org.uk **before** submitting the paper. All subsequent processes involving the author are carried out electronically via the website.

Preparation of Texts

Length. The paper should normally be **approximately 5,000 words**, with longer papers also accepted up to a maximum of 7,000 words. This word count excludes tables, figures and bibliography.

Font and line spacing. Please use Calibri (or similar sans serif) font, 12pt, with 1.5 line spacing in all main text, single space in figure or table captions.

Page layout. All text should be justified to the left hand margin (all margins of at least 2.5cm) with no indents at the start of paragraphs and a line space separating paragraphs. All headings and sub-headings used within the body of the paper should also be left justified. Please do **NOT** use automatic text formatting.

Weblinks. To help our readers to look up cited references or other information available on the web, authors should ensure that all such references are activated.

DOIs. Do NOT include DOIs in the bibliography – they are added during editing as resolvable URLs.

Ensuring an anonymous (blind) review. Please submit papers with a full detailed title page. Once a paper has been submitted to *LLCS* via the website, it will be 'anonymised' by the removal of all author name(s) and institution names on the title page, and any identifying electronic document properties will also be removed. Authors do not need to remove their name(s) from the main text or references but any reference to their work or themselves should be expressed in the third person.

Abstract. The abstract (no subheads or paragraphs) should be no more than 250 words (not part of the main word count).

Keywords. These should be included just below the author list (minimum 3 maximum 10).

Abbreviations. Words to be abbreviated should be spelt out in full the first time they appear in the text with the abbreviations in brackets. Thereafter the abbreviation should be used.

References. Please use the APA 6th edition and refer to examples in the full Guidelines. *Authors not complying with these reference guidelines will be asked to make all the necessary alterations themselves, if their paper is accepted for publication*.

Notes. As a general rule, supplementary notes should be avoided, but if thought to be essential, they should *not* appear on the page as footnotes, but instead, be included as endnotes.

Supplementary material. Supplementary material may be uploaded with the submission, and if the paper is published, will be visible to the reader via a link on the RHS of the screen. ***Submissions containing graphs, tables, illustrations or mathematics.*** All graphs, tables and illustrations should be embedded in the submitted text, and have clear, self- explanatory titles and captions. Mathematical expressions should be created in Word 2003 (or a compatible package) with equation editor 3.0, unless the author has good reason to use other software, in which case please contact sjeal@slls.org.uk. All biological measures should be reported in SI units, as appropriate, followed, in the text, by traditional units in parentheses.

Author citation. If the paper is accepted for publication, a version of the paper with all authors cited in full on the title page will be used. Only individuals who have contributed substantially to the production of the paper should be included.

Copy editing. All accepted manuscripts are subject to copy editing, with reference back to author with suggested edits and queries for response.

Proofs. The corresponding author will be asked to view a layout proof of the article on the website and respond with final amendments within three days.

(Full Author Guidelines at: http://www.llcsjournal.org/index.php/llcs/about/submissions#authorGuidelines**)**

Open Journal System

The ***LLCS*** journal is produced using the Open Journal System, which is part of the Public Knowledge Project. OJS is open source software made freely available to journals worldwide, for the purpose of making open access publishing a viable option for more journals. Already, more than 8,000 journals around the world use this software.

INTRODUCTION

PAPERS

RESEARCH NOTE

STUDY PROFILE

LLCS Journal can be accessed online at www.llcsjournal.org

Published by the Society for Longitudinal and Life Course Studies
info@slls.org.uk www.slls.org.uk