Longitudinal and Life Course Studies:
International Journal

Published by

SLLS Society for Longitudinal and Life Course Studies

LLCS EDITORIAL BOARD

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 http://www.slls.org.uk/#!services/ch6q

All issues are accessible free of charge 12 months after publication.

Print on demand

An attractive printed version of each Issue of the journal, including all back numbers, is 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 http://www.slls.org.uk/#!journal-bookshop/czkc 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.

SLLS disclaimer

The journal's editorial committee makes every effort to ensure the accuracy of all information (journal content). However the Society makes no representations or warranties whatsoever as to the accuracy, completeness or suitability for any purpose of the content and disclaims all such representations and warranties, whether express or implied to the maximum extent permitted by law. The SLLS grants authorisation for individuals to photocopy copyright material for private research use only.

INTRODUCTION

PAPERS

STUDY PROFILE

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

SLLS Society for Longitudinal and Life Course Studies

SLLS Annual International Conference

'Multidisciplinary Collaboration in Longitudinal and Lifecourse Research'

Hosted by The University of Stirling, Scotland, UK
11 – 13 October 2017
(Please note: Post-conference workshops planned for the afternoon of Friday 13 and during Saturday 14 October 2017)

Please mark your calendars for the seventh annual conference of the *Society for Longitudinal and Life Course Studies (SLLS).*

Although the overall conference theme will focus on **multidisciplinary collaboration in longitudinal and lifecourse research** this year, presentations will be from **all areas of longitudinal and lifecourse studies**. These might include (but are not limited to) physical, psychological, social developmental and ageing processes and functioning within and across lifecourse stages from infancy to old age; methods and findings of cohort studies; other sources of longitudinal data such as panel studies and record linkage; international comparisons; household, and income dynamics; intergenerational transfers and returns to learning; gene-environment interactions; 'mixed', and comparative methods; innovative methodology in design, measurement, data management, analysis and research practice (quantitative and qualitative).

KEYNOTE SPEAKERS

Professor Ariel Kalil, University of Chicago
Professor David Bell, University of Stirling

More information about the conference will be posted on the SLLS website in due course - www.slls.org.uk

The fee for conference registration will be discounted for **SLLS members**, so please consider joining if you have not already done so! Visit http://www.slls.org.uk/services for full details and to sign up.

This LLCS Journal will be a possible outlet for the publication of conference papers.

CONFERENCE REGISTRATION OPENS MAY 2017

Editorial: Time to take stock

John Bynner

The most recent editorial committee meeting on 20 April signaled major changes in its membership – see inside the cover of this issue for a full list of members. This is therefore a good time to take stock of the journal's progress.

Progress overview

Good to remember that the origins of the journal lie in a 2007 initiative of the think tank Longview. At a time when national investment in longitudinal data was developing fast, accompanied by an expanding program of research based on it, longitudinal data was seen as under-used, especially by policy makers. There was also a gap in the scientific literature that needed filling. A new scientific journal was seen by Longview as the answer and with the help of a Nuffield Foundation three-year development grant, *Longitudinal and Life course Studies* was launched with a first issue in April 2009.

What was it that needed promoting? As the first Editorial put it "*Longitudinal research* is founded on a set of principles and procedures directed at yielding insights into possible causal connections among social, biological and developmental phenomena. *Life course study* reflects the loosening of disciplinary boundaries in the fields particularly of psychology, sociology, economics, health and history to reflect the growing recognition that the developmental pathways and transitions through which the human life course is constructed are shaped by multiple influences. These include the changing institutions, social structures, and historical era in which development begins."

Longitudinal research is thus the vehicle *par excellence* for mapping the changing human life course within a generation and across generations as development and ageing proceeds through repeated longitudinal studies. Starting in different historical periods, life course study offers a means of showing how the life course is shaped and reshaped through the interactions between changing societal circumstances and individual and collective agency.

Appraisal

Has the ambition of the early days been realised? The signs are that it has but as the data, analytic and policy contexts change at an ever-increasing rate there will always be ground to be made up. From the starting point of three issues a year, we switched in 2015 to four. The current issue is the second in the eighth volume of the journal's output, comprising 159 peer-reviewed papers in 25 individual issues. More than 2,000 registered readers of the journal are now located in 25 different countries across the world and that number is continually expanding. Our output and submissions continue to rise as more and more authors come to see the journal as the vehicle of choice for their longitudinal and life course study publications. Recognition of the journal's achievements is reinforced by its location in such citation databases as Scopus and Thomson Reuters' Emerging Sources.

These progress indicators have paralleled the phenomenal expansion of investment in longitudinal data and in life course study as the means of exploiting it. Despite the setbacks associated with termination of two of the most ambitious programs, such as the US National Children's Study and the UK Life Study, both beginning in pregnancy, the general trend has been upwards with a combination of technological, theoretical and analytic advances making access and use of data – including now 'Big Data' – ever easier. From data enrichment from administrative sources to biomarkers extending to DNA, ever-improving measurement methods, new statistical methods and increasingly flexible software for using and improving them, are transforming the longitudinal research experience and the institutionalisation of life course study as a major science is now assured[2].

Such developments extend in important ways to the next generation through the common strategy of following cohort members' children. As major studies, such as household panels, ageing studies and increasingly birth cohorts harmonise their methods across countries, the life course becomes researchable comparatively. Such *intra-cohort* and *inter-cohort* comparison within multilevel statistical frameworks are perhaps the most significant

developments of all in the new era.

The conclusion must be that by a good luck and judgment the journal was developed at the right time and will continue to thrive.

This issue

The papers in this Issue reflect the conclusions of the previous section through the richness of the journal's coverage.

We start with a study using the US Health and Retirement study data to model the effects of combat exposure during military service as it affects health in later life. The next paper similarly examines health effects, this time arising from the disturbance of sleep patterns as a mediator of chronic medical conditions on depressive symptoms over time. We then move to a comparative study, using Household Panel study data, of the relationship between civic engagement and the education systems operating in England and Germany. A study of wage differentials following a career break using Belgian register data comes next. The final research paper reviews longitudinal methods, in this case comparing the popular sequence analysis with latent class growth models and multistate event history models, as exemplified by Norwegian partnership transition research. A different mode of publication follows in the form of the final paper – a Study Profile devoted to the widely regarded Swiss two-cohort TREE (transition from education to employment) study.

Editorial Committee

The story would not be complete without the most significant development with which this editorial began. As the periods of office of current members of the editorial committee come to an end we have new faces in key positions now ratified by the April meeting of the committee.

As one of the last of the original group of five – Michael Wadsworth, Robert Eriksson, Barbara Maughan and Harvey Goldstein – who established the journal I have now handed over the role of Executive Editor to Heather Joshi (UCL Institute of Education). As Emeritus Professor of economic and developmental demography, previous director of the UCL Centre for Longitudinal Studies and first elected

president of the Society for Longitudinal and Life course studies, Heather is perfectly placed to give the editorial leadership the journal needs.

The other founder member, Harvey Goldstein – co-editor of Statistical Society Series A (Statistics in Society) and a key contributor to the journal's policy development – has for the last three years been sharing with Kate Tilling (Bristol) the Section Editor role for Statistical Sciences and Methodology. He has now completed the full handover to Kate.

Following Richard Layte (Trinity College) having to stand down we now have a new Section Editor for Social and Economic Sciences, Peter Elias, Institute for Employment research University of Warwick. With a background in Applied Labour Economics in Berkeley, past Strategic Adviser for Data Resources to the UK Economics and Social Research Council (ESRC) and a research career drawing extensively on longitudinal data, Peter brings a wealth of invaluable experience to the journal.

New Associate Editors (AEs) appointed last September will also begin their terms of office from 20 April. Health Sciences – Cyrille Delpierre, (Inserm). Michelle Kelly-Irving (Inserm); Statistics and Methodology – Gita Mishra (Queensland), Tim Croudace (Dundee); Behavioural Sciences and Development – Dale Dannefer (Case Western Reserve), Jutta Heckhausen (UC Irvine); Economic and Social Sciences - Bram Vanhoutte (Manchester).

Our gratitude for their great contributions to the journal goes to Richard Layte David Blane, Amanda Sacker, Jan Smit, Lars Bergman, John Hobcraft and Paul Gregg.

Conclusion

In short the journal is on target for achieving its early aims and the new editors will add further to its strengths. It is only left for you, the readers, to promote whenever you get the chance the journal's value as a key scientific resource and the ideal platform for longitudinal and life course research reporting. We value highly your engagement with the journal and welcome your ideas for developing it further.

Military service, combat exposure, and health in the later lives of US men

Alair Maclean   Washington State University Vancouver, US

Ryan D. Edwards   Mills College, Oakland, California, US

Abstract

Researchers have produced mixed findings regarding the relationship between military service, war-zone deployment, combat exposure, post-traumatic stress disorder (PTSD), and physical health at older ages. This article uses data drawn from the Health and Retirement Study (HRS) to estimate growth curve models that predict how self-rated health and life-threatening illness vary across groups of men defined as combat and non-combat veterans, compared to non-veterans. According to the findings, combat veterans have worse health than men who did not experience combat during the draft era decades after their service, while non-combat veterans have health that is similar to if not better than non-veterans. Combat veterans were less healthy than these other men based both on a subjective measure of self-rated health and on an objective count of life-threatening illnesses several decades after service. Studies that simply compare veterans to non-veterans may thus continue to produce mixed findings, because particular types of veterans serve in ways that relate differently to health.

Keywords

Military service, combat exposure, health

Introduction

In the early twenty-first century, 13.7 million American men over the age of 55 (or more than a third) were veterans, many of whom had served in wars stretching from World War II to Vietnam, were sent to foreign lands, and exposed to combat (U.S. Census Bureau, 2014). These men followed different pathways into and had distinct experiences while in the armed forces, which could lead them to have better or worse health than men who did not serve. Some service-members, for example, serve during wartime and are sent overseas to fight in combat, leading to the greater likelihood that they are exposed to trauma. Yet even during wartime, some service members are not sent into battle (MacLean, 2011).

Researchers have produced mixed findings regarding the relationship between military service, war-zone deployment, combat exposure, post-traumatic stress disorder (PTSD), and physical health at older ages. According to previous research, veterans had worse health than non-veterans, but such differences may reflect not an effect of service, but selection into the armed forces (Bedard & Deschenes, 2006; Dobkin & Shabani, 2009; Seltzer & Jablon, 1974). According to other research, veterans had worse health if they were deployed to a combat zone than if they were not, but only within the first five years after they returned from war (Boehmer,

Flanders, McGeehin, Boyle, & Barrett, 2004). Yet others have demonstrated that the negative association of deployment and health persists for longer (McCutchan et al., 2016). Others have focused not on deployment, but on combat exposure, and have demonstrated that men with this military experience were not at increased risk of heart disease decades after their service (Johnson et al., 2010). According to other research, combat veterans and prisoners of war had worse health than other veterans only if they had PTSD, which suggests that this disorder, rather than combat per se, is the relevant exposure (Boscarino, 2006; Kang, Bullman, & Taylor, 2006). Previous research thus leaves open the question of whether combat veterans indeed have worse health   than men who did not see combat, particularly at older ages.

The following article disentangles the impact on older men's health of military service from that of combat exposure experienced decades earlier, by comparing three groups of men: combat veterans, non-combat veterans, and men who did not serve in the armed forces. It evaluates health outcomes among men who served and saw combat from World War II to the Vietnam eras. The analyses are based on retrospective reports of wartime exposure, which have been shown to be closely related to administrative records of conflict (B. Smith et al., 2007). They assess whether combat veterans had worse health based on self-reports and diagnoses of life-threatening illnesses than men who did not see combat, when such effects are measured decades later, while men were in their fifties and older.

Unhealthy veterans: Combat as a negative turning point

Some scholars have pursued research implicitly based on the theory that military service provides a negative turning point for health, due to combat exposure. Previous researchers have argued that combat veterans had higher rates of PTSD and mortality than men who did not see combat decades after their service (Elder, Clipp, Brown, Martin, & Friedman, 2009; Pizarro, Silver, & Prause, 2006; Schnurr & Spiro, 1999). Scholars have speculated that combat veterans have worse health than other men later in life because wartime exposures are associated

with cumulative disadvantage as conveyed by trauma (MacLean, 2010). According to this view, veterans suffer long-term effects of wartime exposure, because they suffer mental and physical trauma during the transition to adulthood. Psychiatrists have argued that the wanderings of Odysseus after the Trojan War continue to echo in the lives of contemporary veterans who find it difficult to readjust to their post-war lives (Shay, 2002). Modern observers have demonstrated that US veterans experienced diminished quality of life when they returned home from fighting since at least the Civil War. During that war, veterans were thought to suffer from a condition called irritable heart (Dean, 1997). In the early part of the twentieth century, the US Congress established a network of medical facilities to rehabilitate veterans who had been negatively affected by combat during World War I, particularly those who had lost limbs (Linker, 2011). While psychiatrists have long been concerned with the psychological consequences of combat, they did not formally label the mental health effects of wartime exposure until 1980, when the Diagnostic and Statistical Manual first included the label of Post-Traumatic Stress Disorder (PTSD) (Laufer, Gallops, & Freywouters, 1984). More recently, scholars have begun to assess the physical and mental injuries experienced by the Americans returning from the contemporary wars in Iraq and Afghanistan (Institute of Medicine, 2013). These injuries may persist in later life, leading directly to ill health. Unfortunately, the data used in these analyses do not indicate whether or not the combatants were wounded.

Negative selection and health outcomes

Veterans may have worse health than other men, however, due not to a causal effect, but to selection. Scholars have long recognised that omitted variable and selection problems complicate inference in studies of military service during the draft era (Angrist, 1990). Even during that era, when service was more widespread than it is today, service-members tended to have different average characteristics than civilians, both because particular types of men may choose to enlist or to evade the draft and because the armed forces choose which potential recruits to accept. Selection fluctuated across the decades of the draft era, with the armed

forces taking relatively more recruits during wartime, particularly World War II and the Vietnam War, whom they might have rejected during peacetime (Flynn, 1993).

Indeed, veterans may suffer worse health if they are exposed to battle due to selection not just into service, but also into combat. Even during wartime, not all service members are sent into battle, nor do all experience combat. Service members have been more likely to serve in combat occupations and fight against the enemy if they have lower ability as measured by their cognitive test scores (Gimbel & Booth, 1996). Combat veterans have also tended to have grown up in families with lower socioeconomic attainment than non-combat veterans (Gimbel & Booth, 1996; MacLean, 2011), which has consistently been associated with worse health (Elo, 2009). Thus, combat veterans may also be more likely to die at young ages and suffer disease due not to their combat exposure, but to their pre-existing characteristics.

Positive selection into service and combat

Veterans may appear more or less unhealthy than non-veterans because of selection into the armed forces and into combat roles based on health. They may therefore have had better health, on average, before their service. People have long been excluded from the armed forces if they are in poor physical health, which includes having asthma or being overweight (National Research Council, 2006) and could also lead veterans to appear relatively healthy. Previous researchers have demonstrated that a variety of outcomes are affected by childhood health, including adult health, socioeconomic attainment, and military service. According to this research, for example, people who were unhealthy as children are unhealthier when they are adults (J. P. Smith, 2009). They also earn less than those who were healthy as children after they enter the labor market (Haas, Glymour, & Berkman, 2011). During the draft era, men were excluded from the armed forces if they had particular health conditions, meaning that service members were healthier, on average, than civilians (Flynn, 1993). Thus, veterans may appear healthier due not to their military experiences, but to

enlistment standards. They would have had better health later in life even if they had not enlisted.

Combat veterans may also be healthier compared to people who did not see combat because of selection into combat occupations and deployment to war zones. Previous researchers have pointed out that the armed forces select troops who are in better health to fight, leading to a "healthy warrior" effect (Armed Forces Health Surveillance Center, 2007). In addition, during World War II, blacks served in segregated units and were therefore less likely to see combat than were whites (MacLean, 2011; Segal & Segal, 2004). Thus, combat veterans may have had fewer mental and physical illnesses than other people even had they stayed at home. They may also suffer worse health than they would have otherwise, but appear in relatively good health because their pre-combat physical and mental fitness mask the effects of any trauma they suffer.

Previous research on combat exposure and health: Primarily negative associations

Researchers have produced findings that are primarily consistent with the view of combat as a negative turning point, though the association appears to change as veterans age. According to some, combat veterans suffer worse health than those who did not experience combat. Scholars have assessed the health of veterans who were recently exposed to combat, demonstrating immediate increases in mental illness and death (Boehmer et al., 2004). In Croatia, people who fought in the war were more likely to experience "distress" in the year after combat ended than those who did not, though this effect appears to hold only for men (Kunovich & Hodson, 1999). In the US, Vietnam veterans were more likely to die of external causes, such as accidents or suicide, in the five years after their service if they deployed to Southeast Asia than if they did not (Boehmer et al., 2004). Yet these short-term effects did not persist; nor were these deployed veterans more likely to die from internal causes, such as illness (Boehmer et al., 2004). Among those serving during the early 2000s, service members and veterans had higher rates of PTSD and depression if they deployed to Iraq and Afghanistan (Hoge & Castro, 2006). These previous findings suggest that veterans may have suffered worse health in the years

immediately after their service if they deployed to war zones than if they served in the United States since at least the Vietnam War.

Other scholars have evaluated the association between combat and health among veterans at middle age, demonstrating alternately a negative association and no association (Boehmer et al., 2004; Boscarino, 2006). They have addressed the question of whether Vietnam and World War II veterans had better or worse health when they were middle aged if they had been sent to war zones and fought in combat when they were younger. According to this research, deployed veterans had higher mortality rates when they were in their thirties and forties compared to those who were not deployed (Boscarino, 2006). They also had higher rates of heavy drinking and drug use, though the evidence is mixed as to whether those rates were associated with mortality (Boehner et al., 2004, Boscarino, 2006). Among veterans who served between the World War II and Vietnam War eras, combat veterans were more likely than other men to report work-related disabilities (MacLean, 2010). Yet, according to other research, deployed Vietnam veterans only had higher rates of mortality in the first five years after their service, and were not distinguishable from non-deployed veterans when they were older (Boehmer et al., 2004).

Still fewer researchers have explored the effects of combat among veterans who were over 50 years old, producing findings that suggest either a negative effect or no effect of wartime exposures at older ages. Looking at these older veterans, researchers have, for example, demonstrated that Vietnam veterans had relatively high rates of PTSD if they experienced combat than if they did not (Schnurr & Spiro, 1999). Similarly, World War II combat veterans had higher mortality rates than veterans who did not fight (Elder et al., 2009). Among Civil War veterans, men had worse health and died at younger ages if they served in companies where more of their comrades were killed (Pizarro et al., 2006). Yet other researchers have demonstrated that combat veterans do not appear to suffer worse physical health than non-combat veterans when they are assessed three or more decades after service (Johnson et al., 2010).

Hypotheses

The following analyses test hypotheses derived from the bulk of the preceding findings, which indicate negative effects of combat on health, potentially varying with age or cohort, along with possible countervailing effects of selection. They therefore test the following hypotheses:

- Hypothesis 1: Combat veterans have worse health than both non-veterans and non-combat veterans.
- Hypothesis 2a: Combat veterans have worse health than both non-veterans and non-combat veterans, and these gaps increase as they age.
- Hypothesis 2b: Combat veterans have worse health than both non-veterans and non-combat veterans, and these gaps decrease as they age.
- Hypothesis 3: Combat veterans have worse health than both non-veterans and non-combat veterans in some but not all cohorts.
- Hypothesis 4: Any associations between health and combat exposure will be explained or suppressed by the pre-service characteristics of combat veterans.

Data and Methods

Data

The analyses are based on the Health and Retirement Study (HRS), with a particular focus on the 2008 wave, when the respondents who were veterans were asked about combat exposure. The HRS was started in 1992, with an original sample of people born between 1931 and 1941. The data were designed to be longitudinal, with survey respondents providing information every two years. In the succeeding years, additional samples have been added to collect data from respondents who were born in both earlier and later years, with the aim of providing data about people who are 50 and over (National Institute on Aging, 2015). Due to the small number of female veterans, the analyses focus on the men who were born between 1908 and 1954 and thus were eligible to serve in the military between the years immediately before World War II through the Vietnam war.

In the HRS overall, the respondents who meet

these criteria amount to a sample of 10,217 men. The analyses are based on samples of men who were included in the 2008 wave who were born in the relevant birth-years and provided data on all of the analysis variables. The samples range between 6,247 and 6,250 respondents depending on the outcome studied. The analysis sample does not differ from the larger HRS sample in their assessment of their childhood health or in the probability that they were black. They do differ in that they were more likely to have been born later in the century, and are thus younger, on average. In addition, they have more educated parents and are more educated themselves. They are also more likely to be Hispanic. Therefore, we ran analyses on samples for which only the dependent variables were missing, substituting the mean when independent variables are missing, along with dummy variables to indicate missing data. The results of these analyses (available by request) do not differ substantively from those presented here.

Dependent variables

The analyses examine the impact of combat exposure by evaluating two dependent variables measured at every survey year. The first measure captures how respondents rate their own health based on responses to the question: "In general, would you say your health is …". Respondents are allocated to the following categories: 1 = poor; 2 = fair; 3 = good; 4 = very good; and 5 = excellent. Many previous researchers have argued that self-rated health reflects objective health. They have shown, for example, that people who rate their health as worse die sooner than do those who rate their health as better (Frankenberg & Jones, 2004). Following Wilmoth, et al. (2010), the models are based on a measure of self-rated health that is continuous.

The second measure assesses whether the respondent has been diagnosed with a life-threatening illness. It is based on that used by Link, Phelan, Miech, and Westin (2008) and reflects whether the respondent reports ever having been diagnosed with one of five serious health conditions: lung problems, cancer, diabetes, heart problems, and stroke. These illnesses, in turn, are associated with subsequent mortality (Link et al., 2008). The analyses are based on a measure of these illnesses that reflects the count. Other analyses (not shown, but available

by request) use measures of both of these dependent variables that are dichotomous and produce results that are substantively similar.

Independent variables

The main independent variable captures a combination of veteran and combat status. The HRS has long asked whether respondents have served on active duty in the armed forces. In 2008, the respondents who had been on active duty were asked if they had ever fired at or were fired on by an enemy. This measure is combined with the previous one to construct a variable that indicates whether the respondent: 0 = did not serve on active duty; 1 = served on active duty but did not see combat; and 2 = served on active duty and saw combat. Unfortunately, the HRS does not include measures that would allow one to examine length of either service or combat exposure.

The models also assess how the health of respondents may vary by either age or cohort. Respondents are included in the data between the ages of 55 and 88. We tested models that included linear and quadratic measures of age, but these results did not differ from those with just the linear measure. They also contain a categorical measure of cohort based on respondents' birth years to reflect the years during which men turned 18, or first became eligible to serve in the military (Wilmoth, Landes, London, & MacLean, Forthcoming). This variable allows the models to compare veterans to non-veterans. The resulting measure places respondents into one of the following 6 categories: 1 = pre-World War II; 2 = World War II; 3 = Post-World War II; 4 = Korean War; 5 = Post-Korea; or 6 = Vietnam War.

The analyses include three other variables that reflect characteristics that have been shown to predict military service during the draft era (MacLean, 2011). The first variable measures socioeconomic background based on reports of the respondent's mother's education. When mother's education is missing, the variable reflects father's education. The second variable specifies the respondent's race and ethnicity, whether the respondent is: non-Hispanic white, non-Hispanic black, or Hispanic. The third variable indicates health selection based on the self-rated health of the respondents reported

retrospectively for their childhood. Scholars have demonstrated that, in the absence of prospective data, retrospective measures of childhood health are both reliable and valid (Haas & Bishop, 2010; J. P. Smith, 2009).

In addition, the analyses include a measure of the respondent's own education, though this variable and health outcomes may be jointly determined by military service and may thus be vulnerable to what Sampson (2008) has labeled "included variable bias." Accordingly, we report results both with and without this measure.

Analytic strategy

The article presents results from multilevel models with random intercepts and random slopes, known as growth curve models, which are estimated in Stata 14 using the "mixed" command. In the current case, the models estimate the average slope of age, while allowing for individual variation around that slope. Respondents can contribute as many as nine observations based on the number of times that they are observed in the data in the specified age range between the 1992 and 2008 waves. The analyses present preferred models that are chosen by

comparing Bayesian Information Criterion (BIC) statistics (Raftery, 1995). Figures are constructed based on these models to represent the marginal effect at the means (or at specified categories) using Stata's "margins" and "marginsplot" commands.

Results

Differences between non-veterans and veterans

Figure 1 presents average self-rated health by age and military status, comparing combat veterans to those who did not see combat and non-veterans. According to the figure, combat veterans rated their health as worse than non-combat veterans and similar to non-veterans when they were in their late 50s. They had self-rated health that was similar to non-combat veterans and better than that of non-veterans in their sixties and seventies. By the time they were in their late eighties, they rated their health again as worse than non-combat veterans and as more similar to that of non-veterans. These non-adjusted figures are partially consistent with hypothesis 1. Combat veterans had worse health at some ages than non-combat veterans, but did not rate their health as worse than non-veterans.

Figure 1. Observed health by age and combat status
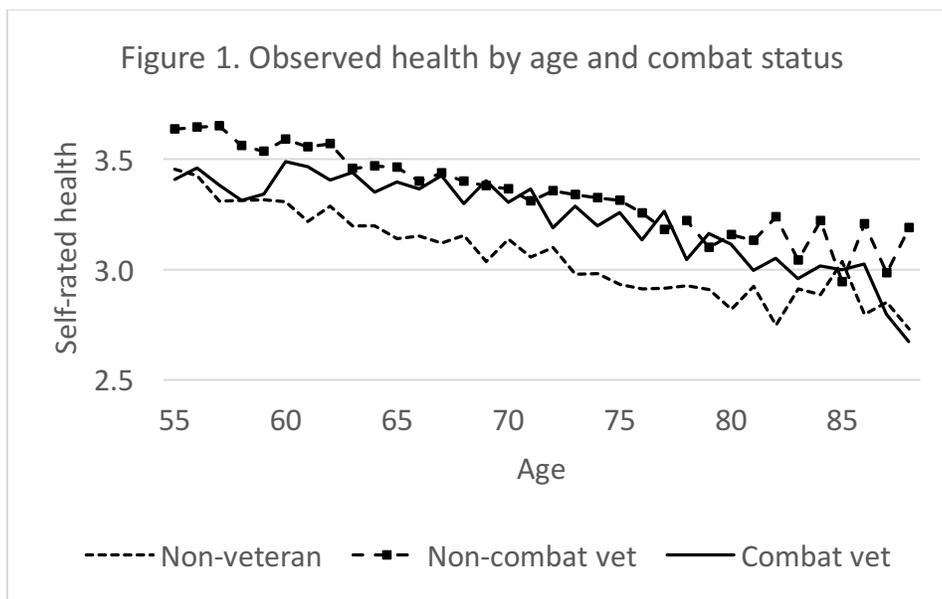
Figure 2 presents the average number of life-threatening illnesses by age and military status. When men were in their late fifties and early sixties, combat veterans had worse health than men in both other groups. Beginning in their early seventies, however, both types of veterans reported more life-threatening illnesses than did non-veterans. These findings also present evidence that is partially consistent with hypothesis 1, combat veterans tended to have worse health than men who did not see combat.

Figure 2. Observed number of illnesses by age and combat status
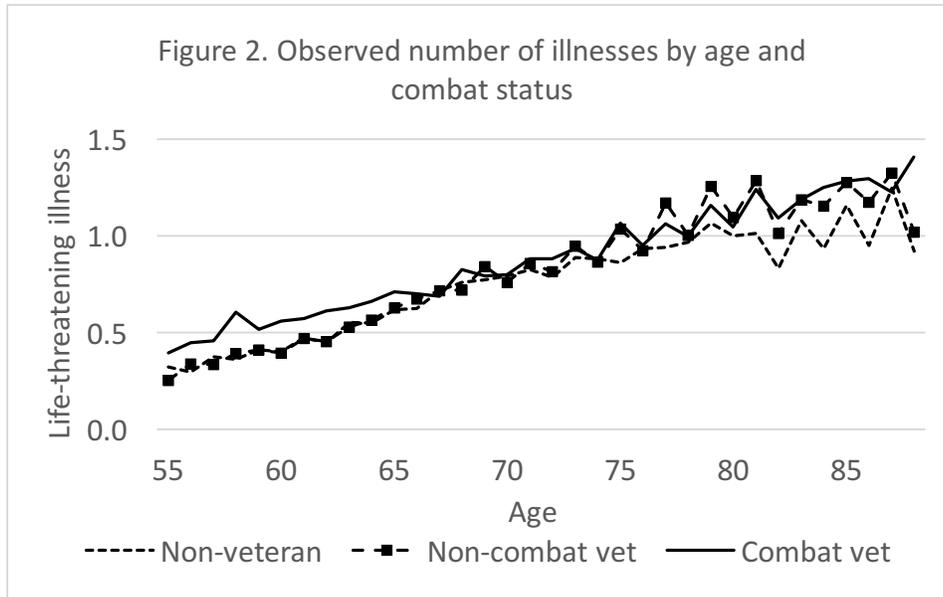
These differences, however, may reflect the differing characteristics of the men in the particular groups. Table 1 presents demographic traits by veteran status and combat exposure. The first two columns present the contrast between veterans and non-veterans. The third and fourth columns contain the same set of statistics for veterans based on whether they did or did not see combat. According to the table, one quarter of the male veterans in the analytic sample indicate that they experienced combat. Among all men in the sample, combat veterans represent approximately 14%.

Table 1. Characteristics of men in the Health and Retirement Study by veteran and combat status (means or proportions)

	Non-veteran	Veteran		Non-combat		Combat		
Mother's years of education	9.455	9.859	***	9.983	***	9.744	*	
	(4.030)	(3.151)		(3.167)		(3.059)		
Respondent's years of education	12.35	13.25	***	13.33	***	13.22	***	
	(3.907)	(2.680)		(2.640)		(2.634)		
Self-rated health (as child)	4.152	4.337	***	4.344	***	4.340	***	
	(1.013)	(0.894)		(0.878)		(0.918)		
Black	0.145	0.094	***	0.0897	***	0.106	**	
Hispanic	0.134	0.041	***	0.0391	***	0.0366	***	
Cohort (years turned 18)								
<1941 (Pre-WWII)	0.051	0.096	***	0.0671	*	0.156	***	†††
1941-1945 (WWII)	0.037	0.135	***	0.108	***	0.204	***	†††
1946-1949 (Post-WWII)	0.061	0.135	***	0.141	***	0.120	***	
1950-53 (Korean war)	0.111	0.155	***	0.172	***	0.108		†††
1954-63 (Post-Korea)	0.391	0.316	***	0.361	*	0.196	***	†††
1964-1973 (Vietnam war)	0.350	0.164	***	0.151	***	0.217	***	†††
Observations	2,995	3,421		2,353		874		

Note: Health and Retirement Study, 1992-2008. Standard deviations in parentheses.

* p<0.05, **p<0.01, *** p<0.001 (two-tailed tests of difference between specified group and non-veterans)

† p<0.05, ††p<0.01, ††† p<0.001 (two-tailed tests of difference between combat and non-combat veterans)

The table indicates that combat veterans were more similar to non-combat veterans before their service in terms of childhood characteristics than they were to non-veterans. According to the first two columns, veterans had more favorable traits than did non-veterans. They came from families in which the mothers had higher education. They had better self-rated health as children. They were less likely to be black or Hispanic. They themselves attained more education. The latter two columns show that combat veterans did not differ from non-combat veterans in terms of these pre-service characteristics. The findings are thus consistent with positive selection into the armed forces.

In addition, veterans came of age in different historical contexts than did non-veterans, while those who saw combat became eligible to serve in eras that differed from those who did not. Veterans were more likely than non-veterans to have been born earlier in the 20th century; they were more likely to have turned 18 in the eras through the Korean war. They were less likely to have come of age after that era. In .

addition, combat veterans were more likely than non-combat veterans to have become eligible to serve in the pre-World War II, World War II, and Vietnam eras.

Preferred Models of Veterans' Health

Table 2 presents fit statistics for select models, which show that the veteran coefficients interacted with neither age nor cohort. The first set of models are those for self-rated health, while the second are those for life-threatening illness. In each of the two sets of models, the first model contains measures of combat status, pre-service characteristics, and age. The second model adds the measure of cohort, which improves the fit of the models. Models 3 and 4 test for interactions between the military status variable and the measures of first age and then cohort. In all cases, model 2 is preferred over the models with interactions, which suggests that military status does not interact with either age or cohort. These findings contradict hypotheses 2a, 2b, and 3, which suggested that the association between military status and health would vary by these attributes

Table 2. Fit statistics for selected models

	DF	BIC	Model Comparison	Difference
Models of life-threatening illness				
Model 1: Combat status, demographics, and age	13	45,207		
Model 2: 1 + cohort	18	44,868	2 - 1	-339.35
Models with interactions with military status				
Model 3: 2 + age	22	44,888	3 - 2	19.52
Model 4: 2 + cohort	28	44,947	4 - 2	78.79
Models of self-rated health				
Model 1: Combat status, demographics, and age	13	92,893		
Model 2: 1 + cohort	18	92,517	2 - 1	-376.03
Models with interactions with military status				
Model 3: 2 + age	22	92,535	3 - 2	17.51
Model 4: 2 + cohort	28	92,608	4 - 2	90.57

Predictors of illness, self-rated health, and depression

Table 3 contains estimates from growth curve models that predict health trajectories, which provide limited evidence of positive selection into the military, and suggest that combat veterans suffered worse health than did all other men. The first three columns reflect how different types of military service are associated with self-rated health, while the next three demonstrate these associations with the number of life-threatening illnesses. Within each set of three models, the first model presents these estimates net of just age and cohort, while the second model adds pre-service characteristics, and the third model adds completed educational attainment.

Table 3. Multilevel models predicting health trajectories

	Predicted self-rated health			Number of life-threatening illnesses		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Veteran type (ref: non-veteran)						
Non-combat veteran	0.236***	0.112***	0.064**	0.021	0.034	0.039
	[0.026]	[0.025]	[0.024]	[0.020]	[0.021]	[0.021]
Combat veteran	0.103**	-0.013	-0.066*	0.124***	0.135***	0.140***
	[0.036]	[0.035]	[0.034]	[0.029]	[0.029]	[0.029]
Age (years after 55)	-0.027***	-0.027***	-0.026***	0.036***	0.036***	0.036***
	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]
Cohort (ref: Korean war)						
Pre-World War II	0.467***	0.523***	0.557***	-0.365***	-0.376***	-0.380***
	[0.058]	[0.056]	[0.054]	[0.051]	[0.051]	[0.051]
World War II	0.205***	0.202***	0.216***	-0.379***	-0.377***	-0.378***
	[0.053]	[0.050]	[0.048]	[0.045]	[0.045]	[0.045]
Post-World War II	0.134**	0.136**	0.127**	-0.202***	-0.199***	-0.199***
	[0.047]	[0.045]	[0.043]	[0.038]	[0.038]	[0.038]
Post-Korean war	-0.107**	-0.153***	-0.185***	0.132***	0.143***	0.146***
	[0.035]	[0.033]	[0.032]	[0.028]	[0.028]	[0.028]
Vietnam war	-0.254***	-0.371***	-0.441***	0.301***	0.327***	0.335***
	[0.040]	[0.038]	[0.037]	[0.031]	[0.031]	[0.031]
Mother's education		0.043***	0.017***		-0.006*	-0.004
		[0.003]	[0.003]		[0.003]	[0.003]
Race/ethnicity (ref: Non-hispanic white)						
Black		-0.287***	-0.187***		0.036	0.025
		[0.035]	[0.034]		[0.029]	[0.029]
Hispanic		-0.220***	-0.037		-0.088*	-0.107**
		[0.044]	[0.043]		[0.036]	[0.037]
Self-rated childhood health		0.205***	0.173***		-0.053***	-0.049***
		[0.012]	[0.011]		[0.010]	[0.010]
Education			0.080***			-0.009**
			[0.004]			[0.003]
Intercept	3.580***	2.447***	1.831***	0.077**	0.350***	0.416***
	[0.035]	[0.064]	[0.069]	[0.027]	[0.053]	[0.059]
Number of observations	37,903	37,903	37,903	37,811	37,811	37,811
Number of respondents	6,250	6,250	6,250	6,247	6,247	6,247

Source: Health and Retirement Study, 1992-2008. Standard errors in brackets.

- p<0.05, **p<0.01, *** p<0.001 (two-tailed tests of significance)
-

According to model 1, non-combat veterans had better health than combat veterans, who, in turn, were healthier than non-veterans net of age and cohort, which reflects the bivariate associations presented in figure 1. After accounting for pre-service characteristics in model 2, however, combat veterans no longer differed from non-veterans in their assessments of their own health. When educational attainment is added in model 3, combat veterans had worse health than both non-veterans and non-combat veterans. (In this, and in every model in the table, the differences between the combat and non-combat coefficients are significant at the .001 level.) These findings provide evidence that veterans were positively selected into service, consistent with hypothesis 4. They are also consistent with hypothesis 1, which suggested that combat veterans suffer worse health compared to men who did not fight.

As shown in models 4 through 6, combat veterans also suffered worse health than did all other men when measured by number of life-threatening illnesses. The estimates of the veteran status variables do not differ statistically across models despite the addition of pre-service characteristics and educational attainment, which is not consistent with hypothesis 4. The associations are not altered by preservice characteristics in the case of illness. Across the various models, however, combat veterans have more life-threatening illnesses, which is consistent with hypothesis 1.

Figure 3 presents predicted trajectories of self-rated health derived from model 3. These predictions are net of background characteristics, education and cohort, reflecting the paths among non-Hispanic white men who turned 18 during the Korean war and who had average childhood health, parental education, and educational attainment. According to the figure, men saw declining health as they aged, with non-combat veterans having better health than both non-veterans and combat veterans. The figure highlights the conclusion that non-combat veterans had better health after their service when compared to non-veterans, while combat veterans did not, decades later.

Figure 3. Fitted self-rated health by age and combat status
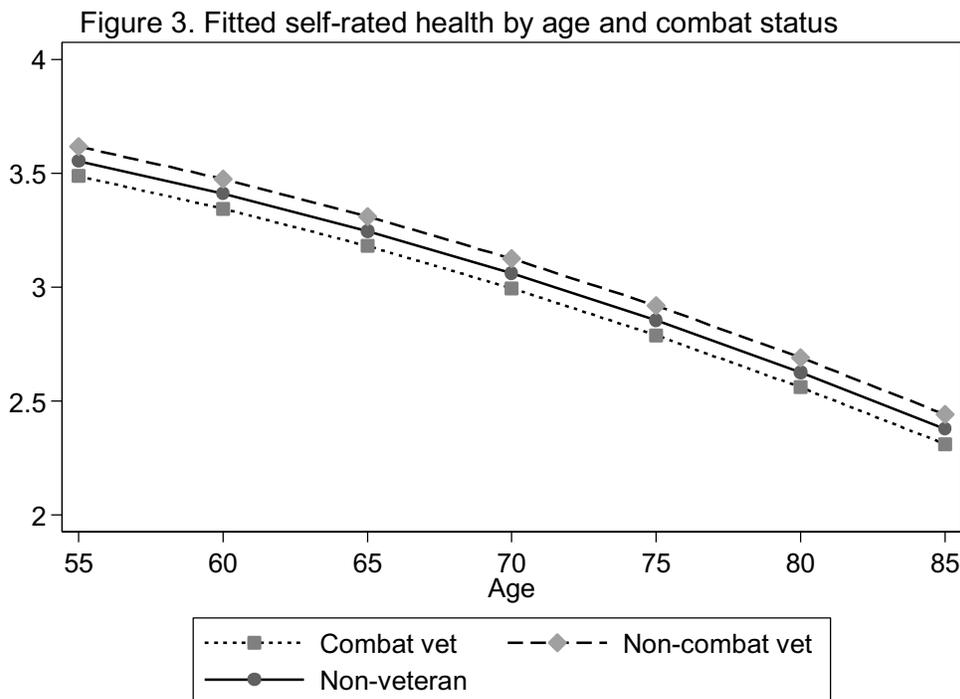
Figure 4 presents trajectories estimated in the same manner as above derived from model 3, but this time for life-threatening illnesses. According to the figure, men became increasingly ill as they grew older. Combat veterans had worse heath than men in both other groups when measured by these diagnoses. Similar to that above, this figure also suggests that combat veterans suffered worse health after their service while non-combat veterans did not.

Figure 4. Fitted illnesses by age and combat status
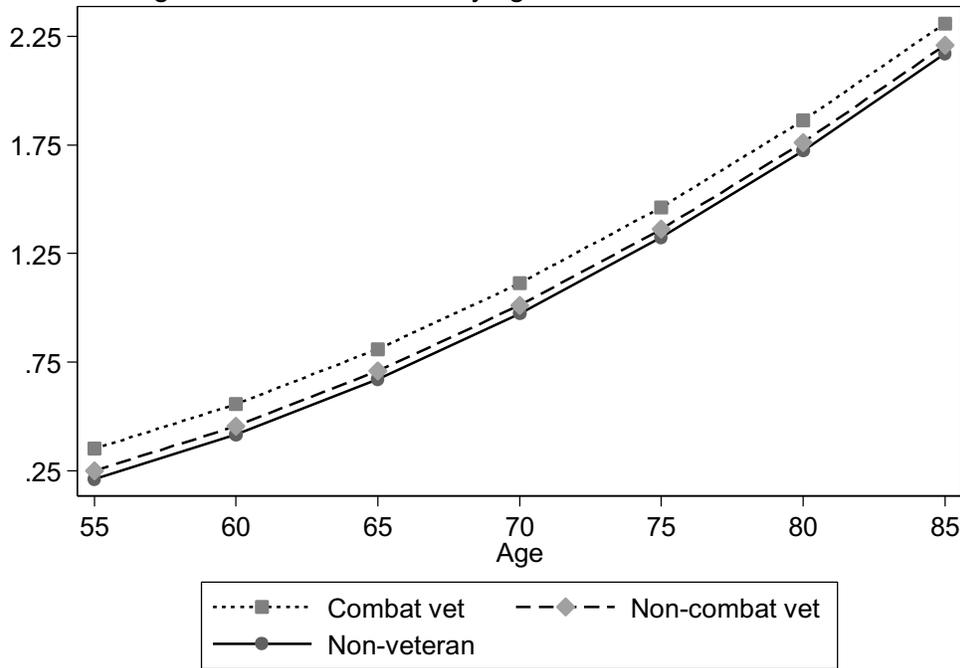
Discussion

This article assesses whether combat has a long-lasting association with health and finds evidence that it does. Combat veterans were less healthy than non-combat veterans based both on a subjective measure of self-rated health and on an objective count of life-threatening illnesses. They also were less healthy than non-veterans, as they were diagnosed with more life-threatening illnesses, net of pre-service characteristics. Yet they did not differ statistically from men who had not served in the military in their subjective assessment based on the measure of self-rated health, until completed education is considered. These findings are consistent with hypothesis 1. Furthermore, they suggest that studies that simply compare veterans to non-veterans may continue to produce mixed findings, because

particular types of veterans serve in ways that relate differently to health.

Yet the analyses suggest that the associations between veteran status and health did not vary across time, at least at these older ages. They did not produce evidence that differences between veterans and non-veterans changed with age, as predicted by hypotheses 2a and 2b. Nor did they indicate that these differences varied across the eras in which men were first eligible to enlist, as predicted by hypothesis 3.

More broadly, scholars have demonstrated that later life outcomes may be affected by traumatic events earlier in life. People are more likely to die at younger ages, for example, if they have been incarcerated (Pridemore, 2014). They may experience long-term effects of a variety of stressors if they grew up in less privileged families (Pearlin,

Schieman, Fazio, & Meersman, 2005). The current findings suggest that combat may fall into this category of a stressor that produces effects not just in the short-term but throughout the life course. In addition, people may be injured in combat, leading to long-term health consequences.

The current set of analyses are limited in at least four ways that would lead them to underestimate the negative association of combat with health. First, the HRS only asked veterans in the core survey about combat exposure in 2008. The analyses are therefore based on information provided by the veterans who remained in the survey until that year. They might therefore underestimate the negative association of combat with health if the veterans who were less healthy were more likely to die or to leave the survey for other reasons than were the unhealthy non-veterans before that wave.

Second, the survey is designed to collect data from people older than 50. As suggested by previous research, veterans may be most negatively affected by combat in the years immediately after they stop serving. If combat veterans suffered worse health and were therefore less likely to participate in the survey, the analyses might further underestimate the negative effects of battle.

Third, the HRS does not include information about PTSD. Researchers have argued that it is this medical condition, rather than combat per se that harms health. In the current set of analyses, combat veterans are included regardless of whether they have PTSD or not, which could lead to an understatement of the impact of PTSD. Nevertheless, the analyses demonstrate that combat itself does

have persistent long-term associations with at least these two measures of health.

Fourth, veterans may have served in the military and in combat due to characteristics that are not fully captured by the independent variables that are included in the current analyses to correct for selection. All veterans may appear to have better health when they are older because the armed forces choose recruits at least partly on the basis of health, excluding potential service-members who are in worse health. Indeed, the preceding findings demonstrate both that veterans had better self-rated health than non-veterans as children and that non-combat veterans rated their health as better than other men when they were older. If service-members were deployed to war zones because they were healthier, as posited by the healthy warrior hypothesis, then the analyses would further underestimate the harmful consequences of combat.

Despite these limitations, the findings have implications for the contemporary era, when more than two million service members have been deployed to the wars in Iraq and Afghanistan. Journalists and lawmakers have recently become concerned that the government is not attending adequately to the health care needs of veterans through the Department of Veterans Affairs (United States Government Accountability Office, 2012). These contemporary veterans have already had to deal with elevated risks of depression and PTSD, as well as suicide. As they grow older, they may also confront increased risks of poor health and life-threatening illness.

References

Angrist, J. D. (1990). Lifetime earnings and the Vietnam era draft lottery - Evidence from Social-Security administrative records. *American Economic Review, 80*(3), 313-336.

Armed Forces Health Surveillance Center. (2007). "Healthy deployers": Nature and Trends of Health Care Utilization during the Year prior to Deployment to OEF/OIF, Active Components, U.S. Armed Forces, January 2002-December 2006. *Medical Surveillance Monthly Report, 14*(3), 2-5.

Bedard, K., & Deschenes, O. (2006). The long-term impact of military service on health: Evidence from World War II and Korean war veterans. *American Economic Review, 96*(1), 176-194. https://doi.org/10.1257/000282806776157731

Boehmer, T. K. C., Flanders, D., McGeehin, M. A., Boyle, C., & Barrett, D. H. (2004). Postservice mortality in Vietnam veterans: 30-year follow-up. *Archives of Internal Medicine, 164*(17), 1908-1916.

https://doi.org/10.1001/archinte.164.17.1908

Boscarino, J. A. (2006). Posttraumatic stress disorder and mortality among US army veterans 30 years after military service. *Annals of Epidemiology, 16*(4), 248-256. https://doi.org/10.1016/j.annepidem.2005.03.009

Dean, E. T. (1997). *Shook over Hell: Post-Traumatic Stress, Vietnam, and the Civil War*. Cambridge, Mass.: Harvard University Press.

Dobkin, C., & Shabani, R. (2009). The health effects of military service: Evidence from the Vietnam draft. *Economic Inquiry, 47*(1), 69-80. https://doi.org/10.1111/j.1465-7295.2007.00103.x

Elder, G. H., Jr., Clipp, E. C., Brown, J. S., Martin, L. R., & Friedman, H. S. (2009). The Lifelong Mortality Risks of World War II Experiences. *Research on Aging, 31*(4), 391-412. https://doi.org/10.1177/0164027509333447

Elo, I. T. (2009). Social Class Differentials in Health and Mortality: Patterns and Explanations in Comparative Perspective. *Annual Review of Sociology, 35*, 553-572. https://doi.org/10.1146/annurev-soc-070308-115929

Flynn, G. Q. (1993). *The Draft, 1940-1973*. Lawrence, Kan.: University Press of Kansas.

Frankenberg, E., & Jones, N. R. (2004). Self-Rated Health and Mortality: Does the Relationship Extend to a Low Income Setting? *Journal of Health and Social Behavior, 45*(4), 441-452. https://doi.org/10.1177/002214650404500406

Gimbel, C., & Booth, A. (1996). Who fought in Vietnam? *Social Forces, 74*, 1137-1157. https://doi.org/10.2307/2580346

Haas, S. A., & Bishop, N. J. (2010). What Do Retrospective Subjective Reports of Childhood Health Capture? Evidence From the Wisconsin Longitudinal Study. *Research on Aging, 32*(6), 698-714. https://doi.org/10.1177/0164027510379347

Haas, S. A., Glymour, M. M., & Berkman, L. F. (2011). Childhood Health and Labor Market Inequality over the Life Course. *Journal of Health and Social Behavior, 52*(3), 298-313.

Hoge, C. W., & Castro, C. A. (2006). Post-traumatic stress disorder in UK and US forces deployed to Iraq. *Lancet, 368*(9538), 837-837. https://doi.org/10.1016/S0140-6736(06)69315-X

Institute of Medicine. (2013). *Returning home from Iraq and Afghanistan: Assessment of readjustment needs of veterans, service members, and their families*. Washington, D.C.: National Academies Press.

Johnson, A. M., Rose, K. M., Elder, G. H., Jr., Chambless, L. E., Kaufman, J. S., & Heiss, G. (2010). Military Combat and Risk of Coronary Heart Disease and Ischemic Stroke in Aging Men: The Atherosclerosis Risk in Communities (ARIC) Study. *Annals of Epidemiology, 20*(2), 143-150. https://doi.org/10.1016/j.annepidem.2009.10.006

Kang, H. K., Bullman, T. A., & Taylor, J. W. (2006). Risk of selected cardiovascular diseases and posttraumatic stress disorder among former World War II prisoners of war. *Annals of Epidemiology, 16*(5), 381-386. https://doi.org/10.1016/j.annepidem.2005.03.004

Kunovich, R. M., & Hodson, R. (1999). Civil war, social integration and mental health in Croatia. *Journal of Health and Social Behavior, 40*(4), 323-343. https://doi.org/10.2307/2676329

Laufer, R. S., Gallops, M. S., & Freywouters, E. (1984). War Stress and Trauma - the Vietnam Veteran Experience. *Journal of Health and Social Behavior, 25*(1), 65-85. https://doi.org/10.2307/2136705

Link, B. G., Phelan, J. C., Miech, R., & Westin, E. L. (2008). The resources that matter: Fundamental social causes of health disparities and the challenge of intelligence. *Journal of Health and Social Behavior, 49*(1), 72-91. https://doi.org/10.1177/002214650804900106

Linker, B. (2011). *War's waste: Rehabilitation in World War I America*. Chicago ; London: University of Chicago Press. https://doi.org/10.7208/chicago/9780226482552.001.0001

MacLean, A. (2010). The Things They Carry: Combat, Disability, and Unemployment among U.S. Men. *American Sociological Review, 75*(4), 563–585. https://doi.org/10.1177/0003122410374085

MacLean, A. (2011). The Stratification of Military Service and Combat Exposure. *Social Science Research, 40*, 336-

348. https://doi.org/10.1016/j.ssresearch.2010.04.006

McCutchan, P. K., Liu, X., LeardMann, C. A., Smith, T. C., Boyko, E. J., Gore, K. L., Freed, M. C., & Engel, C. C. (2016). Deployment, combat, and risk of multiple physical symptoms in the US military: a prospective cohort study. *Annals of Epidemiology, 26*(2), 122-128. https://doi.org/10.1016/j.annepidem.2015.12.001

National Institute on Aging. (2015). Growing Older in America: The Health and Retirement Study.   Retrieved February 17, 2017, 2017, from https://www.nia.nih.gov/health/publication/growing-older-america-health-and-retirement-study/introduction

National Research Council. (2006). *Assessing fitness for military enlistment: Physical, medical, and mental health standards*. Washington, D.C.: National Academy Press.

Pearlin, L. I., Schieman, S., Fazio, E. M., & Meersman, S. C. (2005). Stress, health, and the life course: Some conceptual perspectives. *Journal of Health and Social Behavior, 46*(2), 205-219. https://doi.org/10.1177/002214650504600206

Pizarro, J., Silver, R. C., & Prause, J. (2006). Physical and mental health costs of traumatic war experiences among civil war veterans. *Archives of General Psychiatry, 63*(2), 193-200. https://doi.org/10.1001/archpsyc.63.2.193

Pridemore, W. A. (2014). The Mortality Penalty of Incarceration: Evidence from a Population-based Case-control Study of Working-age Males. *Journal of Health and Social Behavior, 55*(2), 215-233. https://doi.org/10.1177/0022146514533119

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. In P. V. Marsden (Ed.), *Sociological Methodology* (pp. 111-163). Cambridge: Basil Blackwell. https://doi.org/10.2307/271063

Sampson, R. J. (2008). Moving to inequality: Neighborhood effects and experiments meet social structure. *American Journal of Sociology, 114*(1), 189-231. https://doi.org/10.1086/589843

Schnurr, P. P., & Spiro, A. A., III. (1999). Combat exposure, posttraumatic stress disorder symptoms, and health behaviors as predictors of self-reported physical health in older veterans. *Journal of Nervous and Mental Disease, 187*(6), 353-359. https://doi.org/10.1097/00005053-199906000-00004

Segal, D. R., & Segal, M. W. (2004). America's military population. *Population Bulletin, 59*(4), 3-40.

Seltzer, C. C., & Jablon, S. (1974). Effects of selection on mortality. *Am J Epidemiol, 100*(5), 367-372. https://doi.org/10.1093/oxfordjournals.aje.a112047

Shay, J. (2002). *Odysseus in America: Combat Trauma and the Trials of Homecoming*. New York: Scribner.

Smith, B., Wingard, D. L., Ryan, M. A. K., Macera, C. A., Patterson, T. L., & Slymen, D. J. (2007). US military deployment during 2001-2006: Comparison of subjective and objective data sources in a large prospective health study. *Annals of Epidemiology, 17*(12), 976-982. https://doi.org/10.1016/j.annepidem.2007.07.102

Smith, J. P. (2009). Reconstructing Childhood Health Histories. *Demography, 46*(2), 387-403. https://doi.org/10.1353/dem.0.0058

U.S. Census Bureau. (2014). B21001: Sex by Age by Veteran Status for the Civilian Population 18 Years And Over - Universe: Civilian population 18 years and over 2009-2013 American Community Survey 5-Year Estimates.   Retrieved August 7, 2015, 2015, from http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_13_5YR_B21001&prodType=table

United States Government Accountability Office. (2012). *VA Health Care: Reliability of Reported Outpatient Medical Appointment Wait Times and Scheduling Oversight Need Improvement*. Washington, DC: GAO.

Wilmoth, J. M., Landes, S. D., London, A. S., & MacLean, A. (Forthcoming). Appendix: Recommendations for Defining Periods of Military Service and Corresponding Age-18 Cohorts. In A. A. Spiro, III, R. A. Settersten, Jr., & C. M. Aldwin (Eds.), *In Long-Term Outcomes of Military Service: Perspectives on Health and Wellbeing*. Washington, DC: American Psychological Association Press.

The effect of sleep disturbance on the association between chronic medical conditions and depressive symptoms over time

Amanda Leggett   University of Michigan, Ann Arbor, MI, US
leggetta@med.umich.edu
Shervin Assari   University of Michigan, Ann Arbor, MI, US
Sarah Burgard   University of Michigan, Ann Arbor, MI, US
Kara Zivin   University of Michigan and Department of Veterans' Affairs, US

Abstract

Chronic medical conditions (CMC) and sleep disturbances are common among adults and associated with depression. We tested sleep disturbance as a moderator of the effect of CMC on depressive symptoms. The sample includes 3597 adults surveyed up to five times over 25 years (1986-2012) from the nationally representative American's Changing Lives Study (ACL). A multi-level model was estimated to examine sleep disturbance as a moderator of the CMC and depressive symptom association, with a second interaction tested for age as a moderator of the within-person level variability in CMC and depressive symptom association. Sleep disturbance and CMC were associated with depressive symptoms at the between-person level, while only sleep disturbance was associated with depressive symptoms at the within-person level. Sleep disturbance significantly interacted with CMC such that more CMCs were associated with more depressive symptoms among individuals sleeping well, but poor sleep was associated with worse depression regardless of CMC. A second interaction between age and within-person variability in CMC was found significant, suggesting that younger adults had higher symptoms of depression at times of below average CMC relative to older adults. The effect of CMC on depressive symptoms may depend on sleep as well as age. Sleeping restfully may allow individuals with CMC the rejuvenation needed to cope with illness adaptively.

Keywords

Chronic medical conditions, sleep disturbance, depressive symptoms, moderation

Introduction

Chronic medical conditions (CMCs) are of increasing concern and burden as individuals age. The proportion of the population reporting multiple CMCs ranges from approximately 21% of the general adult population to over 60% among older adults (Vogeli et al., 2007). By 2030, an estimated half of the population will have at least one CMC (Anderson, 2010). While advancing treatments for CMCs have led to a decline in risk of associated mortality, individuals are now living longer with these CMCs, and health care utilisation and costs increase significantly with each additional CMC (Lehnert et al., 2011). An added problem is that CMCs increase risk for and are commonly comorbid with mental illnesses such as depression (Bhattarai, Charlton, Rudisill, & Gulliford, 2013; Egede, 2007; Moussavi et al., 2007). Given the association between depressive symptoms and CMCs, determining ways to intervene and improve mental health among these individuals has the potential to enhance quality of life, and reduce overall healthcare costs. The current study explores whether sleep disturbance compounds the effect of CMC on depression and thereby may be a key target for intervention among individuals with CMC.

Mental health consequences of chronic medical conditions

Some researchers have suggested that depression would be less common in later life if not for declines in health that put individuals at increased risk for the disorder (Schnittker, 2005). Moussavi et al. (2007) found as many as 23% of individuals with a physical CMC also had depression, significantly more than the comparison population without a CMC (3.2%). In another study the 12-month prevalence of major depressive disorder ranged from 8% to 17% among adults with different CMCs, and relative to those without chronic medical conditions, those with a chronic condition and depression had greater odds of emergency room and ambulatory visits, time in bed due to illness, and functional impairment (Egede, 2007). The relative rate of depression increased significantly with an increasing number of morbidities (ex. relative rate ratio of 1.63 for one condition versus 2.35 for 3 conditions) (Bhattarai et al., 2013). In prospective evidence, onset of a long-term depressive

episode was predicted by having CMCs and limitations in activities of daily living due to health problems (Geerlings, Beekman, Deeg, & Van Tilburg, 2000). The current study adds to this evidence by taking a life-course developmental approach to consider whether the number of CMCs, as well as within-person variability in multimorbidity over time, is associated with depressive symptoms.

Chronic conditions may be associated with depression for a number of reasons. Depression may result from an individual's psychological response to developing a chronic condition, may be a symptom, complication, or functional impairment that results from the chronic condition, may be a side effect of a medication prescribed to treat a chronic medical conditions, and/or may stem from a pathophysiologic effect of the chronic condition (e.g. inflammation) (Katon, 2003; Konsman, Parnet, & Dantzer, 2002; Ormel, Rijsdijk, Sullivan, van Sonderen, & Kempen, 2002). The theory of vascular depression, for example, posits that cerebrovascular diseases alter brain structure and function, thereby leading to depression (Alexopoulos et al., 1997). Additionally, sleep disturbance may reduce one's energy and capacity to adaptively cope with one's medical condition, and thereby moderate the association between CMC and depression.

Sleep as a potential moderator

Disturbed sleep has also been identified as a contributing factor to the development and expression of depression (Baglioni et al., 2011; Maglione et al., 2012). For example, non-depressed individuals with insomnia have a two-fold increased risk of subsequently developing depression (Baglioni et al., 2011). Even among individuals without an insomnia diagnosis but with insomnia symptoms, the mean subsequent incidence of depression from a meta-analysis of nine studies with mixed age samples was 13.1% compared to only 4% among individuals without sleep problems. Sleep disturbances are highly prevalent in individuals with CMCs, and are associated with decrements in health-related quality of life (Katz & McHorney, 2002). In a community-dwelling sample of adults, 44.1% of individuals with heart disease reported chronic insomnia compared to

only 22.8% of those without heart disease (Taylor et al., 2007).

Disturbed sleep is likely a disruptive symptom and outcome of many CMCs, and may be an ailment that individuals face as a side effect of treatment for their illness. For example, insomnia is commonly experienced by individuals with cancer and is considered one of the most distressing symptoms (Ann M Berger, 2009; Ann Malone Berger & Mitchell, 2008; Bower, 2008; Carpenter et al., 2004; Wielgus, Berger, & Hertzog, 2009). Work by Devins et al. (1993) found that restless sleep in individuals with CMC was associated with illness intrusiveness, defined as the inability to partake in desired activities due to one's illness, and suggested that this may account for some of the association between sleep and depression in individuals with CMC. On its own, poor sleep is also associated with malaise and exhaustion (Parish, 2009), which may make illness management more difficult. Longitudinal studies and clinical trials also suggest that the association between sleep disturbance and pain may be bidirectional, further challenging how individuals may manage and adapt to CMCs. Pain can disturb one's sleep, and likewise, disturbed sleep may intensify pain (Smith & Haythornthwaite, 2004). When someone is sleeping restlessly, they may not be able to manage a CMC as well as those who are maintaining restorative sleep (Hamilton, Catley, & Karlson, 2007). Sleep disturbance, therefore, is directly and negatively related to mental health, and our study uniquely tests whether sleep disturbance may also compound the effect of CMCs on depression by reducing the body's ability to adaptively respond to, cope with, and recover from chronic medical conditions (Hamilton et al., 2007).

Age and illness

As a second aim we consider whether an individual's age interacts with variability in one's level of CMC in predicting depressive symptoms. As we examine both the mean level of CMC and an individual's variability around his/her mean, age is important to consider as a time with less CMC is likely a time where an individual is younger, given the cumulative pattern of CMC. Further, prior research suggests that whether a stressor occurs at an appropriate developmental time, as opposed to a less expected time, relates to its level of impact (Sherbourne, Meredith, Rogers, & Ware, 1992). A younger adult experiencing a CMC may suffer more stress due to its unexpected nature, on the other hand an older adult may experience greater disability from the illness. Additionally older adults tend to use different coping strategies in the face of stress – such as appraising stressors more positively than younger adults and using distancing strategies (Berg & Upchurch, 2007; Brandtstädter & Greve, 1994; Diehl, Coyle, & Labouvie-Vief, 1996; Felton & Revenson, 1987; Folkman, Lazarus, Pimley, & Novacek, 1987; Heckhausen & Schulz, 1995).

The current study

Utilising a longitudinal multi-level approach and multiple observations nested within each respondent, in the current study we first consider chronic medical conditions and sleep disturbance as independent covariates of depressive symptoms over time in a nationally representative sample of US adults. Associations between sleep disturbance, CMCs and depression have been well established; however our longitudinal model furthers this work by examining how sleep disturbance and CMCs vary over time and whether individual variation in these factors is associated with changes in depressive symptoms. Second, our main focus is to examine whether sleep disturbance acts as moderator of the effect of CMCs on depressive symptoms. This aim is in line with studies which have shown a moderating impact of sleep on various risk factors in association with depression and psychological strain, and extends prior work which has looked at the main effects of CMCs and sleep disturbance on depression by considering potential moderation (Leggett, Burgard, & Zivin, 2016; Sanz-Vergel, Demerouti, Mayo, & Moreno-Jiménez, 2011). Finally as a second aim we consider whether there is an interaction between age and an individual's variability in CMCs in association with depressive symptoms.

Methods

Data are drawn from the nationally representative Americans' Changing Lives (ACL) study (House, Kessler, & Herzog, 1990; House, Lantz, & Herd, 2005a; House et al., 1994), a stratified, multistage area

probability sample of non-institutionalised adults age 25 and over, living in the coterminous United States, and followed over a 25-year period. The ACL study was begun from a "stress and coping" framework to examine a broad array of indicators of health, psychosocial and behavioral functioning as individuals age through middle and late adulthood (House, Lantz, & Herd, 2005b). To achieve the nationally representative sample, the sample design started with single counties which were stratified by region and metropolitan statistical area status and selected by probability proportionate to estimated size (PPES). Next area segments (blocks with at least 200-250 persons) were selected by PPES and finally housing units were selected from the area segments. African Americans and adults age 60 and over were over sampled. Interviews were conducted in person in the first two waves, and mainly via telephone thereafter. The first wave of the survey was conducted in 1986 with 3,617 adults. Surviving respondents were re-interviewed in 1989 (N=2867, 83% of survivors), in 1994 (N=2562, 83% of survivors), in 2001/2002 (N=1787, 74% of survivors), and again in 2011/2012 (N=1427, 81% of survivors). The final analytic sample includes 3,597 participants who completed the CES-D items during at least one wave. A sampling weight is incorporated for non-response and a post-stratification adjustment to the 1986 Census estimates of the US population age 25 years and older to ensure that the sample is representative of the age, gender and race distribution of the US population living in the United States in 1986 (House et al., 2005a).

Measures

The primary outcome of interest was *depressive symptoms* measured at each wave with an 11-item short form of the Center for Epidemiological Studies Depression Scale (CES-D) adapted by Kohout, Berkman, Evans, and Cornoni-Huntley (1993) from Radloff (1977)'s original scale. Kohout's 11-item scale was found to be reliable and contain the same four-factor solution as the original 20-item scale. In our analyses the CES-D item "my sleep was restless" was used as a covariate to measure sleep disturbance, so a mean of the 10 remaining CES-D items was used as the measure of depressive symptoms. The Cronbach's

alpha for our 10-item scale also showed strong reliability across waves ranging from .809 to .847. The 10-item scale contains three somatic items (everything was an effort, could not get going, didn't feel like eating), three depressive affect items (felt depressed, lonely, sad), two positive affect items (feel happy, enjoy life; positive items were recoded), and two interpersonal interaction items (people dislike me, were unfriendly). Participants responded on a 1 to 3 scale indicating how frequently the symptom was experienced in the past week (1 = hardly ever, some of the time, or 3 = most of the time).

To assess **chronic medical conditions**, participants reported whether they had a doctor's diagnosis of the following CMCs in the past 12 months: hypertension, diabetes, chronic lung disease, heart attack or other heart trouble, stroke, cancer, and arthritis. Conditions were summed for a maximum score of seven. This approach to assessing chronic conditions is utilised in many large nationally representative surveys such as the National Health and Nutrition Examination Survey (NHANES) (Schnell et al., 2012), the Longitudinal Aging Study Amsterdam (Geerlings et al., 2000), and Health and Retirement Study (HRS) (Lee, Cigolle, & Blaum, 2009). Our count included commonly measured chronic conditions that have previously been associated with depressive symptoms (Bhattarai et al., 2013; Geerlings et al., 2000).

Sleep disturbance was measured using the CES-D item "my sleep was restless". Participants responded on a 1 to 3 scale indicating how frequently the symptom was experienced in the past week (1=hardly ever, 2=some of the time, or 3=most of the time). Though ACL does not have objective measures of sleep, prior work by the ACL study team found this CES-D item to be associated with other predictors in the same way that more detailed measures of sleep were associated with those predictors (Burgard & Ailshire, 2009; Leggett et al., 2016).

Time-invariant demographic controls were taken from wave 1 and included age, gender, race (White, non-White), years of education and functional ability. *Functional ability* was assessed with a Gutman-type scale ranging from one – most severe functional impairment to four – no impairment or high functional ability. This scale was developed from items including being in a bed/chair most/all of the

Leggett, Assari, Burgard, Zivin

The effect of sleep disturbance on the association between chronic medical conditions and depressive symptoms over time

day and difficulty with bathing by oneself, climbing stairs, walking several blocks, and with heavy housework.

Statistical analysis

First, we computed descriptive statistics for all variables. We employed multilevel modeling (Littell, Milliken, Stroup, & Wolfinger, 1996; Raudenbush & Bryk, 2002) using SAS 9.4 PROC MIXED to examine between-person differences (person level) and within-person variability in depressive symptoms across five survey interview waves (wave level) conducted over 25 years. Raw variables were transformed into their between and within-person components prior to analysis. Predictors at the between-person level (Level 2; individual specific- i) were entered as person-specific means (an individual's average across five waves) and within-person level (Level 1; wave by individual specific- wi) predictors were time-varying (deviations in a given wave from an individual's own mean across up to five waves) (Hoffman & Stawski, 2009). This analysis strategy allows us to examine between-person differences in covariates (e.g. are individuals with more CMCs more depressed than individuals with less CMCs?) and also how fluctuations in covariates over time reflect within-person associations between the covariates and depressive symptoms (e.g. at a time when I have more CMCs than my average, am I more depressed?). These multilevel models utilise all waves of data that are available for an individual in the model, a useful feature as missing waves of data are common among individuals in panel studies (Raudenbush & Bryk, 2002).

The model was built in three steps, beginning with the control variables only (model 1). Next, restless sleep and CMCs were added at both the between- and within-person levels (model 2). Finally, the focal interactions were added to test for moderation (model 3). Non-significant interactions were trimmed for parsimony in the final model. The equation presented below represents the final, trimmed model. Depressive symptoms for the wth wave for the ith person are modeled as:

Level 1: Depressive symptoms$_{wi}$ = β_{0i} + β_{1i} (*chronic medical conditions$_{wi}$*) + β_{2i} (*restless sleep$_{wi}$*) + β_{3i} (*chronic medical conditions*age$_{wi}$*)+ e_{wi}

Level 2: β_{0i} = γ_{00} + γ_{01} (*age$_i$*) + γ_{02} (*gender$_i$*) + γ_{03} (*ethnicity$_i$*) + γ_{04} (*education$_i$*) + γ_{05} (*functional ability$_i$*) + γ_{06} (*chronic medical conditions$_i$*) + γ_{07} (*restless sleep$_i$*) + γ_{08} (*chronic medical conditions*restless sleep$_i$*) + u_{0i}

β_{1i} = γ_{10}

β_{2i} = γ_{20}

β_{3i} = γ_{30}

(Model 3)

where the intercept (β_{0i}) represents the mean level of depressive symptoms for each individual (averaged across waves). The first slope (β_{1i}) represents the effect of chronic medical conditions at wave w (w = 1-5) on the respondent's depressive symptoms. The slope parameter (β_{2i}) represents the association of wave w's restless sleep and (β_{3i}) reflects the association of the age by within-person level chronic medical conditions interaction with depressive symptoms. These covariates reflect deviations from an individual's own mean, and were calculated by centering around the person-mean.

At Level 2, person-mean levels of chronic medical conditions (γ_{06}), restless sleep (γ_{07}), and the interaction between chronic medical conditions and sleep (γ_{08}) are entered as between-person covariates predicting the level of depressive symptoms at the intercept (β_{0i}). Each of these covariates indicates an individuals' average level of that characteristic across waves. Age (γ_{01}), gender (γ_{02}), ethnicity (γ_{03}), education (γ_{04}), and functional ability (γ_{05}) were entered as baseline, between-person controls. γ_{10}, γ_{20} and γ_{30} represent the average effects of within-person slopes from level 1 for chronic medical conditions, restless sleep and the within-person level interaction respectively. γ_{00} reflects the group mean level of individual depressive symptom levels and u_{0i} reflects individual deviations from that mean.

Leggett, Assari, Burgard, Zivin

The effect of sleep disturbance on the association between chronic medical conditions and depressive symptoms over time

Results

Descriptive characteristics. Full sample and cross-wave descriptive results can be found in table 1. At baseline, participants were on average 47 years old (range 24-96), 53% were female, 17% were non-White, and on average they had completed just over 12 years of education. The number of chronic medical conditions (range: 0.68-1.38) increased slightly across the waves whereas the mean depression score (range: 1.27-1.37) and restless sleep (range: 1.57-1.65) remained relatively stable.

Model 1, Demographic characteristics and depressiveSymptoms. An intraclass correlation (ICC) for depressive symptoms revealed that 56% of the variance is within-person (for the same person, across survey waves) and 44% is between-person (across respondents). Significant variability at both levels provides justification for the multi-level modeling approach. Table 2 displays results from our two-level multi-level model with coefficients and standard errors. In Model 1, person-level control adjustments revealed that individuals who were younger (γ_{01}; B= -0.003, $p < .001$), female (γ_{02}; B= -0.043, $p < .001$), non-White (γ_{03}; B= -0.101, $p < .001$), less educated (γ_{04}; B= -0.020, $p < .001$), and with poorer functional ability (γ_{05}; B= -0.106, $p < .001$) reported more depressive symptoms.

Model 2, CMC, sleep, and depressive symptoms. In model 2, we tested between- and within-person associations between the number of chronic medical conditions and restless sleep with depressive symptoms. At the wave level, respondents reported more depressive symptoms at waves when they reported more restless sleep (γ_{20}; B= 0.123, $p < .001$) than their average levels. Turning to the person-level findings, individuals who on average reported more restless sleep (γ_{07}; B= 0.248, $p < .001$) and more CMCs (γ_{06}; B= 0.022, $p < .001$) expressed more depressive symptoms.

Model 3, Sleep and age as moderators. In model 3, we considered whether sleep disturbance moderated the effect of CMCs on depressive symptoms. A within-person level interaction was not significant and thus trimmed from the final model for parsimony. However, a significant interaction at the between-person level was found (γ_{08}; B= -0.017, $p<.05$). As displayed in figure 1, while a higher level of chronic medical conditions was associated with a higher level of depressive symptoms among those who slept well, individuals who had restless sleep had a higher level of depressive symptoms regardless of chronic medical conditions. Additionally, in line with our second research question regarding the association between age and an individual's variability in CMC, we found a significant interaction (γ_{08}; B= 0.001, $p<.001$). This interaction suggests that younger adults had higher symptoms of depression at times of below average CMC relative to older adults. However this difference narrows as older adults' depressive symptoms increase with above average CMC levels (figure 2).

Sensitivity Tests. Finally, prior work has found that different CMCs have varying rates of comorbidity with depression, and sleep may impact an individual's adjustment to a more disabling CMC such as cancer relative with one that may carry less symptomatology such as hypertension. Therefore as a sensitivity test we also tested an interaction between sleep and each individual CMC. Only one interaction was significant at the individual level- heart attack and other heart trouble (B= -0.049, $p<.01$), which is consistent with a large literature showing a strong association between cardiovascular disease and depression. However, examining the interactions individually, as opposed to our sum score reflecting multimorbidity, reduced the power to detect other interactive effects.

Discussion

Using a nationally-representative sample of adults from the ACL study, we found that similar to the findings of previous studies (Baglioni et al., 2011; Maglione et al., 2012), sleep disturbance was associated with depression even controlling for CMC and other risk factors. Additionally, sleep disturbance moderated the association between CMCs and depressive symptoms. In particular, more CMCs were clearly linked to higher depressive symptoms among individuals sleeping well; however among individuals reporting poor sleep, depressive symptoms were higher regardless of CMC.

Sleep disturbance has proven treatments, and improvements in sleep may reduce the strength of association between CMCs and depression, and potentially the symptoms of CMCs. There are several

reasons why the combination of sleep and morbidity may have implications for depression. First, Devins et al. (1993) found that sleep disturbance was associated with the "intrusiveness" of one's CMC into daily life. Greater severity of illness may also pose greater challenge to mental stability or even lead to despair if one is not able to adequately adjust. In addition, both sleep disturbance and CMC may be prodromal symptoms or early risk factors for depressive illness, with the compounding of risk leading to greater likelihood of developing depressive illness over time. In line with this hypothesis, future work should consider the moderating impact on development of depression, in other words controlling for baseline depression status. The underlying, putative mechanisms relating to the development of depression may include inflammation, intrusiveness and necessity to adapt to an illness, stress, reduced functioning, pain or another complication of lack of sleep and morbidity (Devins et al., 1993; Konsman et al., 2002; Parish, 2009; Smith & Haythornthwaite, 2004). However, the compounding of these issues likely increases one's risk for increased depressed mood.

We also found support for a second research question regarding whether age alters the effect of CMC on depression. Younger adults had higher levels of depressive symptoms at waves in which they had fewer CMCs than their personal average across the waves. Though one might expect above average CMCs to be associated with depression, this finding may reflect the high stress that comes with an initial chronic medical conditions diagnosis and the fact that chronic medical conditions are less anticipated at younger ages. As individuals age and as CMCs accumulate, coping strategies and outlooks on the illness context may shift. Younger adults may learn to cope more adaptively over time as the initial shock of an illness fades. In addition, our sensitivity test of the moderation of sleep disturbance on the effects of individual CMCs was likely limited due to a lack of power when examining each illness separately. However, a heart trouble or heart attack interaction with sleep was found suggesting that individuals who are sleeping poorly and have heart trouble are at particular risk for depressive symptoms. Future work examining distinct categories of CMC based on symptomatology, pain, or chronicity in interaction

with sleep disturbance may be fruitful to further unpack these associations.

This study has a number of strengths, including a 25-year longitudinal design that allows for examination of between and within-person effects and a nationally representative sample of adults in the United States. However, several limitations to the current study should be noted. The only assessment of sleep available across the five study waves was the single item from the CES-D scale capturing restless sleep. Additional items measuring adequacy of sleep were only added in the most recent measurement wave. Therefore future work should consider validated sleep scales or objective measures of sleep disturbance in relation to CMC and depression. As restless sleep and depressive symptoms were measured in the same wave as distinctive components of the CES-D index, it follows that these constructs will be correlated. However, our key question was to determine whether sleep might moderate the association between CMC and depressive symptoms over time, which our longitudinal, nationally representative data allowed us to consider. It is important to note, however, that a causal link or mediation between sleep and depressive symptoms should not be implied from our findings. Despite the limited one item assessment, prior work has found it to be valid and associated with other predictors in the same way as more extensive sleep measures (Burgard & Ailshire, 2009; Leggett et al., 2016). It is likely that some attrition across study waves is due to CMC and mental illness; however, a strength of the multi-level design is to use all available data waves and thus not exclude individuals who attrited over time. Finally, it is important to note that as we tested our interaction by calculating the significance of multiplicative terms between CMC and sleep disturbance, we cannot statistically distinguish the main effect from the moderator. However, in our conceptual model, CMC and sleep were the main predictor and moderator, respectively. This decision was based on prior research of sleep as a moderator of several risk factors on psychological outcomes, as well as the modifiable nature of sleep disturbance relative to CMC.

Conclusion

In conclusion, in a nationally representative adult US sample, individuals sleeping poorly were more depressed regardless of CMCs. However, among those sleeping well, CMCs were more clearly linked to increased risk of depression, while those free of CMCs were less depressed. Early detection of sleep disturbance among individuals with CMC and preventative pharmacological or behavioral sleep interventions may help to alter both physical and mental health trajectories toward more positive outcomes over time. Further, Bhattarai et al. (2013) found that morbidities combined with comorbidity of depression led to increasing rates of drug prescriptions, primary care consultations, specialist visits, and in-patient hospital visits. The presence of depression at any level of morbidity was associated with increased healthcare utilisation. Given the high utilisation of the health care system among individuals with CMC, depression, and sleep disturbance, early screening and treatment of sleep disturbance may reduce system costs and have enduring public health benefit.

Acknowledgements

References

Alexopoulos, G. S., Meyers, B. S., Young, R. C., Campbell, S., Silbersweig, D., & Charlson, M. (1997). 'Vascular depression' hypothesis. *Archives of General Psychiatry, 54*(10), 915-922. https://doi.org/10.1001/archpsyc.1997.01830220033006

Anderson, G. (2010). *Chronic care: Making the case for ongoing care*. Retrieved from Princeton, NJ: www.rwjf.org/pr/product.jsp?id-50968

Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., . . . Riemann, D. (2011). Insomnia as a predictor of depression: A meta-analytic evaluation of longitudinal epidemiological studies. *Journal of Affective Disorders, 135*(1–3), 10-19. https://doi.org/10.1016/j.jad.2011.01.011

Berg, C. A., & Upchurch, R. (2007). A developmental-contextual model of couples coping with chronic illness across the adult life span. *Psychological Bulletin, 133*(6), 920-954. https://doi.org/10.1037/0033-2909.133.6.920

Berger, A. M. (2009). *Update on the state of the science: sleep-wake disturbances in adult patients with cancer.* Paper presented at the Oncology Nursing Forum. https://doi.org/10.1188/09.ONF.E165-E177

Berger, A. M., & Mitchell, S. A. (2008). Modifying Cancer-Related Fatigue by Optimizing Sleep Quality. *Journal of the National Comprehensive Cancer Network, 6*(1), 3-13.

Bhattarai, N., Charlton, J., Rudisill, C., & Gulliford, M. C. (2013). Prevalence of depression and utilization of health care in single and multiple morbidity: a population-based cohort study. *Psychological Medicine, 43*(07), 1423-1431. https://doi.org/10.1017/S0033291712002498

Bower, J. E. (2008). Behavioral Symptoms in Patients With Breast Cancer and Survivors. *Journal of Clinical Oncology, 26*(5), 768-777. https://doi.org/10.1200/JCO.2007.14.3248

Brandtstädter, J., & Greve, W. (1994). The Aging Self: Stabilizing and Protective Processes. *Developmental Review, 14*(1), 52-80. https://doi.org/10.1006/drev.1994.1003

Burgard, S. A., & Ailshire, J. A. (2009). Putting Work to Bed: Stressful Experiences on the Job and Sleep Quality. *Journal of Health and Social Behavior, 50*(4), 476-492. https://doi.org/10.2307/20617656

Carpenter, J. S., Elam, J. L., Ridner, S. H., Carney, P. H., Cherry, G. J., & Cucullu, H. L. (2004). *Sleep, fatigue, and depressive symptoms in breast cancer survivors and matched healthy women experiencing hot flashes.* Paper presented at the Oncology Nursing Forum. https://doi.org/10.1188/04.ONF.591-598

Devins, G. M., Edworthy, S. M., Paul, L. C., Mandin, H., Seland, T. P., Klein, G., . . . Shapiro, C. M. (1993). Restless sleep, illness intrusiveness, and depressive symptoms in three chronic illness conditions: Rheumatoid

Leggett, Assari, Burgard, Zivin

The effect of sleep disturbance on the association between chronic medical conditions and depressive symptoms over time

arthritis, end-stage renal disease, and multiple sclerosis. *Journal of Psychosomatic Research, 37*(2), 163-170. http://dx.doi.org/10.1016/0022-3999(93)90083-R

Diehl, M., Coyle, N., & Labouvie-Vief, G. (1996). Age and sex differences in strategies of coping and defense across the life span. *Psychology and Aging, 11*(1), 127-139. https://doi.org/10.1037/0882-7974.11.1.127

Egede, L. E. (2007). Major depression in individuals with chronic medical disorders: prevalence, correlates and association with health resource utilization, lost productivity and functional disability. *General Hospital Psychiatry, 29*(5), 409-416. https://doi.org/10.1016/j.genhosppsych.2007.06.002

Felton, B. J., & Revenson, T. A. (1987). Age differences in coping with chronic illness. *Psychology and Aging, 2*(2), 164-170. https://doi.org/10.1037/0882-7974.2.2.164

Folkman, S., Lazarus, R. S., Pimley, S., & Novacek, J. (1987). Age differences in stress and coping processes. *Psychology and Aging, 2*(2), 171-184. https://doi.org/10.1037/0882-7974.2.2.171

Geerlings, S. W., Beekman, A. T., Deeg, D. J., & Van Tilburg, W. (2000). Physical health and the onset and persistence of depression in older adults: an eight-wave prospective community-based study. *Psychological Medicine, 30*(2), 369-380. https://doi.org/10.1017/S0033291799001890

Hamilton, N. A., Catley, D., & Karlson, C. (2007). Sleep and the affective response to stress and pain. *Health Psychology, 26*(3), 288-295. https://doi.org/10.1037/0278-6133.26.3.288

Heckhausen, J., & Schulz, R. (1995). A life-span theory of control. *Psychological Review, 102*(2), 284-304. https://doi.org/10.1037/0033-295X.102.2.284

Hoffman, L., & Stawski, R. (2009). Persons as Contexts: Evaluating Between-Person and Within-Person Effects in Longitudinal Analysis. *Research in Human Development, 6*(2), 97-120. https://doi.org/10.1080/15427600902911189

House, J. S., Kessler, R. C., & Herzog, A. R. (1990). Age, Socioeconomic Status, and Health. *The Milbank Quarterly, 68*(3), 383-411. https://doi.org/10.2307/3350111

House, J. S., Lantz, P. M., & Herd, P. (2005a). Continuity and Change in the Social Stratification of Aging and Health Over the Life Course: Evidence From a Nationally Representative Longitudinal Study From 1986 to 2001/2002 (Americans' Changing Lives Study). *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 60*(Special Issue 2), S15-S26. https://doi.org/10.1093/geronb/60.Special_Issue_2.S15

House, J. S., Lantz, P. M., & Herd, P. (2005b). Continuity and change in the social stratification of aging and health over the life course: Evidence from a nationally representative longitudinal study from 1986 to 2001/2002 (Americans' changing lives study). *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 60*(Suppl 2), 15-26. https://doi.org/10.1093/geronb/60.Special_Issue_2.S15

House, J. S., Lepkowski, J. M., Kinney, A. M., Mero, R. P., Kessler, R. C., & Herzog, A. R. (1994). The Social Stratification of Aging and Health. *Journal of Health and Social Behavior, 35*(3), 213-234. https://doi.org/10.2307/2137277

Katon, W. J. (2003). Clinical and health services relationships between major depression, depressive symptoms, and general medical illness. *Biological Psychiatry, 54*(3), 216-226. https://doi.org/10.1016/S0006-3223(03)00273-7

Katz, D. A., & McHorney, C. A. (2002). The relationship between insomnia and health-related quality of life in patients with chronic illness. *Journal of Family Practice, 51*(3), 229-236.

Kohout, F. J., Berkman, L. F., Evans, D. A., & Cornoni-Huntley, J. (1993). Two shorter forms of the CES-D (Center for Epidemiological Studies Depression) depression symptoms index. *The Journal of Aging and Health, 5*(2), 179-193. https://doi.org/10.1177/089826439300500202

Konsman, J. P., Parnet, P., & Dantzer, R. (2002). Cytokine-induced sickness behaviour: mechanisms and implications. *Trends in Neurosciences, 25*(3), 154-159. https://doi.org/10.1016/S0166-2236(00)02088-9

Lee, P. G., Cigolle, C., & Blaum, C. (2009). The Co-Occurrence of Chronic Diseases and Geriatric Syndromes: The Health and Retirement Study. *Journal of the American Geriatrics Society, 57*(3), 511-516. https://doi.org/10.1111/j.1532-5415.2008.02150.x

Leggett, A., Burgard, S., & Zivin, K. (2016). The Impact of Sleep Disturbance on the Association Between Stressful Life Events and Depressive Symptoms. *The Journals of Gerontology Series B Psychological Sciences and Social Sciences, 71*(1), 118-128. https://doi.org/10.1093/geronb/gbv072

Lehnert, T., Heider, D., Leicht, H., Heinrich, S., Corrieri, S., Luppa, M., Riedel-Heller, S., & PLEASE ADD ANY MISSING AUTORS König, H.-H. (2011). Review: Health Care Utilization and Costs of Elderly Persons With Multiple Chronic Conditions. *Medical Care Research and Review, 68*(4), 387-420. https://doi.org/10.1177/1077558711399580

Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute Inc.

Maglione, J. E., Ancoli-Israel, S., Peters, K. W., Paudel, M. L., Yaffe, K., Ensrud, K. E., & Stone, K. L. (2012). Depressive symptoms and subjective and objective sleep in community-dwelling older women. *Journal of the American Geriatrics Society, 60*(4), 635-643. https://doi.org/10.1111/j.1532-5415.2012.03908.x

Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V., & Ustun, B. (2007). Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet, 370*(9590), 851-858. https://doi.org/10.1016/S0140-6736(07)61415-9

Ormel, J., Rijsdijk, F. V., Sullivan, M., van Sonderen, E., & Kempen, G. (2002). Temporal and reciprocal relationship between IADL/ADL disability and depressive symptoms in late life. *Journals of Gerontology Series B-Psychological Sciences and Social Sciences, 57*(4), 338-347. https://doi.org/10.1093/geronb/57.4.P338

Parish, J. M. (2009). Sleep-related problems in common medical conditions. *Chest, 135*(2), 563-572. https://doi.org/10.1378/chest.08-0934

Radloff, L. (1977). The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401. https://doi.org/10.1177/014662167700100306

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2 ed.). London, England: Sage.

Sanz-Vergel, A. I., Demerouti, E., Mayo, M., & Moreno-Jiménez, B. (2011). Work–home interaction and psychological strain: The moderating role of sleep quality. *Applied Psychology, 60*(2), 210-230. https://doi.org/10.1111/j.1464-0597.2010.00433.x

Schnell, K., Weiss, C. O., Lee, T., Krishnan, J. A., Leff, B., Wolff, J. L., & Boyd, C. (2012). The prevalence of clinically-relevant comorbid conditions in patients with physician-diagnosed COPD: a cross-sectional study using data from NHANES 1999–2008. *BMC Pulmonary Medicine, 12*(1), 26. https://doi.org/10.1186/1471-2466-12-26

Schnittker, J. (2005). Chronic illness and depressive symptoms in late life. *Social Science & Medicine, 60*(1), 13-23. https://doi.org/10.1016/j.socscimed.2004.04.020

Sherbourne, C. D., Meredith, L. S., Rogers, W., & Ware, J. E. (1992). Social support and stressful life events: age differences in their effects on health-related quality of life among the chronically ill. *Quality of Life Research, 1*(4), 235-246. https://doi.org/10.1007/BF00435632

Smith, M. T., & Haythornthwaite, J. A. (2004). How do sleep disturbance and chronic pain inter-relate? Insights from the longitudinal and cognitive-behavioral clinical trials literature. *Sleep Medicine Reviews, 8*(2), 119-132. https://doi.org/10.1016/S1087-0792(03)00044-3

Taylor, D. J., Mallory, L. J., Lichstein, K. L., Durrence, H., Riedel, B. W., & Bush, A. J. (2007). Comorbidity of chronic insomnia with medical problems. *SLEEP-NEW YORK THEN WESTCHESTER-, 30*(2), 213.

Vogeli, C., Shields, A., Lee, T., Gibson, T., Marder, W., Weiss, K., & Blumenthal, D. (2007). Multiple Chronic Conditions: Prevalence, Health Consequences, and Implications for Quality, Care Management, and Costs. *Journal of General Internal Medicine, 22*(3), 391-395. https://doi.org/10.1007/s11606-007-0322-1

Wielgus, K. K., Berger, A. M., & Hertzog, M. (2009). *Predictors of fatigue 30 days after completing anthracycline plus taxane adjuvant chemotherapy for breast cancer.* Paper presented at the Oncology Nursing Forum. https://doi.org/10.1188/09.ONF.38-48

Table 1. Sample characteristics and comparison of key variables across waves

	Wave 1 (1986)	Wave 2 (1989)	Wave 3 (1994)	Wave 4 (2001/2)	Wave 5 (2011)
Age					
Mean	47.11				
Range	24-96				
SD	16.45				
Female (%)	52.9%				
Education (years)					
Mean	12.36				
Range	0-17				
SD	3.14				
Non-White (%)	16.5%				
Functional ability	N=3617	N=2867	N=2559	N=1785	N=1427
Mean	3.73	3.74	3.66	3.64	3.46
Range	1-4	1-4	1-4	1-4	1-4
SD	0.70	0.70	0.87	0.93	1.18
Chronic medical conditions	N = 3617	N = 2867	N = 2559	N = 1688	N = 1164
Mean	0.68	0.70	0.83	0.88	1.38
Range	0-5	0-6	0-6	0-6	0-5
SD	0.96	0.99	1.06	1.09	1.12
Restless sleep	N = 3585	N = 2852	N = 2390	N = 1682	N = 1316
Mean	1.65	1.61	1.57	1.62	1.59
Range	1-3	1-3	1-3	1-3	1-3
SD	0.70	0.67	0.74	0.78	0.82
CES-D 10	N = 3605	N=2859	N = 2394	N = 1682	N = 1316
Item Mean	1.37	1.34	1.28	1.27	1.28
Range	1-2.9	1-2.9	1-3	1-2.8	1-3
SD	0.35	0.35	0.34	0.35	0.40

Note. Descriptive statistics are weighted.

Table 2. Linear mixed model of demographic and health characteristics, sleep disturbance, and chronic medical conditions on depressive symptoms

	Depressive Mood					
	\multicolumn{2}{c}{Model 1}	\multicolumn{2}{c}{Model 2}	\multicolumn{2}{c}{Model 3}
	B	*(SE)*	*B*	*(SE)*	*B*	*(SE)*
Fixed Effects						
Intercept, γ_{00}	2.21***	(0.038)	1.51***	(0.040)	1.49***	(0.041)
Within-Person Covariates[a]						
Chronic Medical Conditions (CMC), γ_{10}	--	--	-0.002	(0.004)	-0.045***	(0.013)
Restless Sleep, γ_{20}	--	--	0.123***	(0.005)	0.123***	(0.005)
CMC*Age, γ_{30}	--	--	--	--	0.001***	(0.0003)
Between-Person Covariates[b]						
Age, γ_{01}	-0.003***	(0.0003)	-0.002***	(0.0003)	-0.002***	(0.0002)
Male, γ_{02}	-0.043***	(0.009)	-0.021*	(0.008)	-0.020*	(0.008)
White, γ_{03}	-0.101***	(0.011)	-0.093***	(0.010)	-0.092***	(0.010)
Education, γ_{04}	-0.020***	(0.002)	-0.015***	(0.001)	-0.015***	(0.001)
Functional Ability, γ_{05}	-0.106***	(0.007)	-0.058***	(0.006)	-0.060***	(0.006)
CMC, γ_{06}	--	--	0.022***	(0.006)	0.053***	(0.015)
Restless Sleep, γ_{07}	--	--	0.248***	(0.008)	0.267***	(0.012)
CMC*Sleep, γ_{08}	--	--	--	--	-0.017*	(0.008)
Random Effects						
Intercept, σ^2_{u0}	0.041***	(0.002)	0.027***	(0.001)	0.027***	(0.001)
Residual, σ^2_{e1}	0.065***	(0.001)	0.059***	(0.001)	0.059***	(0.001)
-2 Log Likelihood	7951.3		6211.0		6216.6	

Notes. Step 1: Participant *N* = 3,605, Observation *N* = 11,834; Step 2 & 3: Participant *N* = 3,597, Observation *N* = 11,546. [a]Person-mean centered scores (i.e., time-varying). [b]Person-mean across waves (i.e., time-invariant). *$p < .05$. **$p < .01$. *** $p < .001$

Figure 1. Restless sleep as a significant moderator of the chronic medical conditions and depressive symptom association (Between-person level interaction, γ_{08})
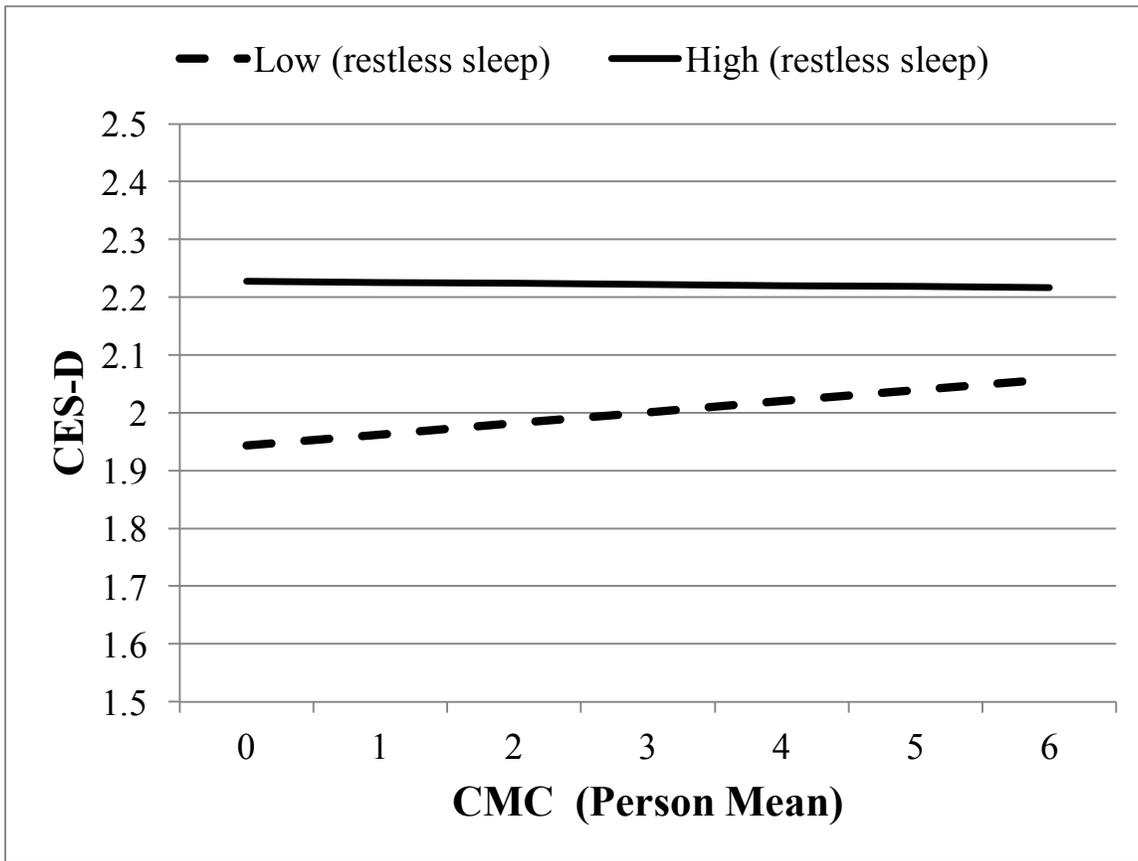
Notes. High and low restless sleep indicating 1 SD above and below the overall sample (between-person level) mean. Chronic medical conditions reflect the summed total range of the sample.

Figure 2. Age as a significant moderator of the chronic medical conditions and depressive symptom association (Within-person level interaction, γ_{30})
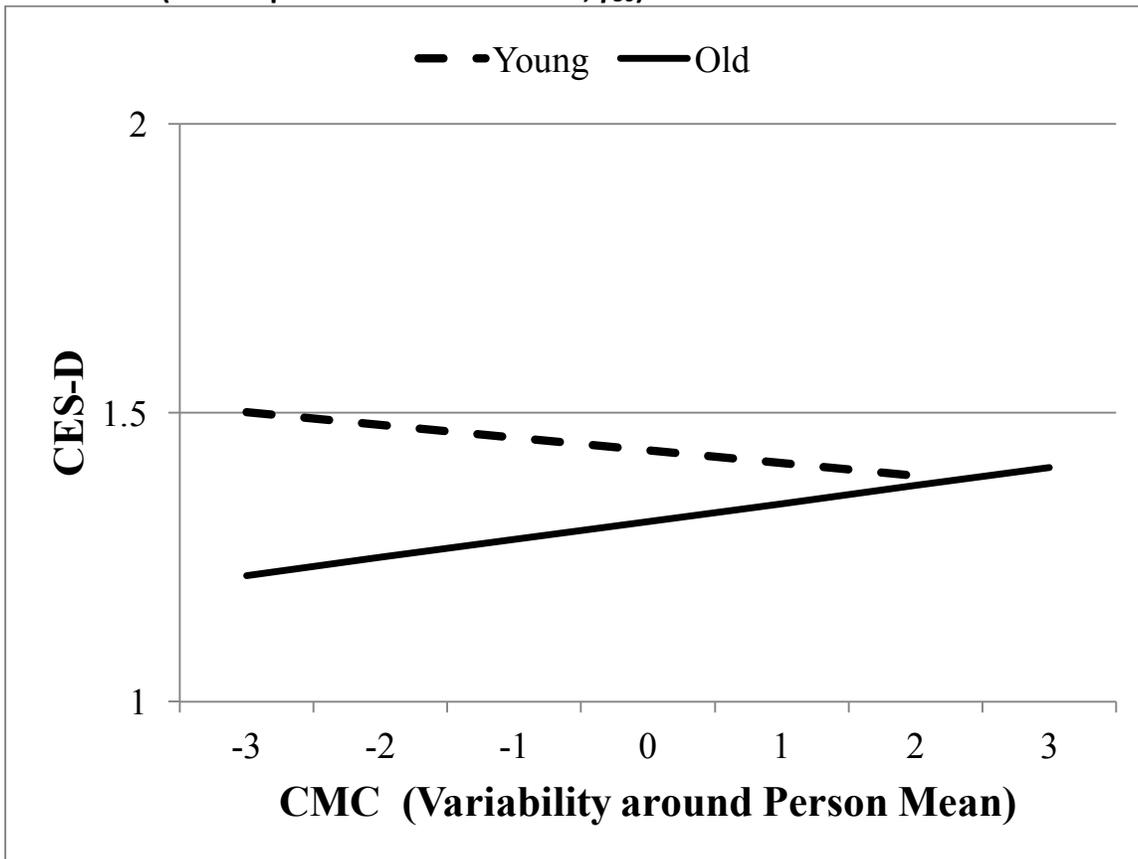
Notes. Young represents 2 SD below the sample mean (age 25) and 2 SD above the sample mean (age 81). Chronic medical conditions variability around an individual's person mean with zero reflecting the person mean, a positive number reflecting an above average wave and a negative number reflecting a below average wave.

Education and civic engagement: A comparative study of the benefits of post-compulsory education in England and Germany

Emma Salter   University of Sussex, UK
e.j.salter@sussex.ac.uk
Angelika Kuemmerling   Institut Arbeit und Qualifikation, University of Duisburg-Essen, Germany
Rod Bond   University of Sussex, UK
Ricardo Sabates   University of Cambridge, UK

Abstract

This paper examines the role of different types of post-compulsory education in determining civic engagement (political interest and election participation) in England and Germany. The educational systems of England and Germany provide ideal comparators for investigating the social benefits of education, in particular those that accrue from vocational education. The paper uses two longitudinal panel surveys, the British Household Panel Survey and the German Socio-Economic Panel for empirical purposes. Contrary to our expectations, our findings revealed few differences between the two countries: the level of political interest is the same for youth who had a vocational degree as those without any further qualifications, in both England and Germany. Similarly, greater levels of interest in politics were observed in adulthood for youth who had achieved academic qualifications in both countries. Likewise voting behaviour in particular was associated with the achievement of academic qualifications in Germany and to some extent with the achievement of mixed vocational and academic qualifications in England.

Keywords

Introduction

A free and democratic state is reliant on the active participation of its citizens. By participating in elections and other civic engagement citizens are able to influence politics; the involvement of all citizens in democratic processes and the extent to which they are enabled to be involved are crucial for ensuring that all citizens have an equal chance of participation and political integration (Weßels, 2008). Individuals with higher levels of political interest are in general more likely to vote (Prior, 2010), therefore the converse follows: in a society of citizens with low or no interest in politics, voting

levels would drop; a situation that is anathema to a democracy (Brennan, Durazzi & Sene, 2013). Thus, it is important for a society to understand the mechanisms that contribute to individuals' propensity for civic engagement.

Since Aristotle's writing in *The Politics and the Constitution of Athens* (Aristotle translated by Everson, 1996) about the importance of education in the creation of a democratic society, there have been debates around the role of education in this process (Grenier & Wright, 2006; Lochner, 2011). Previous research has shown that higher levels of education correlate positively with political interest

and political participation (OECD, 2014; Lochner, 2011). But those studies tend to focus on general (academic) education or years of education (Ogg, 2006; OECD, 2014; de Walque, 2007); much less research has addressed differences depending on the *type* of education. While some researchers have examined the role of vocational education and training in the formation of civic engagement (Sabates et al., 2010), whether vocational education and training has a similar or different role to academic education in forming politically interested youth has yet to be researched. This study seeks to fill in the gaps and broaden our understanding of the relationship between education and civic participation by comparing the links of vocational and academic education with civic outcomes, namely political interest and voting practices in two different educational contexts: England and Germany.

Rationale for the comparative study

Education takes many different forms even across Western Europe, encompassing systems of compulsory and post-compulsory education, vocational education and training (VET) and academic education. The educational contexts of Germany and England are two such examples of contrasting systems, which will enable us to understand more about the role of post-compulsory education in the formation of civic engagement.

Before introducing our theoretical framework, it is important to set out the major differences between the educational systems of these two countries which may lead to differences or similarities in civic engagement. First, the German education system is more stratified according to academic ability than the English system. In Germany, high academic ability pupils are streamed as early as the age of 10 in the Gymnasium, which is a more traditional academic route and enables school leavers to enter university[i]. In England, some degree of stratification exists in particular with wealthier or better-educated parents using the system to their advantage to gain access to the 'better' state schools (Ball, 2003). The A-level results in England, taken around the age of 18, are used by universities as entrance criteria. Since academic ability is associated with higher civic participation (Hillygus, 2005), it is important to account for the different forms of stratification of

the education systems to partially deal with academic ability bias.

Secondly, there are important intergenerational mechanisms which affect both parental selection of education for children (Fleury & Gilles, 2015) and parental influence on children's future civic engagement (Gidengil, Wass & Valaste, 2016). In Germany, the educational background of parents is a strong predictor of the school their children attend. University graduates are more likely to put their children into the Gymnasium, where the Abitur examination is taken at age 19 (approximately), which is a precondition for entering University (Solga & Dombrowski, 2009). In England, there are also large gaps in participation in higher education as measured by socioeconomic disadvantage (Chowdry, Crawford, Dearden, Goodman & Vignoles, 2013). It is likely therefore that family background is a strong predictor of educational outcomes and potentially of young people's civic engagement in both England and Germany.

Finally, there are important differences between the vocational education and training systems of the two countries which are worth highlighting (see Bosch, 2015; Brockmann, Clarke & Winch, 2009 for further discussion). The renowned dual-system of Germany is an example of *Berufliche Bildung*, with strong links between education and industry and clear paths for young apprentices to follow to gain accreditation in their particular field (Rauner, 2006). Vocational education in Germany is both valued and necessary for many occupations, having a broader focus in this respect than the English system and garnering more status as a result (Hillmert, 2002). The curriculum of the VET system in Germany is knowledge-based and designed to develop multi-dimensional competencies, from theory to practice, from civic education to personal development as opposed to the English system which concentrates on a narrow set of skills without a focus on personal development (Brockmann et al. , 2009). Indeed in the early 1990s about two thirds of school leavers (including those with an Abitur) took up an apprenticeship (BIBB, n.d.).

In contrast, the English vocational system is a 'market of qualifications' (Rauner, 2006), where students gain accreditation via a complex array of courses. The system of vocational education and training in England has undergone myriad reforms in the past few decades, with apprenticeship

schemes withdrawn and later reinstated and qualifications offered by a number of non-governmental organisations with little in the way of coherence or standardisation. Primarily these changes have been driven by attempts to raise the status of vocational education and generate more 'parity of esteem' between vocational and academic education, but with little effect (Wolf, 2011). Critics have also suggested that reforms have been more about successive attempts to reduce youth unemployment than in improving the content and structure of vocational education and training (Shavit & Müller, 2000).

Civic participation and education: theory and evidence

While there is consensus that education benefits individuals and society, there has been much debate about how these benefits occur (OECD, 2007). Disciplines as diverse as economics, psychology and sociology have identified key channels through which engaging in education can impact on individuals: social capital, human capital and identity capital. These 'three capitals' have been developed into an integrated framework in relation to the wider benefits of education (Schuller, Preston, Hammond, Brassett-Grundy and Bynner, 2004) and are relevant to the formation of civic participation and engagement as follows.

Firstly, the educational context increases an individual's social capital by introducing them to new peer groups and developing relationships between students and with teachers (Glaeser, Ponzetto & Shleifer, 2007). Recent research by Kawashima-Ginsberg and Levine (2014) indicates that young people who attend racially mixed high schools tend to be less politically engaged later in life. However, the authors suggest that school-based participation in social issues and mixing with peers who share common interests tend to be positively associated with political participation, in particular voting. Secondly, individuals acquire skills and knowledge, in particular civic knowledge (Hillygus, 2005) through the learning process, which can be applied to other areas of their lives, for example, engaging in their community or wider society (Schuller et al., 2004). This is known as human capital and the capabilities that a student gains through education make them more effective citizens, more able to solve problems (Mirowsky & Ross, 2005), and make informed decisions

(Grossman, 2005), which enable them to participate more fully in civic society. Human capital also encompasses the qualifications that people achieve in their educational history, which can translate into economic and social capital, giving those individuals greater access to political processes (Bourdieu, 1986).

Finally, the learning process itself can bring about increased self-esteem and a sense of identity (identity capital), which can create greater social cohesion (Côté, 2005). Identity capital refers to a sense of identity that not only accompanies learning, but is 'inseparable' from it, particularly in an apprenticeship setting where students 'become' a metier as well as learn it (Lave & Wenger, 1991). In this regard researchers found that vocational trainees became included or excluded by the extent to which "they orientate themselves to a vocational habitus" (Colley, James, Diment & Tedder, 2003; p.293). Linked to the concept of identity capital is Erikson's *psychosocial moratorium* referring to a period of time in which late adolescents/emerging adults can work out their identity (Erikson, 1956). The period of time that a student spends at university for example, can provide such an opportunity. However, being engaged in a low status role in the labour market might hinder the process (Schwartz, Côté & Arnett, 2005).

In general terms, there is evidence of similarities between the countries in terms of trends in civic engagement. In Germany, notable differences were found between the younger and older generations, with younger people expressing much lower levels of interest than older members of society. A similar gap was found with regard to participating in elections (Weßels, 2008). Likewise research in the UK has shown that older people were more politically interested than the under 35s at each educational level (Ogg, 2006). There are also within-generation differences (Keating, Green and Janmaat, 2015), in terms of class and gender: women were less likely to be interested in politics than men and two thirds of people in higher social classes expressed an interest in politics compared with less than a fifth of those in the lower classes (Grenier and Wright, 2006).

With regard to the evidence on the relationship between education and civic participation, higher levels of education in the UK have been linked to greater levels of interest in politics (Brennan et al., 2013; OECD, 2010; Ogg, 2006). Siedler (2010) found

no positive associations of education on political interest in Germany, whereas the Autorengruppe Bildungsberichterstattung (2014) describes profound differences between people with a university degree and those with no further education with regard to both political interest and voting. The 2014 edition of *Education at a Glance* reported that a third of people in Germany with tertiary education felt they had a say in government, compared to 13% who completed up to grade 10 (around age 16) of secondary schooling. In the UK, the figures were 42% and 22% respectively (OECD, 2014). In 2012, the OECD cited a 50% 'voting gap' between the proportions of tertiary educated people aged 25-34 in Germany who said they would vote and those who completed up to grade 10 of secondary education. The gap in the UK was smaller but still sizeable at 32% (OECD, 2012).

The majority of research in this area has focussed on the role of years of study or highest academic qualification. Studies commissioned by the European Centre for the Development of Vocational Training (CEDEFOP) examined the role of VET for individuals, firms and nations. The outcome of one of such studies on the benefits of VET for adults aged 16 and over revealed that across Europe, positive increases in civic participation, as measured by higher proportion of individuals reporting membership of civic organisations, was greater for adults who had undertaken vocational education and training than for those who were not in education (Sabates et al., 2010). However, the authors found that when the same analysis was undertaken at the country level, results showed that these benefits of VET on civic participation were found only in Southern European countries (Greece, Italy, Spain and Portugal) as well as in Austria and Sweden, but not in Germany or the UK. A further limitation of the empirical literature is the fact that a longitudinal dimension is lacking from the research to date, as is any controlling for confounding factors except income and gender when undertaking comparative research across countries. Our analysis addresses these gaps, by using longitudinal panel data, differentiating by type of post-compulsory education and controlling for a range of background measures.

From theory, we propose that education has the potential to increase an individual's capital (social, human and identity), which may result in increased social status, which in turn enables individuals to access and engage with political processes in society (OECD, 2010). However, the impact of education on people's capitals is tempered by factors within the system of education in which it takes place (Shavit & Müller, 1998). In a cross-European study, the beneficial effects of vocational education and training were shown to be dependent on the national context; civic participation outcomes were likely to be enhanced by vocational education in countries where these were rewarded or acted as a safety net for the unemployed (Sabates, Salter & Obolenskaya, 2012). Likewise, greater civic engagement and a stronger interest in politics was discovered among German vocational trainees, (Fouad & Bynner, 2008; Bynner, 2011) than their English counterparts.

A comparative analysis of outcomes within the differing national contexts of England and Germany enables us to propose hypotheses about the role of education in building social capital, which can be realised through civic participation and engagement. Against the outlined theoretical background, it can be hypothesised that an apprenticeship model such as the German dual-system provides a better milieu for increasing identity capital than the British model of vocational education and training, which in comparison has been more piecemeal. Despite the higher status of vocational education and training in Germany, in both countries, an accredited academic qualification still confers a greater amount of social capital, in terms of social status and access to political institutions (Feinstein, Hammond, Woods, Preston & Bynner, 2003). On this basis we would expect improved outcomes for civic participation for all types of education, but within this we would expect more nuanced results: an academic education will be associated with higher levels of political interest in both countries, but vocational education and training will be associated with higher levels of civic engagement in Germany than in England.

Thus our main research questions are:
1. Does the type of post-compulsory education (that is academic education, vocational training or a combination of both) influence civic outcomes later in the life course?

2. Do different national educational systems result in different outcomes? Specifically, does academic or vocational education result in different outcomes with regard to political interest or engagement in England and Germany?

Hence, while the aim of the study is to investigate the role of both vocational and academic education on civic outcomes, key to its success is the comparative nature of the analysis. The selection of these two countries, and their differing educational context, allow us to understand whether the post-compulsory education context is important in developing an interest in politics.

Data and variables

We used the British Household Panel Survey (BHPS) and the German Socio-Economic Panel (SOEP) survey to study the benefits of different post-compulsory educational trajectories on civic participation compared with those who left education after completing compulsory schooling. The BHPS started in 1991, initially interviewing 10,264 individuals in 5,511 households in Great Britain (Taylor, Brice, Buck & Prentice-Lane, 2010). The SOEP commenced in 1984 with an initial sample of 12,245 individuals in 5,921 households (Kroh & Spieß, 2008) in both East and West Germany. The main reason for selecting these datasets was their comparability in terms of structure, time period and measures. Additionally, both datasets are rich in information on educational qualifications enabling us to construct detailed educational trajectories for all individuals. The BHPS started in 1991 and we chose the same reference year for Germany, thus confining the observations from Germany to the post re-unification years.

In order to obtain the cohort of young people with the longest history in the datasets, we included young people aged between 16 and 21 years old in 1991 in England and between 17 and 21 in Germany[ii], following them annually until 2010, when they were aged between 35-40 years old. This enables us to look at the original sample members of the datasets, that is, those that were interviewed in the initial wave, with the advantage that they show less attrition than their peers who entered the survey in subsequent waves. The BHPS sample was restricted to include young people who

studied in the English education system to make neater cross-country comparisons between England and Germany. In order to minimise the amount of missing data, the sample for analysis excluded people who dropped out of the survey before they reached twenty years of age. Our final sample for estimation purposes comprised 732 and 878 young people in England and Germany, respectively.

Explanatory variable: Educational trajectories

The explanatory variable in the model is the type of education that the young people completed during their transition to adulthood. The longitudinal nature of the dataset enabled us to construct educational trajectories for the young people in each country using three steps. Firstly, we coded the annual information they gave regarding their post-compulsory educational qualifications as either academic or vocational following the International Standard Classification of Education so that qualifications in England and Germany were comparable. Secondly, we constructed educational trajectories for respondents in each country, taking into account the differences outlined earlier between the two systems in terms of compulsory school leaving age and in age of graduation. The compulsory school leaving age in England in the period under study was 16, while in Germany it was 18. In Germany, before the Bologna Process came into force, the average age at graduation from university was 28, while in the UK it was as early as 21, with some degrees taking a few years more (OECD, 2002)[iii]. Consequently, the information used for constructing trajectories for the English cohort were between the ages of 16 and 25 and those for the German cohort between 17 and 28 years of age.

Within the age range between post-compulsory schooling and completion of higher education, educational qualifications achieved during a specific year were classified as either academic or vocational. In doing so, we encounter repeated episodes of either vocational or academic qualifications. The final step was to use this longitudinal data on educational qualifications to categorise the young people's educational trajectories into one of four types: academic only, vocational only, mixed (people who had achieved both academic and vocational education and training in any order) and no further education. Figure 1 shows the proportion of young people who comprised these categories in each country.

Figure 1: Proportion of respondents in each educational category in England (n=732) and Germany (n=878)
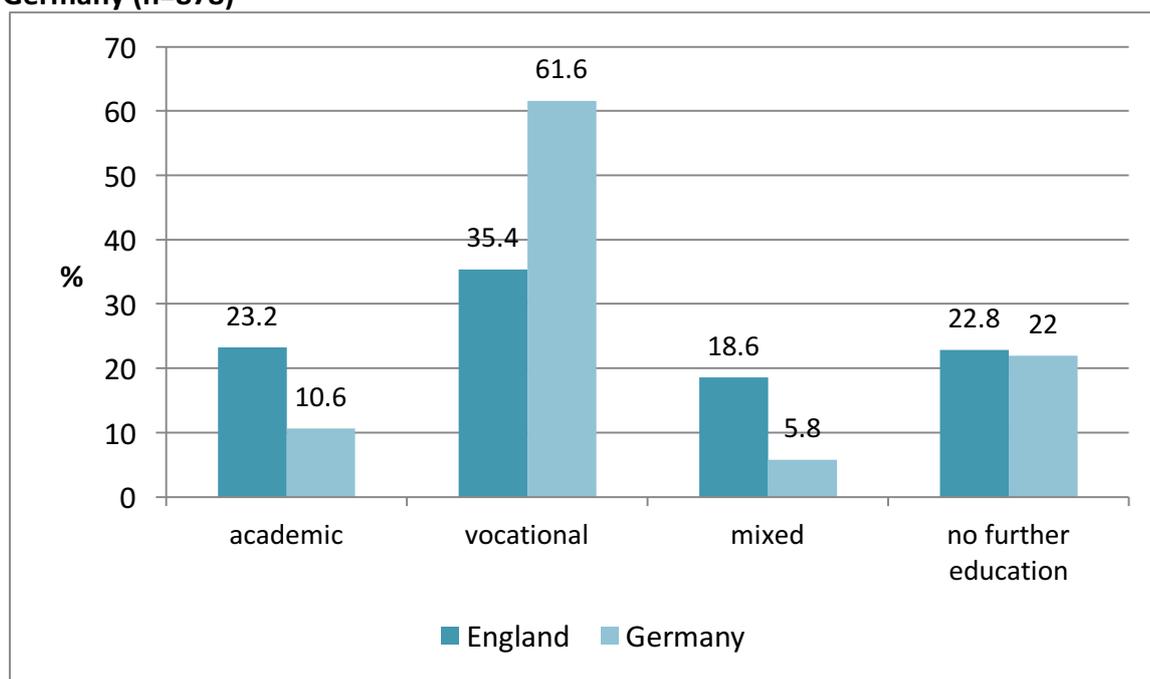
Figure 1 reveals already profound differences in the allocation of the educational categories between the two countries, thus reflecting the different educational systems. In the nineties, in England the proportion following the academic track was more than twice as high as in Germany; on the other hand the proportion with vocational training was much higher in Germany. In both samples, however, one in five respondents left school without following further education.

Dependent variables: Interest in politics and voting

In line with previous research civic engagement was operationalized by two measures of civic participation: political interest and voting.

Political interest was measured by the question 'how interested are you in politics?' with four possible responses from 'very interested' to 'not interested'. The responses were reduced to a three-point scale in both datasets as the number of respondents in the category 'very interested' was

too small. The outcome was measured at age 25 and age 35 to examine the immediate and long-term roles of education.

Voting was measured in the general election years in both countries (1997 and 2005 in England and 2005 and 2009 in Germany) with the questions: "did you vote in the last election?" for the BHPS, which captures actual voting behaviour earlier in the year, and "would you vote if there was an election on Sunday?" in the SOEP, which captures expectation to vote. Table 1 shows descriptive statistics for these outcome variables in Germany and England. Political interest is generally higher in England than in Germany; the proportion expressing fairly or very interested in politics is about ten per cent higher. The proportion of people who voted is about 16-18% higher in England than in Germany, however, since one survey measured intent to vote and the other actual voting behaviour, this difference may be due in part to the nature of the questions posed.

Table 1: Descriptive statistics of outcome variables measuring civic engagement in BHPS and SOEP (%)

|		Germany		England	
Level of political interest		Age 25	Age 35	Age 25	Age 35
fairly or very interested		26	30	36	41
not very interested		60	53	34	30
not interested at all		14	17	30	29
Voting		2005	2009	1997	2005
Would you vote? (Germany)	Yes	42	45	n.a.	n.a.
Did you vote? (England)	Yes	n.a.	n.a.	58	63

Source: SOEP & BHPS.

Controls: Background factors

We introduced a number of socio-demographic factors into our models in order to see whether any possible statistical associations of post-compulsory education were driven by background factors occurring earlier in the life course. The socio-demographic variables included out of theoretical considerations and in line with previous research were parental education and parental occupational class and prior education (Plutzer, 2002). A small difference in the parental occupational class variable is that in Germany we used father's occupational class as the majority of mothers in the German sample had been housewives when occupational class was measured. Prior education enables us to capture prior ability to some degree (Hillygus, 2005), and it was measured in England by GCSE attainment (five or more GCSEs at grades A-C) and in Germany by whether or not students participated in Gymnasium. All regressions include gender as a factor to differentiate between differences in political interest and voting between young men and young women (Schoon, McCulloch, Joshi, Wiggins & Bynner, 2001). Year of birth, coded as a set of dummy variables, was included as a control in all years since the sample spans six birth years, to control for any cohort effects.

Analysis

Three linear models were run: firstly, the outcomes at age 35 were regressed on post-compulsory education, controlling only for gender and birth year, to see whether there were any differences by type of education. The second model analysed the relationship of post-compulsory education on outcomes at age 35, with the inclusion of political interest at age 25, to investigate if the relationship of education were visible at an earlier time point, again, controlling only for gender and birth year. Thirdly, the final model (see figure 2) was the same as model 2 with the addition of background controls, to see if these had any effect on the relationship between education and political interest. In addition, we tested for the significance of the indirect associations of further education on the outcome at age 35 via the outcome at age 25 using bootstrapped standard errors (Mackinnon, 2008). In all models, the post-compulsory education category of 'no further education' was the reference category; that is, each model compared the outcomes for those who had continued in any of the three routes through post-compulsory education with those who had left the education system.

Figure 2: Model 3
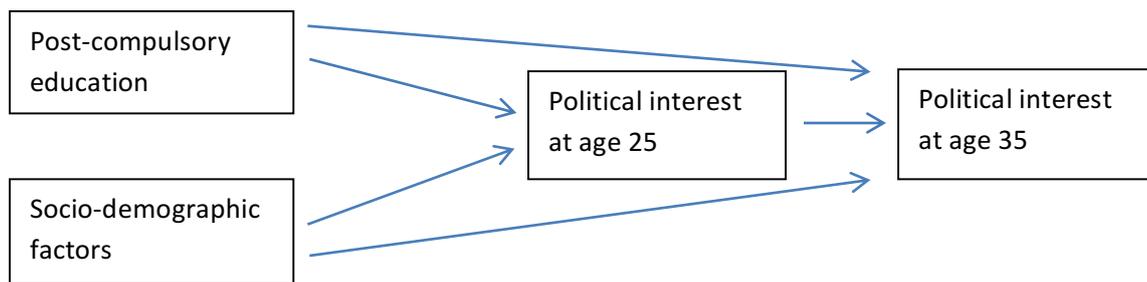
All analyses were run in MPlus (Muthén & Muthén, 1998-2015). In order to deal with missing data in the two datasets, multiple imputation was used. One advantage of multiple imputation is that we can include more variables in the model to derive imputed values than are included in our substantive models, and therefore include variables that cause missingness but are not of substantive interest. For instance, in the BHPS, household income and scores on the general health questionnaire (a measure of psychological wellbeing) between ages 21 to 35 were included as potential predictors of missingness. We imputed twenty datasets as this is deemed a sufficient number (Enders, 2010; Schafer & Graham, 2002). Both political interest and voting were treated as categorical dependent variables and models used a robust weighted least squares estimator, with bootstrapped standard errors based on 2000 bootstrapped samples.

Results

Political interest

In England, model 1 in table 2 shows that individuals who gained academic qualifications or a mixture of academic and vocational qualifications reported higher levels of interest in politics at age 35 than their peers who left school at 16 without qualifications. Model 2 shows, however, that academic education and mixed education are associated with greater political interest at age 25 and that direct effects of type of education on political interest at 35 are no longer significant. Model 3 shows that with the introduction of background controls, individuals with academic education and a mixture of academic and vocational qualifications remain statistically significantly associated with higher levels of political interest at 25 and indirectly at 35 than individuals who left school at 16 without qualifications. In none of the models shown in table 2 do we find statistical differences in the levels of political interest between individuals who achieved vocational qualifications and individuals who left school at age 16. Finally, table 2, model 3 shows that men are more likely to be interested in politics at age 35 and age 25 compared with women and that parental education and educational attainment at age 16 were statistically significant predictors of political interest at age 35, but not at age 25.

Table 2: Linear panel results for effects of post-compulsory education on political interest in England

	Model 1	Model 2	Model 3
Effect on political interest at 35			
Academic	0.606**	0.117	-0.112
Vocational	-0.067	-0.087	-0.113
Mixed	0.609**	0.314	0.148
Gender = male	0.323**	0.469**	0.474**
Political interest at 25		0.553**	0.559**
Parental education			0.288*
Parental social class			-0.165
5+ GCSEs A-C			0.371*
Year controls	Yes	Yes	Yes
Mediated effects via political interest at 25			
Academic			0.441**
Vocational			0.022
Mixed			0.266*
Effect on political interest at 25			
Academic		0.886**	0.716**
Vocational		0.039	0.039
Mixed		0.531*	0.428*
Gender=Male		0.277*	0.282*
Parental education			-0.134
Parental social class			0.223
5+ GCSEs A-C			0.238
Year controls	Yes	Yes	Yes

Source: BHPS. Notes: Asterisks *, ** represent statistical significance at 5%, 1% respectively.

In Germany, model 1 in table 3 shows that individuals at age 35 with an academic education reported a greater interest in politics than those with no further education. Similarly, individuals with mixed educational qualifications reported greater interest in politics at age 35 compared with individuals with no further educational qualifications. Model 2 shows that there is a relationship between individuals with academic and mixed qualifications and higher levels of political interest at age 25, but there is also a direct effect of academic qualifications on political interest at age 35. Vocational trainees were not more likely to be interested in politics at age 25 than those who did not have further educational qualifications. Model 3 in table 3 shows that with the inclusion of background factors having an academic education and having mixed educational qualifications still have direct associations on levels of political interest at age 25 and indirectly with levels of political interest at age 35. In Germany, we also find that men are more likely to be politically interested at age 35 and at age 25 than women and that parental education and achievement of a Gymnasium are statistically significantly associated with political interest at age 25 only.

Table 3: Linear panel results for effects of post-compulsory education on political interest in Germany

	Model 1	Model 2	Model 3
Effect on political interest at 35			
Academic	0.663**	0.294*	0.254*
Vocational	0.011	-0.024	-0.039
Mixed	0.387*	0.036	-0.008
Gender = male	0.398**	0.221*	0.227**
Political interest at 25		0.507**	0.498**
Parental education			-0.008
Parental social class			0.074
Gymnasium			-0.013
Year controls	Yes	Yes	Yes
Mediated effects via political interest at 25			
Academic			0.302**
Vocational			0.002
Mixed			0.206**
Effect on political interest at 25			
Academic		0.727**	0.570**
Vocational		0.070	0.003
Mixed		0.694**	0.520**
Gender=Male		0.350**	0.345**
Parental education			0.132*
Parental social class			0.095
Gymnasium			0.495**
Year controls	Yes	Yes	Yes

Source: SOEP. Notes: Asterisks *, ** represent statistical significance at 5%, 1% respectively.

In sum, there are similarities between the two countries: in both England and Germany, people who have an academic or mixed trajectory were more interested in politics than their peers who left education, while vocational trainees had similar levels of interest to their peers who have no further education. Additionally, in both countries, prior educational achievement in the form of GCSE attainment in England and attending Gymnasium in Germany is predictive of higher levels of political interest, although this is apparent earlier in adulthood, at age 25 in England, and later, at around age 35 in Germany. In both countries men reported, on average, higher levels of political interest than women.

Voting

Table 4 shows results for the likelihood of voting in the 2005 elections in England. Model 1 shows that individuals with academic or mixed qualifications were more likely to vote in the 2005 elections than individuals who left education at age 16. Individuals with vocational qualifications were not statistically different in their likelihood to vote in the 2005 elections from individuals who left education at age 16. Similar results with respect to educational attainment and likelihood to vote in the 1997 elections were found in model 2, and that educational attainment did not have a direct effect on voting in 2005 once voting in 1997 was taken into account. In model 3, with the inclusion of

background factors, we found that the statistical associations between education and likelihood of voting in the 1997 or 2005 elections were mostly conditioned out, and only a positive association between having mixed education and voting in 1997 relative to having no further education after the age of 16 remained statistically significant. We found that the inclusion of previous voting

behaviour fully mediated the association between educational qualifications and voting in 2005 whereas the inclusion of prior educational attainment, achieving five or more A-Cs at GCSE, conditioned out the association between the attainment of academic qualifications and voting in the 1997 elections

Table 4: Linear panel results for effects of post-compulsory education on voting in England

	Model 1	Model 2	Model 3
Effect on voting in 2005			
Academic	0.420*	0.128	-0.099
Vocational	0.007	-0.070	-0.083
Mixed	0.539**	0.235	0.079
Gender = male	-0.228	-0.145	-0.130
Voting in 1997		0.574**	0.555**
Parental education			0.006
Parental social class			0.015
5+ GCSEs A-C			0.340*
Year controls	Yes	Yes	Yes
Mediated effects via voting in 1997			
Academic			0.138
Vocational			0.053
Mixed			0.198
Effect on voting in 1997			
Academic		0.509**	0.232
Vocational		0.134	0.089
Mixed		0.531**	0.333*
Gender=Male		-0.145	-0.134
Parental education			0.016
Parental social class			0.086
5+ GCSEs A-C			0.295*
Year controls	Yes	Yes	Yes

Source: BHPS. Notes: Asterisks *, ** represent statistical significance at 5%, 1% respectively.

Finally, table 5 presents results for likelihood of voting in Germany. Model 1 shows that individuals with academic or mixed qualifications were more likely to vote in 2009 than their peers who had left education. We again find that vocational trainees were similar in their likelihood to vote in 2009 compared with individuals who had left education. When voting in 2005 is added into the model, we

continue to find a statistical association between academic or mixed qualifications and likelihood to vote in 2009 relative to individuals without post-compulsory qualifications (model 2, table 5). But only individuals with academic qualifications were more likely to vote in 2005 compared with individuals who had left education. Finally, when we added background factors, model 3 shows a direct

association between achievement of academic qualifications and mixed qualifications with likelihood to vote in 2009, relative to no educational qualifications; a direct association between achievement of academic qualifications with likelihood to vote in 2005, relative to no educational qualifications; and an indirect association of achievement of academic qualifications and likelihood to vote in 2009. Attending Gymnasium was significantly associated with likelihood to vote in 2009 and 2005 whereas parental education was significantly associated with likelihood to vote in 2005.[iv]

In the case of voting therefore, there are differences between the two countries. While post-

compulsory education had limited associations with individuals' voting habits in England, academic or mixed trajectories were associated with a greater likelihood of voting in Germany. In both countries, vocational trainees exhibited similar voting practices to their contemporaries who had no further education. As with political interest, prior educational attainment had a positive association on voting habits early and later in adulthood in both England and Germany. In neither country did we find the parameter for gender to be statistically significant, which indicates that there are no differences in likelihood to vote between men and women in either England or Germany.

Table 5: Linear panel results for effects of post-compulsory education on voting in Germany

	Model 1	Model 2	Model 3
Effect on voting in 2009			
Academic	0.896**	0.450*	0.519*
Vocational	0.057	0.063	0.047
Mixed	0.841*	0.635*	0.707*
Gender = male	-0.058	0.012	0.024
Voting in 2005		0.552**	0.547**
Parental education			0.034
Parental social class			-0.108
Gymnasium			0.487*
Year controls	Yes	Yes	Yes
Mediated effects via voting in 2005			
Academic			0.345*
Vocational			-0.120
Mixed			0.051
Effect on voting in 2005			
Academic		0.704**	0.571*
Vocational		-0.107	-0.196
Mixed		0.220	0.088
Gender=Male		-0.200	-0.178
Parental education			0.323**
Parental social class			-0.157
Gymnasium			0.475**
Year controls	Yes	Yes	Yes

Source: SOEP. Notes: Asterisks *, ** represent statistical significance at 5%, 1% respectively.

Discussion and conclusions

Our research set out to investigate whether the type of post-compulsory education (that is either academic education or vocational training) followed by young people influenced their civic outcomes later in the life course and whether these associations differed in different educational systems. The results of our analyses show that the type of post-compulsory education *does* matter in terms of people's civic outcomes in adulthood. In both countries, people who followed an academic or mixed trajectory through post-compulsory education were more interested in politics and more likely to vote than their peers who had left education.

We hypothesised that both types of youth education would increase civic engagement through the development of the three capitals in the learning process and that the dual system of VET in Germany was more likely to have an effect on civic participation than VET in England. Our analysis instead found that vocational education does not increase young people's interest in politics or likelihood of voting compared to those who only complete compulsory education. While this was half-expected in the English context, it was an unexpected result for the dual-system of Germany.

While cross-sectional studies have repeatedly found associations between increased levels of education and increased levels of political interest, such studies do not take a longitudinal perspective and thus are unable to investigate the role of education in affecting civic outcomes at different points in life as well as the role of individual and family background factors in conditioning out the relationship of education to civic outcomes. We find that much of the initial association of post-compulsory education on civic outcomes later in life is explained by either the direct or indirect relationship of education on civic outcomes earlier in life. Gender impacts on political interest (men reported, on average, higher levels of political interest than women) but does not influence the likelihood to vote in both countries. The inclusion of family background variables, as well as previous educational attainment, contributes to conditioning out some of the differences between the attainment of post-compulsory academic qualification and greater civic participation. Nonetheless, these factors are not sufficient to condition out the potential academic ability bias,

that is, academic ability simultaneously affecting educational attainment and political participation (Hillygus, 2005).

Our theoretical model aids in anchoring our results based on the hypothesis proposed. First, we expected to find an association between vocational education and civic participation in Germany due to the potential benefits of the dual system in terms of identity capital and thus increased civic participation. The fact that those who pursue vocational education and training in both countries are no more politically interested than those who do not gain post-compulsory qualifications questions to what extent this is due to the nature of individuals undertaking these qualifications as much as to the content of such qualifications. Despite the differences outlined earlier in the paper between the two educational systems, there are clear selection mechanisms in both countries, where parental background and parental education are highly predictive of the route that young people pursue in post-compulsory education. This raises the question of whether the kind of vocational education and training provision can readdress the potential consequences that socioeconomic disadvantage has on lack of civic and political engagement.

Secondly, our consistent results on the positive associations between the attainment of academic qualifications and civic engagement in both countries could be partially explained by the benefits of education due to increased skills and knowledge, in particular civic knowledge, which is more likely to be provided in academic environments. We acknowledge the limitation that it is not possible to ascertain from the results the specific content of academic provision, which is the main contributor to the level of political interest of those taking that route.

A further limitation is our inability to deal with bias due to omitted factors since we did not have the relevant data on unobservable variables such as social and cognitive skills or personality traits which are associated with educational attainment and with civic participation. The nature of our study, however, was not intended to measure causality, but to propose direction of associations and estimation of parameters using structural equations. A lack of data on civic interest at an earlier time point, such as at age 16, precluded us from seeing if there was an effect of pursuing

vocational education and training which we were not able to capture in the current study.

The fact that vocational education was not associated with civic outcomes in two different environments but academic qualifications or a mixture of academic and vocational qualifications were associated with these outcomes in both countries therefore suggests two possibilities: either there is a lack of civic interest in the type of people pursuing vocational education which vocational education and training does not redress

or that vocational education and training does not currently promote such interest, in either country. The inclusion of citizenship education at school-level and within vocational education and training in both countries may increase civic participation and interest, and as a result foster more democratic and cohesive societies. Nonetheless, further research is required to understand how the educational experience can enhance a more politically sensitive and civically engaged society.

Acknowledgements

References

Aristotle translated by Everson, S. (1996). *The Politics and the Constitution of Athens.* Cambridge, UK: Cambridge University Press.

Autorengruppe Bildungsberichterstattung (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: Bertelsmann Verlag.

Ball, S.J. (2003). *Class Strategies and the Educational Market: the middle classes and social advantage*. London: Routledgefalmer. https://doi.org/10.4324/9780203218952

BIBB (n.d). Datenbank Auszubildende" des Bundesinstituts für Berufsbildung auf Basis der Daten der Berufsbildungsstatistik der statistischen Ämter des Bundes und der Länder (Erhebung zum 31. Dezember), Berichtsjahre 1993 bis 2010, und Bevölkerungsfortschreibung des Statistischen Bundesamtes, Berichtsjahre 1993 bis 2010. Absolutwerte aus Datenschutzgründen jeweils auf ein Vielfaches von 3 gerundet. Berechnungen des Bundesinstituts für Berufsbildung

Bosch, G. (2015). Different national skill systems. In K. Mayhew, C. Warhust (Eds.), *Oxford Handbook of Skills and Training*. Oxford: Oxford University Press (in press).

Bourdieu, P. (1986). The Forms of Capital. In Richardson, John G. (Ed.) *Handbook of Theory and Research for the Sociology of Education.* New York: Greenwood.

Brennan J., Durazzi, N. & Sene, T. (2013). *Things we know and don't know about the Wider Benefits of Higher Education: A review of the recent literature*. BIS Research Paper Number 133. London: Department for Business Innovation and Skills.

Brockmann, M., Clarke, L & Winch, C. (2009). Difficulties in recognising vocational skills and qualifications across Europe. *Assessment in Education: Principles, Policy & Practice* 16, *Special Issue: The Challenge of Cross-Border Qualifications Recognition*, 97-109. https://doi.org/10.1080/09695940802704153

Bynner, J. (2011). Youth transitions and apprenticeships: A broader view of skill. In T. Dolphin & T. Lanning (Eds.), *Re-thinking apprenticeships* (pp.17-28). London: IPPR.

Chowdry, H., Crawford, C., Dearden, L., Goodman, A. and Vignoles, A. (2013). Widening Participation in Higher Education: Analysis using Linked Administrative Data. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176* (2), 431–457. https://doi.org/10.1111/j.1467-985X.2012.01043.x

Colley, H., James, D., Diment, K. & Tedder, M. (2003). Learning as becoming in vocational education and training: class, gender and the role of vocational habitus. *Journal of Vocational Education & Training*, *55*(4), 471-498. https://doi.org/10.1080/13636820300200240

Côté, J. E. (2005). Identity capital, social capital and the wider benefits of learning: generating resources facilitative of social cohesion. *London Review of Education, 3*(3), 221-237. https://doi.org/10.1080/14748460500372382

de Walque, D. (2007). Does education affect smoking behaviors? Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics, 26*, 877–895. https://doi.org/10.1016/j.jhealeco.2006.12.005

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York: Guilford Press.

Erikson, E.H. (1956). The problem of ego identity. *Journal of the American Psychoanalytic Association*, *4*, 56-121. https://doi.org/10.1177/000306515600400104

Feinstein, L., Hammond, C., Woods, L., Preston, J. & Bynner, J. (2003) *The Contribution of Adult Learning to Health and Social Capital,* Wider Benefits of Learning Research Report No. 8. London: Centre for Research on the Wider Benefits of Learning, Institute of Education.

Fleury, N. & Gilles, F. (2015). A meta-regression analysis on intergenerational transmission of education: publication bias and genuine empirical effect. *Travail, Emploi et Politiques Publiques (TEPP) Working Paper 2015-02*, Paris, France: TEPP.

Fouad, N. & Bynner, J. (2008). Work Transitions. *American Psychologist, 63*(4), 241-251. https://doi.org/10.1037/0003-066X.63.4.241

Gidengil, E., Wass, H. & Valaste, M. (2016). Political Socialization and Voting. The Parent–Child Link in Turnout. *Political Research Quarterly*, *69* (2), 373-383. https://doi.org/10.1177/1065912916640900

Glaeser, E.L., Ponzetto, G.A.M. & Shleifer, A. (2007). Why does democracy need education? *Journal of Economic Growth, 12*, 77-99. https://doi.org/10.1007/s10887-007-9015-1

Grossman, M. (2005). *Education and Non-market Outcomes*. NBER Working Paper 11582. Cambridge, MA: National Bureau of Educational Research. https://doi.org/10.3386/w11582

Grenier, P. & Wright, K. (2006). Social Capital in Britain: Exploring the Hall Paradox, *Policy Studies*, *27* (1), 27-53. https://doi.org/10.1080/01442870500499900

Hillmert, S. (2002). Labour Market Integration and Institutions: An Anglo-German Comparison. *Work Employment Society, 16* (4), 675-701. https://doi.org/10.1177/095001702321587424

Hillygus, S. (2005). The Missing Link: Exploring the relationship between higher education and political engagement. *Political Behaviour*, *27* (1), 25-47. https://doi.org/10.1007/s11109-005-3075-8

Kawashima-Ginsberg, K. & Levine, P. (2014). Diversity in classroom: The relationship between deliberative and associative opportunities in schools and later electoral engagement. *Analyses of Social Issues and Public Policy*, *14* (1), 394-414. https://doi.org/10.1111/asap.12038

Keating, A., Green, A. & Janmaat, G. (2015). *Young Adults and Politics Today: disengaged and disaffected or engaged and enraged?: The latest findings from the Citizenship Education Longitudinal Study (CELS).* Working Paper.

Kroh, M. & Spieß, M. (2008). Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2007). DIW Berlin Data documentation No 39.

Lave, J. & Wenger, E. (1991). *Situated Learning*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Lochner, L. (2011). *Non-productive benefits of education: Crime, Health and Good Citizenship.* NBER Working Paper 16722. Cambridge, MA: National Bureau of Economic Research.

MacKinnon, D.P. (2008). *Introduction to statistical mediation analysis.* New York: Erlbaum.

Mirowsky, J. & Ross, C.E. (2005). Education, learned effectiveness and health. *London Review of Education, 3* (3), 205-220. https://doi.org/10.1080/14748460500372366

Muthén, L.K. & Muthén, B.O. (1998-2015). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.

OECD (2002). *Education at a Glance*. Paris: OECD Publishing. https://doi.org/10.1787/eag-2002-en

OECD (2007). *Understanding the Social Outcomes of Learning.* Paris: OECD Publishing.

OECD (2010). What are the social outcomes of education? In OECD, *Education at a Glance 2010: OECD Indicators*. Paris: OECD Publishing. https://doi.org/10.1787/eag-2010-en

OECD (2012) *Education at a Glance.* Paris: OECD Publishing.

OECD (2014) *Education at a Glance.* Paris: OECD Publishing.

Ogg, J. (2006). A Brief Profile of the New British Establishment. *The Political Quarterly, 77* (1), 81-89. https://doi.org/10.1111/j.1467-923X.2006.00783.x

Plutzer, E. (2002). Becoming a habitual voter: inertia, resources, and growth in young adulthood. *American Political Science Review*, *96*(1), 41–56. https://doi.org/10.1017/S0003055402004227

Prior, M. (2010). You've Either Got It or You Don't? The Stability of Political Interest over the Life Cycle. *The Journal of Politics*, *72*(3), 747-766. https://doi.org/10.1017/S0022381610000149

Rauner, F. (2006). Berufliche Bildung—die europäische Perspektive. In Grollmann, P., Spöttl, G. & Rauner, F. (eds) *Europäisierung Beruflicher Bildung— eine Gestaltungsaufgabe.* Hamburg: Lit Verlag, 127–153.

Sabates, R., Salter, E., Bromberg, T., Voss-Dahm, D., Obolenskaya, P. & Bosch, G. (2010). *Social Benefits of Vocational Education and Training for Individuals: Concepts, Contexts and Empirical Results.* Thessaloniki, Greece: CEDEFOP.

Sabates, R., Salter, E. & Obolenskaya, P. (2012). The social benefits of initial vocational education and training for individuals in Europe. *Journal of Vocational Education & Training*, *64*(3), 233-244. https://doi.org/10.1080/13636820.2012.691530

Salter, E., Kuemmerling, A., Bond, R. & Sabates, R. (2015). *Post-compulsory education: what are the benefits? A comparative study of the social outcomes of education in England and Germany.* Brighton: Centre for International Education, University of Sussex.

Schafer, J.L. & Graham, J.W. (2002) Missing data: our view of the state of the art. *Psychological Methods, 7*, 147-177. https://doi.org/10.1037/1082-989X.7.2.147

Schoon, I., McCulloch, A., Joshi, H., Wiggins, R. & Bynner, J. (2001). Transitions from school to work in a changing social context. *Young,* 9 (1): 4-22. https://doi.org/10.1177/110330880100900102

Schuller, T., Preston, J., Hammond, C., Brassett-Grundy, A. & Bynner, J. (2004). *The Benefits of Learning: The impact of education on health, family life and social capital*. London: RoutledgeFalmer.

Schwartz, S.J., Côté, J.E. & Arnett, J.J. (2005). Identity and agency in emerging adulthood: Two developmental routes in the Individualization Process. *Youth Society*, 37 (2): 201-229. https://doi.org/10.1177/0044118X05275965

Shavit, Y. & Müller, W. (1998). *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations*. Oxford: Oxford University Press.

Shavit, Y. & Müller, W. (2000). Vocational Secondary Education: Where diversion and where safety net? *European Societies, 2* (1*)*, 29-50. https://doi.org/10.1080/146166900360710

Siedler, T. (2010). Schooling and citizenship in a young democracy: evidence from post-war Germany. *Scandinavian Journal of Economics, 112* (2), 315-338. https://doi.org/10.1111/j.1467-9442.2010.01604.x

Solga, H. & Dombrowski , R. (2009). Soziale Ungleichheiten in schulischer und außer-schulischer Bildung. Arbeitspapier 171.

Taylor, M. F., Brice, J., Buck, N. & Prentice-Lane, E. (2010) *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices.* Colchester: University of Essex.

University of Essex. Institute for Social and Economic Research, *British Household Panel Survey: Waves 1-18, 1991-2009* [computer file]. *7th Edition.* Colchester, Essex: UK Data Archive [distributor], July 2010. SN: 5151.

Weßels, B. (2008). Politische Integration und politisches Engagement. In: Statistisches Bundesamt (Eds.), *Datenreport 2008. Ein Sozialbericht für die Bundesrepublik Deutschland (pp. 363-369)*. Bonn: SFG Servicecenter Fachverlage.

Wolf, A. (2011) *Review of Vocational Education – The Wolf Report.* London: Department for Education.

Endnotes

[i] As in Germany education is organized at a regional level, the transition from primary school to the Gymnasium is not standardized nationally. In some regions teachers' recommendation (based upon grades) is decisive, in others it is parental choice.

[ii] The SOEP is only conducted with adults aged 17 and above, so the German sample consists of 17-21 year olds in 1991 instead.

[iii] It should be noted that graduate students in Germany in the 1990s usually left university with a full master degree or its equivalent.

[iv] Models for voting remain unchanged if we include political interest at 25 as control instead of voting in the previous election. We opted for voting in previous election as this is an observed behaviour rather than an attitude.

Wage differentials after a career break: A latent growth model using Belgian register data

Dimitri Mortelmans   University of Antwerp, Belgium
dimitri.mortelmans@uantwerpen.be
Dorien Frans   University of Antwerp, Belgium

Abstract

This article analyses income differentials after re-entry into the labour market between people who have had a career break and people who have not by applying latent growth modelling to a sample of longitudinal register data. The results suggest that when comparing the incomes of those who return from a break with those who did not have a break there are significant initial income differences to the disadvantage of the former. Moreover, the income differentials between men were greater than those between women. In addition, significant additional income growth was found after the break for women but not for men. The evidence suggests that such leave is more socially acceptable for women but leads to significant negative income differentials among men.

Keywords

Introduction

In Belgium, the government-supported Career Break Scheme allows employees to interrupt their career temporarily while receiving a limited substitute income. The underlying idea was to create more space in the labour market for the unemployed. Over the years, the scheme gradually shifted from economic to more individual-oriented measures such as thematic leave (i.e. parental leave, leave for medical support or leave for palliative care). In the international literature, little is known about the consequences of these types of career break schemes. It is generally believed that employees who interrupt their careers will be penalised in the labour market. Some studies have shown the negative effects of taking a career break on career

development in terms of fewer promotions and wage depreciation in comparison to individuals with continuous working patterns (e.g. Spivey, 2005). This article contributes to the literature by focusing on wage differentials after a career break. Using longitudinal register data, we will analyse wage differentials between those who took a break and a control group that did not interrupt their career during the same period.

Background to the time credit scheme in Flanders

The Belgian Career Break Scheme was introduced in 1985 in response to the high unemployment rate, offering employees a limited substitute income during the period of work interruption. In 2002, the

programme was modernised to become a general Time Credit Scheme in the private sector. The new scheme was promoted as a way to increase quality of life, with a specific emphasis on a work-life balance. The Time Credit Scheme enjoys great popularity and the number of people entering the system has increased rapidly over the years: from 156,584 in 2003 (3.4% of the total labour force) to 276,301 in 2014 (5.5% of the total labour force). The majority of people taking a career break (65%) are located in the northern part of the country (the Dutch-speaking region of Flanders), where the regional government provides additional financial incentives to take a career break.

There are several options that people can choose from when taking a career break: full-time or part-time breaks, and specific thematic leave. A full-time career break allows employees to interrupt their careers and cease work activities completely for a limited period of time for whatever reason. In the private sector such a career break is called 'Time Credit'. The distinction between the private and public sectors and the accompanying name change was introduced in 2002 when the system of career breaks was restructured. However, apart from the name and extended duration in the public sector (six years rather than five), both systems are similar. In 2003, less than 8,500 people took a full-time career break. After the reforms, the number of full-time breaks remained rather low in both sectors and even diminished slightly over the years. In 2015, nearly half of the people (43.6%) taking a full-time career break, whether in the public or the private sectors, were between 25 and 40 years old.

The second option is to take a partial career break. Each employee is free to choose the extent to which he or she wants to reduce their workload, whether by 50%, 33%, 25% or 20%. We will not make a distinction between these categories in this article. A partial reduction in working hours is by far the most popular option across all types of career break. The popularity of the part-time scheme in the private sector has grown spectacularly since its introduction in 2002. A part-time career break is especially significant for people aged 50 years or older, as it gives them the opportunity to work part-time until their retirement. This partially explains why there is a low labour market participation rate for people over 50 in Belgium. Almost 66% of those who took a part-time break in 2010, either in the public or private sectors, were 50 years or older.

The third option involves specific categories of career breaks that can be taken based on predefined conditions. Among these 'thematic' forms of leave are three important types: parental leave, leave for medical assistance or leave to undertake palliative care. Of these three possibilities, parental leave is by far the most popular (85%), while leave for medical assistance accounts for less than 15% and leave for palliative care is very limited, at less than 1%.

The majority of people who take a career break choose to do so on a part-time basis. In 2014, almost 64,000 people (1.3% of the total labour force) took a part-time career break, either in the private or the public sectors. There are also important differences between men and women in the take up rates. More than 66% of people taking a career break are women. This female majority becomes even more apparent among the full-time breaks and thematic leave (77% and 72%, respectively). However, this percentage is lower among the part-time breaks (63%), where the part-time options are gradually becoming more attractive to men. This increase seems to be primarily due to the rise in part-time breaks among men who are 50 years or older.

Previous research has also shown that the intentions behind the use of a break are gender related. Desmet, Glorieux, and Vandeweyer (2007), for example, showed that the main motive for taking a break for women was childcare (especially those with young children), while for men, ending their career on a part-time basis was clearly the most dominant reason. Nevertheless, some men also use the scheme to try out a new job or to start their own business.

Literature review: Consequences of career breaks

A work-life balance is a priority in a series of EU policy initiatives on childcare, the amount of time people work and the need for flexible arrangements. The EU considers work-life balance policies to have a direct effect on labour participation and quality of life (Eurofound, 2004). Flexible arrangements such as the Career Break Scheme are considered useful

instruments to achieve a better work-life balance (Hyman, 2005). Moreover, the involvement of fathers in childcare has been shown to increase when paternal leave is available (Bünning, 2015; Nepomnyaschy & Waldfogel, 2007).

There are indications from previous research that there are potential unwanted negative consequences on subsequent wages due to career breaks (e.g. Arun, Arun, & Borooah, 2004). A common element in all these studies is the focus on maternal and paternal leave. These studies looked at the influence of leaving the labour market after a first or subsequent birth and consistently showed that mothers experience a wage penalty (Gangl & Ziefle, 2009; Waldfogel, 1998a, 1998b). While this wage gap is partly due to differences between the background characteristics of women, a considerable part can also be attributed to occupational segregation (Manning & Petrongolo, 2008). For men, there is little research data on this topic. The dominant strand in recent literature looks at the effect of paternal quotas on subsequent earnings. Both in Scandinavia (Norway, Sweden) and in Germany, legislation encourages fathers to take parental leave after the birth of their first child (Geisler & Kreyenfeld, 2011). The effects of uptake on earnings were found to be negative in Norway (Rege & Solli, 2013) and negative but insignificant in Sweden (Ekberg, Eriksson, & Friebel, 2013). For Germany, studies only looked at uptake but not wage consequences (Trappe, 2013). Thus, most studies point to clear income losses as a consequence of a career break. Immediately after a career interruption, wages will be relatively lower than before (Mincer & Ofek, 1982).

The effect of career breaks also seems to be dependent on the timing and the length of the break. Corcoran and Duncan (1979), for example, found that there was only a negative effect on wages when the break occurred at the beginning of a career. They also found evidence of recovery from an initial income decrease after the break, as wages increased more rapidly subsequently (Corcoran, Duncan, & Ponza, 1983). Other studies found that longer interruptions could lower wages. In addition, short interruptions had a rather small or no impact on further career development compared to longer interruptions (Mincer & Ofek, 1982; Schönberg & Ludsteck, 2007; Theunissen, Verbruggen, Forrier, & Sels, 2009).

A career break leads to a significant wage loss during the break because the substitute income is much lower than the wage being paid. Moreover, a career break not only brings the accumulation of human capital to a halt during the break, but can also lead the existing human capital to deteriorate. This can have a negative effect on the productivity of such employees (Becker, 1985). According to human capital theory, wages and job opportunities reflect the productivity of individuals. This productivity increases as education, training and job experience increase (Becker, 1964). Therefore, it can be assumed that a career break interrupts the accumulation of work experience and can even lead human capital to deteriorate during the break. As a consequence, it can lead to a worker having a lower productivity and may reduce their career prospects. This may explain the relative lower wages paid to people on their return to work after a career break (Gangl & Ziefle, 2009). Our first hypothesise, therefore, is that after a career break, we will find *a lower income for those who took a break compared to those who worked without a break* (H1).

The wage loss experienced immediately after the break is followed by a recovery period. This recovery period is a phase where previously eroded human capital might be restored (i.e. a catch-up effect). Given that such a process would be quicker and less costly than building up new human capital, this process and the wage growth will eventually slow down and continue to grow at the rate of a worker without a career break (Mincer & Ofek, 1982). During this recovery period, we hypothesise that the wage growth of people who had taken a career break will be higher than those who did not take a break because there would be a catch-up effect in relation to work experience. In the long run, human capital theory expects wage losses to be compensated. Therefore, our second hypothesis is that *the income growth rates will be significant and positive for people with a career break compared to people without a career break* (H2).

In some studies, the loss of human capital could not explain all of the wage differentials. The wage losses resulting from career breaks were found to exceed losses due to the work experience foregone (i.e. an indicator of human capital) (Beblo & Wolf, 2002). Moreover, some studies found gender

differences in the effects of career breaks on income that suggest other explanations are important because the loss in human capital can be assumed to be equal among men and women (Albrecht, Edin, Sundstrom, & Vroman, 1999). These studies refer to signalling theory to explain these gender differences, which suggests that a temporary career break can send a signal to employers that such people are less committed to their job and prioritise their family over their work. This signal can play an important role in decisions regarding promotions where higher commitment is expected. People who interrupt their career are, in fact, stigmatised for doing so (Beblo & Wolf, 2002). Men, in particular, are seen to be less committed to their jobs if they take a break. Employers suggest that they would respond with lower wages. Women, however, are traditionally seen as the ones to take a break for family reasons. As a consequence, they do not send the equivalent signals to their employers when taking a career break. The smaller wage loss found among women can thus be explained as a loss in human capital, while the wage loss for men is attributed more to the stigmatisation process (Albrecht et al., 1999).

Men having a career break may therefore be expected to be penalised more in the labour market than women, with the wage differentials between men with and without a break expected to be higher than the wage differentials between women with and without a break. Since a career break is generally used by women, it is suggested that this is already reflected in the lower wages of women compared to men (Gangl & Ziefle, 2009). Men, however, are not expected to take a career break and would more likely be stigmatised for doing so. Therefore, our third hypothesis is that *an interaction effect between gender and career break is significant, where the initial income differential after the break between men with or without a break is higher than the income differential between women with or without a break* (H3).

Since most of the research in this field concentrates on first births and maternity leave, little is known about potential age effects. When looking at the uptake of parental leave schemes, human capital is used to explain the uptake of leave by younger partners in a couple. This relative resources hypothesis states that the partner who has the least

work experience will lose less by taking a break (Geisler & Kreyenfeld, 2011). Since our study uses data across the life course, and births are not the prime focus, we need to take age into account. We do not follow the relative resources hypothesis in our study, as we consider workers in different age groups rather than age differences within a couple. Since older workers take leave as a pathway to retirement, we consider the age of 50 as a crucial turning point in a Belgian career. Although the retirement patterns of men and women are different (De Preter, Van Looy, & Mortelmans, 2014), we can take the same age as a turning point for both men and women in the labour market. However, we do not consider the age effect to be independent of gender, with women taking leave more when they become a parent, and men considering the system to provide an end of career strategy. Therefore, our fourth hypothesis is that *the interaction effect between gender and career break is age dependent, with the initial income differential after the break between younger men higher than the income differential between older men with or without a break* (H4).

In general, the empirical evidence is not unequivocal with regard to the financial consequences of career breaks. The actual effect of a career break on subsequent income remains unclear, neither do the results show a clear picture of the underlying processes determining income trajectories after a career break. There are possible selection effects. For example, employees who have taken a career break might have followed a similar development in their career path even if they had not taken this break; in other words, employees might have less upward career mobility regardless of a career break because they do not have the ambition or an interest in building a 'career'. People motivated to achieve the highest levels in their career will probably work full-time, work over-time and never interrupt their careers. Someone prioritising family life, however, is more likely to work fewer hours (part-time if possible), stay at home to look after a sick child and take up options such as a career break. As far as the data allows us, we will thus take the potential selection effects into account in our models (cf. Ekberg et al., 2013).

In summary, previous studies are often limited in scope. Almost all previous research relates to the

gender gap in wage differentials and focuses primarily on the effects of parental leave for women (e.g. Beblo & Wolf, 2002; Dex, 1999; Pylkkanen & Smith, 2004). Few studies have examined the differences between men and women or taken a male perspective (e.g. Ekberg et al., 2013; Trappe, 2013). In this article, we want to go beyond these shortcomings by examining the impact of a career break on the income of both men and women across the life course. Additionally, we will use a large sample of reliable longitudinal data to estimate the effects, rather than survey measures of income.

Data and methodology

Data

Longitudinal data from a representative sample of workers taking a career break during the observation period was required to test these hypotheses. We obtained such a sample from the Datawarehouse Labour Market and Social Security of Belgium. This database combines data from various social security institutions in the country. The sample population consists of people living in Flanders who were active in the labour market in the year of sampling. The data at our disposal spans the period from the second quarter of 1998 to the fourth quarter of 2006. A great advantage of such register data is the accurate and complete income data. We can examine the income differentials after a break as well as the differentials in subsequent income growth. We also have information regarding the length and number of career breaks in this period, and the sample also provides enough statistical power for our models, whereas data on career breaks in general-purpose surveys would usually be numerically insufficient to test our hypotheses.

A total of 90,414 individuals were randomly selected, and the sample was drawn on a yearly basis. For 2000 to 2006, a 10% sample of all individuals taking a career break (in each of these seven years) was taken. In order to be able to compare the incomes and income growth, a 'control group' was created, with the same number of people working without having had a career break randomly selected each year (also totalling 90,414 individuals). This control group was sampled to match the career interrupters in terms of age and gender for every quarter in the first sample.

For the purposes of this article, we only used people taking one career break (leaving approximately 80% of the original sample: 69,023). In order to model the influence of the income trajectory before and after the break, we limited the sample to breaks ending in 2003. This decision further limited the sample to 5,537 people. From the control sample (initially 67,906 individuals), we randomly selected a 2003 sample of 8,259 individuals, who were all working in that year. The selection of the 2003 subsample was done for several reasons. Firstly, the subsample allowed us to control for possible periodic effects. Because the individual cases are spread over the years, possible income differences could be due to the different times at which the breaks were observed. Secondly, the selection allowed us to examine part-time career breaks that extended for four or five years. We were also able to include the new career break scheme (from 2002), albeit for a short period of time (maximum one year).

Measures

Income is the dependent variable in the multilevel models presented below. We used the gross income measured on a quarterly basis. For each quarter, the incomes were adjusted for inflation, with 2004 as the base year. To linearise the individual income trajectories and to address the ways in which income data is commonly skewed, the natural logarithm of income was modelled in our analyses (ln_inc). For the interpretation of the results, income was translated back into the original scale in euros (by taking the antilog).

Other variables available from the register data were the presence of children in the household (introduced as a time-varying dummy in the model), age (age^2 was also included in the models) and a time-constant dummy indicating whether the individual was younger or older than 50 when taking the career break. For the control group, this dummy registers whether the person was younger or older than 50 in the first quarter of 2003.

Analytic strategy

The research questions were examined using multilevel models of change, also known as longitudinal growth models (Singer & Willet, 2003).

This is a better way to examine the data than traditional regression analyses because longitudinal data can be seen as multilevel data with repeated measures nested within individuals. This also allowed us to be more flexible in the assumptions imposed on the models. The classic assumptions in OLS regression analyses, i.e. independent and identically distributed residuals, and homoskedastic variance across occasions and individuals, are less credible in longitudinal data (Singer & Willett, 2003).

Our model is essentially a multilevel model with two levels. On the first level, we located the repeated measurement of different variables within one person. As such, we postulated a level one model that represents the expected change in income level within one individual. This is also referred to as the individual income trajectory. On the second level, we located individual-specific variables to explain some of the variability in income trajectories between individuals.

We were interested in the change in income before, surrounding and after the career break. Because the Datawarehouse Labour Market and Social Security of Belgium data spans the period from the second quarter of 1998 to the fourth quarter of 2006, we were able to estimate the income trajectory over a 35-quarter time-span. This observational window is represented in figure 1 with the arrow labelled 'Quarter'. In the tables, we will refer to this general time trajectory as TIME1.

Figure 1

Time Variables that constitute the Individual Income Trajectory (example with career break between quarter 1 and 5)

Due to the event of a career break, we expected a discontinuous income trajectory rather than a linear one, which was modelled using the following basic variables: 'Break' (TIME2), 'Return' (TIME3) and 'Recovery' (TIME4). This regression specification is in line with the work of Kletzer and Fairlie (2003) but adapted to this particular time trajectory (see also: Couch & Placzek, 2010; Jansen, Mortelmans, & Snoeckx, 2009). Figure 1 illustrates the parameterisation of time.

In the unconditional growth model – the model without any covariates except for those mentioned above – the intercept estimates ln_inc at Time zero, the start of the observation window for all respondents irrespective of the event of a career break. As 'Quarter' is a continuous time variable, the estimate for the slope of the curve associated with the general growth indicates the linear trend in income during the administrative panel. 'Break' (TIME2) is essentially a time-varying dummy variable, which has the value zero for all employees not interrupting their career throughout the observation window. For the respondents with a career break, it changes to one in the quarter that the break was

taken and continues to have the value one during the whole post-break period. The break effect thus assesses the impact of the career break based on the income level in the quarter the break was taken. For all those not taking a career break, it is assigned zero throughout the observation window.

'Return' (TIME3) is a comparable parameterisation for the quarter in which a career break is ended. It receives the value one in the quarter a career break ends and continues to have the value one until the end of the observation window. Again, non-interrupters have a zero-value during the whole panel. The slope associated with 'Recovery' (TIME4) assesses the linear growth in income trajectory in the post-break period. This variable is given the value one, one year after returning from a career break and increases from there on. We refer to this effect as the income growth after the break because the slope gives an insight into the deviation for individuals who took a break compared to the general income trajectory of all individuals. For example, a positive slope indicates a larger increase in income growth, compared to the general income growth throughout the sample.

Table 1. Taxonomy of multilevel models for change

Level 1/Level 2 specification

	Level 1 model	Level 2 model	Composite model
Model 1	$Ln\_y_{ij} = \pi_{oi} + E_{ij}$	$\pi_{oi} = \gamma_{00} + \zeta_{oi}$	$Ln\_yij = \gamma_{00} + (\zeta_{oi} + E_{ij})$
Model 2	$Ln\_y_{ij} = \pi_{Oi} + \pi_{1i}*TIME1_{ij} + \pi_{2i}*TIME2_{ij}$ $+ \pi_{3i}*TIME3_{ij} + \pi_{4i}*TIME4_{ij}$ $\Pi_{5i} * SELECTION_{ij} + E_{ij}$	$\pi_{Oi} = \gamma_{00} + \xi_{0ij}$ $\pi_{1i} = \gamma_{10} + \xi_{1ij}$ $\pi_{2i} = \gamma_{20} + \xi_{2ij}$ $\pi_{3i} = \gamma_{30} + \xi_{3ij}$ $\pi_{4i} = \gamma_{40} + \xi_{4ij}$ $\pi_{5i} = \gamma_{50} + \xi_{5ij}$	$Ln\_y_{ij} = \gamma_{00} + \gamma_{10}*TIME1_{ij} + \gamma_{20}*TIME2_{ij}$ $+ \gamma_{30}*TIME3_{ij} + \gamma_{40}*TIME4_{ij} + \gamma_{50}*SELECTION_{ij}$ $+ (\xi_{0ij} + \xi_{1ij}*TIME1_{ij} + \xi_{2ij}*TIME2_{ij} + \xi_{3ij}*TIME3_{ij}$ $+ \xi_{4ij}*TIME4_{ij} + \xi_{5ij}*SELECTION_{ij} + E_{ij})$

We also modelled the instantaneous impact of gender, age and other covariates on the income level in the quarters during and following the break, as well as the impact of these covariates on the deviation of the general linear growth rate. In order to achieve this, we added interaction terms between the

explanatory variables and the break, return and recovery effect, respectively.

Finally, we also included a time-constant dummy variable indicating whether the respondent did take a career break in this sample. This strategy allowed us to account for possible differences between those

who do and those who do not interrupt their career before the actual break is taken (this is called the selection effect in figure 1). Alternatively, selection effects could be ruled out using a fixed effects approach. Like the random effects model we propose here, fixed effects models deal with the problem of dependence. Such models focus on explaining the within-person variability only by controlling for all stable characteristics of respondents, observed or not. By adopting such an approach, we would, by definition, lose all time-constant explanatory variables. Looking for a balance between controlling for unobserved heterogeneity and insight into the processes at hand, we believe the multilevel model of change, corrected for selection, is a sound alternative to a fixed effect model.

Findings

Sample characteristics

The sample of people taking a career break was compared to the control group in 2003 with regard to certain characteristics (see table 2). We looked at both groups according to age and gender. Apart from the fact that there are more people in the control group (8,259 vs 5,537), we can see that the distribution is comparable in both groups. In the career break group, data on age was missing in six cases. We can see from the table that most men are 50 years or older (54.67% and 52.93%, respectively) and there are slightly more women of 50 years or older in the career break group (23.97%) compared to the control group (16.35%). According to the household typology LIPRO, in 2003 (results not shown), we find that the career break group consists of slightly more married people than the control group, at 19.0% and 16.0%, respectively. There are also slightly more married people with children among the career break group (53.3%) compared to the control group (45.4%). The control group, however, consists of more single people (11.0%) than the career break group (6.9%). The other categories are rather low in both groups.

Table 2. People with a career break and people without a career break across age and gender

	With Career Break						Without Career Break					
	Men	%	Women	%	Total	%	Men	%	Women	%	Total	%
< 25	66	4.75	336	8.10	402	7.26	146	5.64	505	8.91	654	7.88
25-49	564	40.58	2817	67.93	3381	61.06	1073	41.43	4237	74.74	5310	64.29
>= 50	760	54.67	994	23.97	1754	31.68	1371	52.93	927	16.35	2298	27.82
Total	1390	100.00	4147	100.00	5537	100.00	2590	100.00	5669	100.00	8259	100.00

In table 3, we can see how our sample of people taking a career break is divided across the age categories, and the type of career break taken by men and women. Most people taking a career break are female (4,147 or 74.9%). There are considerable differences across the type of career break. Most men take a part-time break, especially in the public sector. Moreover, the part-time breaks taken by men are primarily situated among those aged 50 years or older, while the full-time options are most likely to occur among the middle-aged category (25-49 years).

For men, thematic leave is also primarily taken by those who are middle-aged. Most women also take a part-time career break. While this type of break is also popular among women aged 50 or older, contrary to men, the majority of those taking a part-time break lie in the middle-aged category. Significantly fewer women opt for a full-time break, with the main full-time users being middle-aged. Thematic leave is also very popular among women, with 89.04% belonging to the middle-aged category.

Table 3. Gender and age differences in each type of career break

		Part-time (public)	%	Full-time (public)	%	Thematical	%	Part-time (private)	%	Full-time (private)	%	Total	%
Men	< 25	22	3.99	8	6.06	5	2.36	13	4.5	18	8.78	66	4.75
	25-49	94	17.03	74	56.06	191	90.09	73	25.26	132	64.39	564	40.58
	>= 50	436	78.98	50	37.88	16	7.55	203	70.24	55	26.83	760	54.68
	Total	552	100.00	132	100.00	212	100.00	289	100.00	205	100.00	1390	100.00
Women	< 25	172	8.35	28	6.50	94	7.92	20	8.70	22	9.17	336	8.10
	25-49	1145	55.58	291	67.52	1056	89.04	132	57.39	193	80.42	2817	67.93
	>= 50	743	36.07	112	25.98	36	3.04	78	33.91	25	10.41	994	23.97
	Total	2060	100.00	431	100.00	1186	100.00	230	100.00	240	100.00	4147	100.00

These findings are in accordance with the descriptive statistics presented above. It seems that the public sector is overrepresented among those taking a career break, and particularly among individuals using the part-time option. From our analysis in the previous section, we saw that this ratio (public/private) changes throughout the years, where part-time breaks in the private sector become more important than the public sector and that thematic leave also becomes more popular among men.

The unconditional means model

In the analyses, we will begin by fitting two simple models: the unconditional means model and the unconditional growth model. These unconditional models partition and quantify the outcome variation in two important ways: the first model describes the amount of variance on each level across people regardless of time, and the second model does so across both people and time by including time as a predictor, allowing us to examine income growth at re-entry to the labour market after a break. These models allow us to determine whether there is any systematic variation in outcome which is worth exploring and where the variation is located (between or within people) (Singer & Willett, 2003). Both models also provide baselines for subsequent comparisons of other models to which we add substantive predictors, such as career break type and gender, on the second level to explain the variance between individuals.

The unconditional means model (model 1 in table 4) does not describe change in the outcome over time. There are no predictors included in this model. In this null model, only variations in intercept are allowed. The income trajectory (ln_inc or ln_y_{ij}) of a certain individual is flat since there are no slopes. This model only gives time-constant averages. The average of ln_y_{ij} for individual i is π_{oi}. The average of ln_y_{ij} for all individuals is γ_{OO}. E_{ij} is the variance on level one, which is the spread around the person-specific ln_inc level over the quarters. ζ_{oi} is the variability of the average of ln_inc between persons on level two that can possibly be explained by introducing covariates on this level.

In the null model, the average ln_y_{ij} is 8,356 throughout the observational time for both the control group and the group of people who took a career break. The average income is €4,256 ($e^{8,356}$)/quarter or €1,419/month during the whole time period. At time j, ln_y_{ij} varies from the average of individual i (π_{oi}) by E_{ij}. All parameters are statistically significant. We therefore reject the null hypothesis for each group (p < 0.001) that the average income of people between 1998 and 2006 is zero.

A null model is always estimated first because it describes and partitions the outcome variation. The residual on level one represents the within person (σ_E^2) variation and the residual on level 2 represents the between person (σ_o^2) variation. Both variance components were significantly different from zero (σ_E^2: 0.317 and σ_o^2: 0.546 with p < 0.01). To look at the amount of variation on each level, we looked at the intraclass correlation coefficient (ICC) from the unconditional model.

Intraclass correlation coefficient (ICC) = ρ = $\sigma_o^2 / (\sigma_o^2 + \sigma_E^2)$ = 0.63

Of the variation in income (ln_inc), 63% is due to variation between individuals and 37% to variation within individuals over time. In the observed quarters from 1998 to 2006, we can say that 37% of the variation in income can be ascribed to variation of income within the trajectory of each individual, or how each person's income changes over time, while 63% of the variation of income is due to variation between individuals, or how these changes differ across people.

The unconditional growth model

The unconditional growth model only has time as a predictor, with no other covariates taken into account (results not shown). This model fits the data better than the previous model, as both variance components declined in value. Because people can have different initial incomes as well as different income growths, we included the selection effect and the four time effects, both fixed and at random in the model (model 2). This did not considerably alter the estimates, but fit the data significantly better (lower Chi²/deviance, AIC and BIC). The variance components of the intercept, as well as the selection

effect and the time variables, differed significantly from zero. The level two residual variances σ_o^2 and σ_1^2 now summarise between-person variability in initial status and rates of change. A significant intercept variance implies that the individuals have different initial status/values (σ_E^2: 0.177 with p < 0.001).

Different variations of time indicate a significant variation between individuals in general income growth (TIME1), income drop (TIME2), income return (TIME3) or income recovery (TIME4). The overall intercept remains more or less the same (€8,716 or €2,033/month). The first TIME variable shows the overall income growth in our time window. In general, individual incomes rise by 0.005 per quarter. This is equivalent to €10 ($e^{(8.716+0.005)}$)/quarter or €41/year. TIME2 models the income drop at the start of the break. This effect is negative and shows a significant instant drop in income of €1,353 ($e^{(8.716+-0.407)}$/4)/month. The effect of returning to the previous job (TIME3) shows a non-significant positive effect of

0.004 or €2,041 ($e^{(8.716+0.004)}$). The income effect of returning after a break is a little smaller than the income drop at the start of the break. After the return, we see no significant difference in income growth between those who took a break and the control group. The effect of TIME4 is positive and shows a steady recovery of income after the initial return effect.

The selection shows the income level of those individuals who will take a break in the observation window, compared to the control group. The effect is negative and significant. This implies that those who decide to take a break are already earning less than the individuals in the control group. The difference amounts to €443/month ($e^{(8.716-0.246)}$/4).

Using the models below, we will try to explain the income differences before, during and after the break by including covariates on the second level and bringing them into interaction with our time variables.

Table 4. Multilevel models for wage differentials after a career break (n=13802)

		Model 1	Model 2	Model 3	Model 4	Model 5
FIXED EFFECTS						
Initial status	Intercept	8.356 ***	8.716 ***	8.521 ***	8.528 ***	8.773 ***
	Men			0.627 ***	0.604 ***	0.513 ***
	Age					-0.001 ***
	Age²					0.000 ***
	Younger than 50 (dummy)					-0.269 ***
	Has children					-0.036 ns
1. General income growth	Intercept (TIME 1)		0.005 ***	0.005 ***	0.005 ***	0.006 ***
	Time 1 * Men				- 0.004 ***	-0.004 ***
2. Break	Intercept (TIME 2)		-0.407 ***	-0.399 ***	-0.408 ***	-0.4144 ***
	Time 2 * Men				0.029 ns	-0.032 ns
3. Return	Intercept (TIME 3)		0.004 ns	0.010 ns	0.059 ***	0.060 ***
	Time 3 * Men				-0.187 ***	-0.188 ***
4. Recovery	Intercept (TIME 4)		0.004 ***	0.004 ***	0.008 ***	0.008 ***
	Time 4 * Men				-0.018 ***	-0.018 ***
5. Selection	Break (ref. = no break)		-0.246 ***	-0.159 ***	-0.165 ***	-0.127 ***
	Break * Men			-0.218 ***	-0.200 ***	-0.219 ***
	Break * Younger than 50					-0.068 ns
	Break * Has children					-0.014 ns

VARIANCE COMPONENTS

Level 1 Within-person	In general income level	0.317 ***	0.177 ***	0.177 ***	0.177 ***	0.177 ***
Level 2 Between-persons	In general income level	0.546 ***	0.466 ***	0.417 ***	0.417 ***	0.396 ***
	In general growth (time 1)		0.000 ***	0.000 ***	0.000 ***	0.000 ***
	In break effect (time 2)		0.326 ***	0.317 ***	0.317 ***	0.320 ***
	In return effect (time 3)		0.387 ***	0.385 ***	0.378 ***	0.374 ***
	In post-break growth (time 4)		0.004 ***	0.004 ***	0.004 ***	0.004 ***
Fit						
	-2LL	785699.6	603688.8	602614.7	602435.9	600277.8
	Chi²-diff test		***	***	***	***
	AIC	785703.6	603700.8	602626.7	602447.9	600289.8
	BIC	785718.6	603746.8	602671.9	602493.1	600335.0

Growth models with covariates

In models 3 and 4, we included gender as a predictor of income. Since career breaks are gendered to a high degree, we wanted to measure the impact of career breaks on incomes and income growth for both men and women. In model 3, we only included gender to explain the general income differences and the selection effect. The positive main effect shows that, on average, men in our sample earn €1,459/month more than women ($e^{(8.521+0.627)}$). The negative selection effect from model 2 decreases slightly, but model 3 shows that men who take a career break earn less at the start of our observation window than women. In model 3, the estimated initial income in 1998 for people without a career break is €1,673/month ($e^{8,521}/3$). The estimated difference in income in 1998 between men and women with a career break is €280/month ($e^{8,716-0.159}/3 - e^{8,521-0.159-0.218}/3$). Model 3 is significantly better than model 2 because the deviance is significantly lower in the latter (a deviance of 1,074 for a difference of 2 degrees of freedom in a Chi-square distribution is significant at a 99.9% probability level). This was confirmed by the lower AIC and BIC for model 3 compared to model 2. This suggests that adding gender significantly improves the model.

Model 4 further extended the gender analysis by including interaction effects with the time variables. As such, we can explore the income changes experienced by men and women during and after a career break.

The main effects of the time variables and the selection effect do not change in terms of significance or direction. Only the main effect of gender decreases from 0.627 to 0.604, still showing a higher general level of income for men. The interaction terms of gender with time are significant for TIME1, TIME3 and TIME4. The effect of TIME1 shows that the general income growth for men is smaller than that of women, which came as a surprise. The drop in income for those men who took a career break is non-significant. This implies that men drop about an equal amount of income as women. The recovery effect is negative for men, indicating that their income increases less sharply after returning to work. Again, this is surprising, but as we will show in model 5, this is related to the omission of some important control variables. Finally, we can see that the interaction effect of MEN*TIME4 is significant. Just like the lower return rate, the recovery rate of men is also lower after taking a career break.

Because the Belgian labour market shows a high take up of career breaks by individuals at the end of their career, we decided to include a dummy variable that captured the fact that the person concerned is aged 50 or older. Together with some general control variables such as age, age² and the presence of children, model 5 estimated the effect of this dummy on the income trajectory. Because of the complexity of the model, we only included the age-dummy and the presence of children as a main effect and as an interaction effect with our selection effect. The main effects were all in the expected direction. Individuals younger than 50 have lower incomes (-0.269) and we see no effect of the presence of children on the general income level. There is also no selection effect of parents taking a career break on the income level (-0.014) nor from the age dummy.

The important influence of gender and the symbolic age of 50 years have led us to examine their interrelation more closely. In table 5, we stratify our analyses in two directions. Firstly, we undertook separate gender analyses in order to see the gender-specific effect of the age boundary between men and between women. Secondly, we did the opposite by focusing on gender within groups younger and older than 50.

Growth models for gender and age separately

In model 6, we analysed the income differentials among the different types of career breaks for men and women separately. The overall effects remained more or less the same. The trajectory of incomes over time follows the same effect directions, with only minor differences in effect sizes between men and women. One clear difference was found in the selection effect. Men who interrupt their career have a lower income compared to men who continue working. The effect for women is also negative but to a far lesser degree. A second gender difference concerns the presence of children. For men, there is an overall positive effect on their income once a father, while women with children earn less than women without children.

Table 5. Multilevel models for wage differentials after a career break (n=13802)

		Model 6		Model 7	
		Men	Women	Minus 50 yrs	50 and older
FIXED EFFECTS					
Initial status	Intercept	9.206 ***	8.819 ***	8.513 ***	8.771 ***
	Men			0.554 ***	0.455 ***
	Younger than 50 (dummy)	-0.222 ***	-0.281 ***		
	Age	-0.001 **	0.000 ns	0.001 ***	0.001 ns
	Age²	0.000 ns	0.000 ***	0.000 ***	-0.000 ns
	Has children	0.096 **	- 0.093 ***	-0.054 *	-0.053 ns
1. General income growth	Intercept (TIME 1)	0.003 ***	-0.003 ***	0.008 ***	-0.003 ***
	Men			-0.001 ns	-0.000 ns
	Younger than 50 (dummy)	0.010 ***	-0.011 ***		
2. Break	Intercept (TIME 2)	-0.351 ***	-0.427 ***	-0.418 ***	-0.424 ***
	Men			-0.009 ns	-0.072**
	Younger than 50 (dummy)	-0.081 ns	-0.010 ns		
3. Return	Intercept (TIME 3)	-0.206 ***	-0.114 ***	0.121 ***	-0.113 ***
	Men			-0.136 ***	-0.093 **
	Younger than 50 (dummy)	0.195 ***	0.236 ***		
4. Recovery	Intercept (TIME 4)	-0.014 ns	0.005 *	0.0125 ***	0.005 **
	Men			-0.017 ***	-0.008 ***
	Younger than 50 (dummy)	0.009 **	0.0177 ***		
5. Selection	Break (ref. = no break)	-0.314 ***	-0.126 **	-0194 ***	-0.099 *
	Break * Men			-0.211 ***	-0.245 ***
	Break * Younger	-0.087 ns	-0.079 ns		

	than 50				
	Break * Has children	-0.512 ns	0.001 ns	0.024 ns	-0.007 ns

VARIANCE COMPONENTS

Level 1 Within-person	In general income level	0.116 ***	0.200 ***	0.208 ***	0.002 ***
Level 2 Between-persons	In general income level	0.299 ***	0.431 ***	0.437 ***	0.285 ***
	In general growth (time 1)	0.001 ***	0.001 ***	0.001 ***	0.001 ***
	In break effect (time 2)	0.213 ***	0.362 ***	0.364 ***	0.204 ***
	In return effect (time 3)	0.326 ***	0.372 ***	0.357 ***	0.606 ***
	In post-break growth (time 4)	0.003 ***	0.004 ***	0.004 ***	0.372 ***

Fit

-2LL		118886.3	469435.7	479050.2	98881.6
AIC		118898.3	469447.7	479062.2	98893.6
BIC		118936.1	469490.9	479105.3	98931.4

Source: sample of administrative data (KSZ) in Flanders
*** p-level <0.001 **p-level <0.01 *p-level <0.05 ns=not significant

Splitting our analyses according to the age-50 dummy (Model 7) did not reveal any differences in the overall effects and the main growth trajectory. Employees above and below 50 have a positive growth curve with a decline in income for those taking a career break. This effect is only significant for those aged 50 and older. A major difference, however, is apparent in the return effect (TIME3). While individuals below 50 follow the general path of increasing their income after they return to work (0.121), older employees do not experience an increase but a further decrease their income when returning. We assume that these older workers reduce their working hours after their break and therefore show no gain after returning to the workplace. A second result in which the age divide clearly plays a role can be found in the selection effect. The overall selection effect for employees below 50 is significant and negative, indicating that those with a lower income take career breaks more often. Above 50, this selection effect is only significant to the 0.05 level.

Discussion and conclusion

When looking at income trajectories before and after a career break, we observed a clear effect of the break, resulting in a relatively lower income after the break compared to the situation before (H1). As human capital theory predicted, we saw a significant recovery effect in the years after the break (H2). There is a catch-up effect in relation to work experience and slowly the gap closes between the incomes of those who took a career break and their colleagues who did not.

The income differentials are much higher between men with and without a break than between women with and without a break. This is partly due to the initial income gap between men and women. Women lose less when taking a career break because their income is lower to begin with. However, when we controlled for the presence of children, we found that while women with a break showed significantly larger income growth after the break than women without a break, there were no significant differences in income growth between men, whether they took a break or not. In terms of income, men seem to be 'penalised' more for taking a career break than are women (H3).

This could be due to the fact that men are stigmatised more for taking a break. As it is not common for men to do so, this may send a signal that they are less committed to their job, which then leads to lower wages (signalling theory).

According to statistical discrimination theory, the lower wage loss found for women could also be due to the fact that women are penalised by lower wages regardless of a career break because it is already anticipated that they will take a break. By controlling for a selection effect, this turned out to be the case, but even more so for men taking a break than for women. Men on lower incomes tend to take a career break more often than women on lower incomes. We observed a reversed 'double standard' working against men (Foschi, 2000). Evaluations and expectations differ for men and for women and taking a career break might be subject to different evaluative processes within the labour market. Although we do not possess data showing the influence of these implicit standards, research has already shown the presence of gendered evaluative processes (Beblo & Wolf, 2002; Foschi, 1996).

Women taking a career break had a significantly lower income than women without a break, after controlling for age and child-related covariates. The fact that the child-related variables did not have a significant effect on the income differentials of men was remarkable, which might suggest that men take career breaks for personal or work-related reasons. For women, we saw a reverse selection effect for mothers. Mothers tend to have a higher income before they take a break. We do not have data on educational levels in the Datawarehouse but we assume that this effect could be due to an educational gradient in the taking of career breaks among women.

Because the analysis concerns Belgian data, special attention was paid to the age limit of 50 years. Beyond this age, employees can take career breaks in a system that allows them to ease out of the labour force as they approach retirement. The use of this system is partly responsible for the low labour market participation of those over 50 in Belgium.

The interaction model indeed showed that the income trajectories, especially for men, are different for younger employees. After 50, the selection effect of taking a career break disappears, indicating that all

men over 50 use the system in order to reduce their working hours at the end of their career. Below 50, less well paid men take career breaks. The gendered consequences are also more men-friendly after 50. While we still found negative consequences for men (and women) after 50, they were much more modest. Men who take a career break before 50 are particularly liable to experiencing negative consequences in terms of income trajectory, both during and after a career break (H4). Taking the specific Belgian context into account, it can be concluded that the results concerning employees over 50 were influenced more by the Belgian end-of-career measures when compared to other countries. These measures are unique to the country, such that we doubt the generalisability of the longitudinal income trajectories for people over 50. For the major part of the sample, the results were not different from those studies of other career break systems. This implies that the gendered outcome of taking a career break reveals the negative outcomes in terms of income for men compared to women.

Despite the huge dataset and the reliability of the income data, this study has several limitations that need to be taken into account. A first limitation concerns the income data used in this study. The registers provided us with quarterly income data, and while this income is a reflection of the hourly wage of the individuals in the sample, the aggregated data does not allow us to convert the income back into hourly data.

As a result, we could not account for actual changes in work hours in our models. The changes in income are a proxy for changes in work hours, as Belgian legislation requires employers to allow employees to return with their original contracts. Income trajectories therefore reflect changes in hours worked or changes in the workplace (which is also limited in the rigid Belgian work force) (Kovalenko & Mortelmans, 2014), although the models cannot account for the actual hourly wages behind the income trajectories.

A second limitation of the study was the absence of data on educational level. The registers do not include this characteristic of the individuals in question. Data on other more subjective work-related issues, such as career aspirations, were also not available. As a consequence, these kinds of analyses do have great statistical power but remain rather descriptive. Finally, a third limitation is related to this and concerns the possible presence of selection effects. It might be the case that specific groups of employees select themselves in the career break scheme. They could be individuals with a lower labour market commitment and less ambition for promotion. It might be the case that these personal characteristics are only partly represented by these workers' wages. We did include a selection effect in the models as a way to control for lower wages before the break, but this might not be the complete story. The income differences found were, nevertheless, significantly large, especially for men. Future research on the dynamics of these processes is required.

References

Albrecht, J. W., Edin, P. A., Sundstrom, M., & Vroman, S. B. (1999). Career interruptions and subsequent earnings: A reexamination using Swedish data. *Journal of Human Resources, 34*(2), 294-311. https://doi.org/10.2307/146347

Arun, S. V., Arun, T. G., & Borooah, V. K. (2004). The effect of career breaks on the working lives of women. *Feminist Economics, 10*(1), 65-84. https://doi.org/10.1080/1354570042000198236

Beblo, M., & Wolf, E. (2002). *Wage penalties for career interruptions: an empirical analysis for West Germany*. Retrieved from

Becker, G. S. (1964). *Human Capital: A theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press. https://doi.org/10.1086/298075

Becker, G. S. (1985). Human Capital, Effort, and the Sexual Division of Labor. *Journal of Labor Economics, 3*(1), S33-S58. https://doi.org/10.1086/298075

Bünning, M. (2015). What Happens after the 'Daddy Months'? Fathers' Involvement in Paid Work, Childcare, and Housework after Taking Parental Leave in Germany. *European Sociological Review, 31*(6), 738-748. https://doi.org/10.1093/esr/jcv072

Corcoran, M., & Duncan, G. J. (1979). Work history, labor force attachment, and earnings differences between the races and sexes. *Journal of Human Resources, 14*(1), 3-20. https://doi.org/10.2307/145535

Corcoran, M., Duncan, G. J., & Ponza, M. (1983). A longitudinal analysis of white women's wages. *Journal of Human Resources, 18*(4), 497-520. https://doi.org/10.2307/145441

Couch, K. A., & Placzek, D. W. (2010). Earnings Losses of Displaced Workers Revisited. *American Economic Review, 100*(1), 572-589. https://doi.org/10.1257/aer.100.1.572

De Preter, H., Van Looy, D., & Mortelmans, D. (2014). Retirement Timing of Dual-Earner Couples in 11 European Countries? A Comparison of Cox and Shared Frailty Models. *Journal of family and economic issues, 36*(3), 396-407. https://doi.org/10.1007/s10834-014-9403-6

Desmet, B., Glorieux, I., & Vandeweyer, J. (2007). *Wie zijn de loopbaanonderbrekers? Socio-demografische kenmerken, motivaties en arbeidshouding van loopbaanonderbrekers*. Brussel: VUB.

Dex, S. (1999). Careers and motherhood: policies for compatibility. *Cambridge Journal of Economics, 23*(5), 641-659. https://doi.org/10.1093/cje/23.5.641

Ekberg, J., Eriksson, R., & Friebel, G. (2013). Parental leave — A policy evaluation of the Swedish "Daddy-Month" reform. *Journal of Public Economics, 97*, 131-143. https://doi.org/10.1016/j.jpubeco.2012.09.001

Eurofound. (2004). *Living to work - working to live: Tomorrow's work-life balance in Europe* (Vol. 8). Brussels: European Foundation for the Improvement of Living and Working Conditions,.

Foschi, M. (1996). Double standards in the evaluation of men and women. *Social Psychology Quarterly, 59*(3), 237-254. https://doi.org/10.2307/2787021

Foschi, M. (2000). Double standards for competence: Theory and research. *Annual Review of Sociology, 26*, 21-42. https://doi.org/10.1146/annurev.soc.26.1.21

Gangl, M., & Ziefle, A. (2009). Motherhood, labor force behavior, and women's careers: an empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States. *Demography, 46*(2), 341-369. https://doi.org/10.1353/dem.0.0056

Geisler, E., & Kreyenfeld, M. (2011). Against all odds: Fathers' use of parental leave in Germany. *Journal of European Social Policy, 21*(1), 88-99. https://doi.org/10.1177/0958928710385732

Hyman, J. (2005). Getting on or getting by?: Employee flexibility and coping strategies for home and work. *Work, Employment & Society, 19*(4), 705-725. https://doi.org/10.1177/0950017005058055

Jansen, M., Mortelmans, D., & Snoeckx, L. (2009). Repartnering and (Re)employment: Strategies to Cope With the Economic Consequences of Partnership Dissolution. *Journal of Marriage and Family, 71*(5), 1271-1293. https://doi.org/10.1111/j.1741-3737.2009.00668.x

Kletzer, L. G., & Fairlie, R. W. (2003). The long-term costs of job displacement for young adult workers. *Industrial & labor relations review, 56*(4), 682-698. https://doi.org/10.1177/001979390305600408

Kovalenko, M., & Mortelmans, D. (2014). Does career type matter? Outcomes in traditional and transitional career patterns. *Journal of vocational behavior, 85*(2), 238-249. https://doi.org/10.1016/j.jvb.2014.07.003

Manning, A., & Petrongolo, B. (2008). The Part-Time Pay Penalty for Women in Britain*. *The Economic Journal, 118*(526), F28-F51. https://doi.org/10.1111/j.1468-0297.2007.02115.x

Mincer, J., & Ofek, H. (1982). Interrupted Work Careers - Depreciation and Restoration of Human-Capital. *Journal of Human Resources, 17*(1), 3-24. https://doi.org/10.2307/145520

Nepomnyaschy, L., & Waldfogel, J. (2007). Paternity Leave and Fathers' Involvement with Their Young Children. *Community, Work & Family, 10*(4), 427-453. https://doi.org/10.1080/13668800701575077

Pylkkanen, E., & Smith, N. (2004). The impact of family-friendly policies in Denmark and Sweden on mothers' career interruptions due to childbirth. *IZA Discussion Paper, 1050*, 1-36.

Rege, M., & Solli, I. F. (2013). The impact of paternity leave on fathers' future earnings. *Demography, 50*(6), 2255-2277. https://doi.org/10.1007/s13524-013-0233-1

Schönberg, U., & Ludsteck, J. (2007). Maternity leave legislation, female labor supply, and the family wage gap. *IZA Discussion Papers, 2699*, 1-64.

Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis. Modeling change and event occurence*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195152968.001.0001

Spivey, C. (2005). Time off at what price ? The effects of career interruptions on earnings. *Industrial & labor relations review, 59*(1), 119-140. https://doi.org/10.1177/001979390505900107

Theunissen, G., Verbruggen, M., Forrier, A., & Sels, L. (2009). Career Sidestep, Wage Setback? The Impact of Different Types of Employment Interruptions on Wages. *Gender, Work & Organization*. https://doi.org/10.1111/j.1468-0432.2009.00471.x

Trappe, H. (2013). Fathers Who Claim Parental Leave Benefits: Only a Matter of Economic Considerations? *Zeitschrift für Soziologie, 42*(1), 28-81. https://doi.org/10.1515/zfsoz-2013-0104

Waldfogel, J. (1998a). The family gap for young women in the United States and Britain: Can maternity leave make a difference? *Journal of Labor Economics, 16*(3), 505-545. https://doi.org/10.1086/209897

Waldfogel, J. (1998b). Understanding the "Family Gap" in Pay for Women with Children. *The Journal of Economic Perspectives, 12*(1), 137-156. https://doi.org/10.1257/jep.12.1.137

Longitudinal methods for life course research: A comparison of sequence analysis, latent class growth models, and multi-state event history models for studying partnership transitions

Júlia Mikolai   University of St Andrews and University of Southampton, UK
Julia.Mikolai@st-andrews.ac.uk
Mark Lyons-Amos   London School of Hygiene and Tropical Medicine and University of Southampton, UK

Abstract

This paper qualitatively compares and contrasts three methods that are useful for life course researchers; the more widely used sequence analysis, and the promising but less often applied latent class growth models, and multi-state event history models. The strengths and weaknesses of each method are highlighted by applying them to the same empirical problem. Using data from the Norwegian Generations and Gender Survey, changes in the partnership status of women born between 1955 and 1964 are modelled, with education as the primary covariate of interest. We show that latent class growth models and multi-state event history models are a useful addition to life course researchers' methodological toolkit and that these methods can address certain research questions better than the more commonly applied sequence analysis or simple event history analysis.

Introduction

In the last half century, family life courses have changed considerably. For example, the transition to parenthood has been delayed, non-marital cohabitation and non-marital childbearing have become more common, as have union dissolution and re-partnering. Additionally, the timing and sequencing of family life events has changed resulting in more complex and less predictable family life courses. These changes have generated an increased interest in the applicability of different methods for modelling life courses with their complexities. Although a number of methods are available to study the family life course, discussion has mainly been limited to comparing simple event

history models and sequence analysis (Billari, 2001b, 2005; Billari & Piccarreta, 2001, 2005; Piccarreta & Billari, 2007).

Simple event history analysis is commonly used to examine single or competing events (Heuveline & Timberlake, 2004; Perelli-Harris & Gerber, 2011; Perelli-Harris, Sigle-Rushton, et al., 2010). Applications of event history analyses vary in focus and complexity. For example, recent studies (Baizán, Aassve, & Billari, 2003, 2004) applied simultaneous equations models to study the determinants of several concurrent life course transitions. Others used multilevel multiprocess models to account for correlated event histories (Steele, Kallis, Goldstein, & Joshi, 2005). These 'event based' approaches

primarily focus on the (causal) influence of certain covariates on particular events. Simultaneous models improve upon simple event history models by accommodating possible interdependencies between several events via modelling joint processes and unobserved heterogeneity. Even so, they limit attention to studying a specific segment of the life course.

Others have promoted the use of sequence analysis arguing that, unlike event history analysis, this 'holistic approach' examines the life course trajectory as a whole meaningful unit. This technique creates 'ideal-types' of trajectories that categorise and describe different life course patterns (Billari, 2001a, 2001b, 2005; Billari & Piccarreta, 2005; Piccarreta & Billari, 2007). It is then possible to assess how different covariates influence the probability of an individual to belong to one of these 'ideal-types'.

Despite the availability of other techniques only a few studies have investigated their applicability to life course research. For example, Barban and Billari (2012) have compared and tested the consistency of sequence analysis and latent class analysis. Additionally, Bonetti, Piccarreta, and Salford (2013) proposed an extension of multi-state models to studying the family life course. This paper aims to contribute to this line of research by qualitatively comparing the strengths and weaknesses of sequence analysis and two other promising techniques: latent class growth models, and multi-state event history models. These methods combine the properties of the event based and the holistic approaches by focusing on several consecutive events and thus are ideal to examine the family life course.

By applying these methods to a real life example, the differences and similarities as well as the strengths and weaknesses of these approaches are emphasised. Our application focuses on the role of education on changes in partnership status (i.e. being never partnered, transition to first cohabitation and first marriage, the dissolution of a first cohabitation or a first marriage, and forming a new partnership after union dissolution) of Norwegian women born between 1955 and 1964. We ask the following questions, pertinent to life course research: how can sequence analysis, latent class growth models and multi-state event history models be used for studying the influence of education on partnership transitions over the early

family life course? What types of research questions can be answered using these methods? And are these methods applicable to the same problems to the same extent or is one of them better than the other and if so in which situation?

Data

We illustrate similarities and differences between sequence analysis, latent class growth models, and multi-state event history models using data from the first wave of the Norwegian Generations and Gender Survey [i] (GGS) from 2007/2008. We examine the influence of educational attainment on changes in partnership status of women born between 1955 and 1964 (N = 1290). The dataset includes extensive retrospective information on the start and end date (year and month) of up to five cohabiting and marital unions as well as union dissolutions. Cohabitation is defined as a co-residential relationship which lasted for at least three months. For this application Norwegian data are used because in Norway variation in partnership experiences is one of the largest compared to other European countries (Elzinga & Liefbroer, 2007). This enables us to demonstrate the strengths and weaknesses of the examined methods using a rich dataset with more complex partnership experiences.

Although the Norwegian GGS provides cross-sectional weights, not all applied methods are able to incorporate these. Therefore, the analyses presented in this paper do not incorporate weights. This implies that the results might not be representative of the overall distribution of the examined partnership formation behaviours in Norway. However, this is not a major limitation because the aim is not to provide population estimates but to explore how the different methods can be applied to the same problem.

Methods and Modelling Strategy

Sequence Analysis

Sequence analysis (SA) represents each individual life course by a sequence (i.e. a character string, which indicates the order and duration of states occupied by an individual in each month). For example, the sequence SSSCCMMMM means that the respondent was never partnered [ii] (S) for three months followed by two months of cohabitation (C) and four months of marriage (M). Due to the large possible number of combinations of states, usually

very few individuals experience the exact same sequence. To reduce the number of sequences, Optimal Matching Analysis (OMA) is used.

OMA is a technique that measures the dissimilarity between sequences by identifying how similar pairs of sequences are. Similarity is defined in terms of the number, order, and duration of states within sequences. The algorithm calculates the similarity/dissimilarity between two sequences by taking into account three possible operations: replacement (one state is replaced by another one), insertion (an additional state is added to the sequence), and deletion (a state is deleted from the sequence). The fewer operations are needed to turn one sequence into the other, the more similar two sequences are and vice versa. Furthermore, to each operation, a certain cost can be attached. Therefore, identifying the relative cost of all operations is critical to determine the (dis)similarity between sequences. These require *a priori* definition by the researcher with little objective measure of the correct specification, and the results can be highly sensitive to these specifications (Brzinsky-Fay & Kohler, 2010). Then, the distance between two sequences is defined by the minimum cost of the operations that are necessary to transform one sequence into the other (Abbott & Tsay, 2000). The distances are recorded in a dissimilarity matrix.

Then, in order to find existing patterns in the data, hierarchical cluster analysis is performed on the dissimilarity matrix. The aim of the cluster analysis is to minimise the within-cluster and maximise the between-cluster distance. The researcher needs to specify the number of clusters to be extracted from the data either *a priori* or by using fit statistics. Once the clusters are formed, they can be described with respect to the variables used to create the clusters (in this example partnership experiences). The clusters can be used both as independent and dependent variables in further analyses. Additionally, sequences can also be compared based on the number of episode changes within once sequence, the length of the sequences, or the number of different events in a sequence (Brzinsky-Fay & Kohler, 2010).

In our application, clusters are created based on women's monthly partnership trajectories between age 15 and 40. Women can be in the following partnership states in a given month: single, cohabiting, married, and separated. Individuals who experience a new partnership following separation can be in the 'marriage' or 'cohabitation' state. After performing OMA with equal costs assigned to indel operations (i.e. insertion and deletion; in this instance 1), individuals are allocated to clusters based on Ward's distance. Since the results can be sensitive to the chosen indel and substitution costs, we performed sensitivity analyses varying the indel costs to be 0.5, 1.0 and 1.5, and using both a constant substitution matrix as well as a matrix based on the frequency of transitions. The findings indicated that the number and composition of the obtained clusters remain consistent across the different specifications. We assess the number of clusters based on two measures of average cluster linkage; the Calinski–Harabasz pseudo-F index (Calinski & Harabasz, 1974) and the Duda–Hart index (Duda & Hart, 1973). These statistics help to determine the optimal number of clusters by comparing the ratio of the within-cluster distances to the between-cluster distances. Additionally, the Duda-Hart index also produces a pseudo T-statistic[iii]. Once the optimal number of clusters is established, cluster allocation is used as the response variable in a multinomial logistic regression. The models are estimated using the SQ-Ados for Stata 12 (Brzinsky-Fay, Kohler, & Luniak, 2006).

Latent Class Growth Models

Latent Class Growth Models (LCGMs) are an extension of conventional growth curve models with the assumption that individuals are drawn from different subpopulations (latent classes) that have different growth trajectories (Perelli-Harris & Lyons-Amos, 2015). Similarly to SA, these models have an individual centred perspective, thus they seek to identify relationships between individual response patterns and form groups based on these patterns (Jung & Wickrama, 2008). Individuals' relationship histories are recorded at each age and then grouped into latent classes. The response (in this application partnership state in each year) is defined as the random variable $y_{i,age}$ with the following categories: never partnered (0), persistent cohabitation (1), marriage (2), and separated (3). After separation, individuals are allowed to re-enter cohabitation or marriage. Respondents move between these partnership states between age 15 and 40[iv].

The specification of the growth curves and the robustness checks are similar to those in Perelli-Harris and Lyons-Amos (2015). Classes are formed

based on yearly partnership histories[v]. Individuals' partnership histories form trajectories which are combined to form the latent classes. Latent classes describe different partnership patterns across individuals' life courses. Each woman has a probability of belonging to each latent class. The closer an individuals' partnership history is to the class trajectories, the more likely she is to belong to a particular latent class. The probability of being in partnership state s at a given age is defined as $\pi_{i,age}^s = P(y_{i,age} = s)$ where i stands for individuals. The probability of each partnership (compared to marriage which is the reference category) across the life course is modelled as a growth equation (see Equation 1). A separate growth equation is specified for each class C_j ($j = 1 \dots 5$) which are defined by a class-specific intercept (α_j^s) and class-specific slope parameters ($\beta_{1\dots4,j}^s$).

$$\ln\left(\frac{(\pi_{i,age}^s | C_j = j)}{(\pi_{i,age}^{s=2} | C_j = j)}\right) = \alpha_j^s + \beta_{1,j}^s age_i +$$
$$\beta_{2,j}^s age_i^2 + \beta_{3,j}^s educ + \beta_{4,j}^s educ * age_i$$
$$(1)$$

In order to examine the relationship between the latent classes and education LCGMs offer two possibilities. First, education can be used to predict the probability of belonging to a certain latent class (Wang, Hendricks Brown, & Bandeen-Roche, 2005). This approach is comparable to SA. Where LCGMs have an advantage over SA is that they can also allow for education to alter the shape of the growth curves. The parameter $\beta_{3,j}^s$ expresses how the inclusion of educational level alters the intercept while $\beta_{4,j}^s$ indicates the influence of education on the slope of the growth curves. An additional advantage of LCGMs compared to SA is that a variety of fit statistics are available for deciding the optimal number of classes and this choice can be

validated via simulation since the estimates are model based. However, the different criteria and test statistics (such as AIC, BIC or Lo-Mendell-Rubin Likelihood Ratio Test) can lead to different and sometimes contradictory conclusions (Nylund, Asparouhav, & Muthen, 2007).

In our application, we explore a set of two, three, four, and five class models and perform the Lo-Mendell-Rubin-Likelihood Ratio Test (LMR-LRT) for all classes. This test examines the improvement in model fit for a J class model compared to a J-1 class model. In case of a two class model, it is equivalent to testing whether the LCGM performs better than a simple latent growth curve model, which assumes that one growth curve is enough to describe women's partnership behaviours. The models are estimated in Mplus 6.2 for Linux. Note that we do not explore models with more than five classes[vi]. Due to the specification of partnership state as a nominal variable, the implementation of this model is not part of the main Mplus language. As a result, model estimation is computationally intense both due to the difficulty of the calculations required and the volume of data to be read.

Multi-state Event History Models

Multi-state event history models differ from SA and LCGMs in that they do not aim to classify or group individuals[vii]. It is a variable-centred approach where the main purpose is to establish statistical relationships between the independent variable(s) and several transitions. Multi-state event history models are an extension of simple event history models; rather than examining one transition, this approach allows individuals to move among different states over time. These movements are assumed to be stochastic and are modelled by means of transition probabilities. Thus, multi-state event history models allow for examining covariate effects on several transitions within the same model.
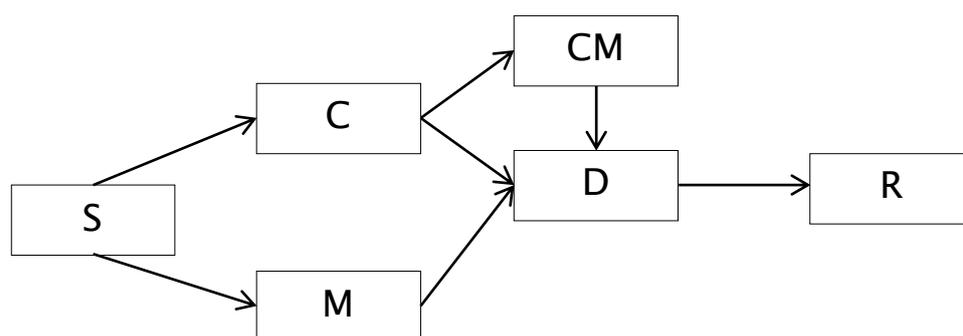
Note: S – never partnered, C – cohabitation, M – direct marriage, CM – marriage preceded by cohabitation with the same partner, D – union dissolution, R – re-partnering.

Figure 1. Multi-state event history model

Another distinct advantage of this method is the possibility to include time-varying covariates and thereby examine how the influence of a variable of interest changes over the family life course. For example, it is possible to examine the influence of educational attainment, which may change over the life course, on several family life transitions. This cannot be done using simple event history models, SA, or LCGMs. Multi-state models assume the Markov property; that is that the present behaviour of an individual is enough to predict their future behaviour (Andersen & Keiding, 2002; Hougaard, 1999). For example, it would assume that the transition probability from marriage to union dissolution is the same for all individuals regardless of whether they have cohabited before marriage. As life course theory emphasises that earlier transitions play an important role in later transitions, this assumption is not realistic. In order to be able to examine the partnership transitions in a dynamic way, the model can be extended. We do so by defining the state 'CM' to differentiate between direct marriage and marriage that was preceded by cohabitation. One disadvantage of multi-state event history models is that as the number of states increases and as individuals move along the life course, one might end up with small cell sizes and thus, with unreliable estimates of the transition hazards.

Figure 1 shows the multi-state event history model, where the following states are defined: never partnered (S), cohabitation (C), direct marriage (M), marriage that was preceded by

cohabitation with the same partner (CM), union dissolution (D) and re-partnering (R). We do not distinguish between cohabitation and marriage as a form of re-partnering due to relatively small cell sizes and to keep the models comparable. The multi-state event history model is estimated using a stratified continuous-time Cox model with a non-parametric baseline hazard where each transition is represented by a different stratum (de Wreede, Fiocco, & Putter, 2011; Putter, Fiocco, & Geskus, 2007; Putter, van der Hage, de Bock, Elgalta, & van de Velde, 2006). Covariates are incorporated as transition-specific covariates to allow for their effect to differ across transitions. The transition hazard of individual *k* is given by:

$$\lambda_{ij}(t|\mathbf{Z}(k)) = \lambda_{ij,0}(t) \exp\left(\boldsymbol{\beta_{ij}}^T \mathbf{Z}(k)_{ij}\right)$$

(2)

where *ij* indicates a transition from state *i* to state *j*, $\lambda_{ij,0}(t)$ is the baseline hazard, *Z(k)* is the vector of covariates at baseline for individual *k* and $\mathbf{Z}(k)_{ij}$ is the vector of transition-specific covariates for individual *k*.

In this application, we use monthly information on partnership experiences. To estimate the model, an augmented dataset is used with one row per transition that the individual is at risk of. Women are observed from age 15, when they are never partnered until age 40, the time of the survey, or the time when they experience re-partnering, whichever happens earlier. As educational

attainment is defined as a time-varying categorical variable, additional episode splitting is performed where an educational transition happens within an at-risk period. The models are estimated using the *mstate* package in R (de Wreede et al., 2011).

Variables

Level of education. In all three models, the highest level of education at the time of the survey is measured by a variable with the following categories: low (ISCED 0 to ISCED 2), medium (ISCED 3 and ISCED 4), and high education (ISCED 5 and ISECD 6). High education is the reference category in all three models. In the multi-state event history models, education is measured as a time-varying variable which is created using information on the year and month of reaching the highest level of education. We assume continuous education from age 15 and that secondary education takes four years while tertiary education takes three years on average. Missing information (7.9%) on the year and/or month of reaching the highest level of education was imputed using information on the median age of finishing education by educational level. In LCGMs and SA, education is time-constant and indicates the highest level of education at the time of the survey.

Results

Sequence Analysis

Table 1 presents the Calinski–Harabasz and the Duda–Hart indices for two to seven cluster models. On the Calinski–Harabasz and Duda–Hart indices, higher values indicate more distinct clustering, whereas for the related Duda–Hart Pseudo T-square measure, lower values are indicative of more distinct grouping.

Table 1. Calinski–Harabasz and Duda–Hart indices for k cluster specifications

Number of clusters (k)	Calinski–Harabasz Pseudo-F	Duda–Hart indices	
		Je(2)/Je(1)	Pseudo T-square
2	51.39	0.97	12.96
3	33.72	0.95	42.33
4	35.73	0.52	227.69
5	**81.37**	0.98	8.91
6	67.79	**0.99**	**1.22**
7	56.79	0.99	1.51

Note: Numbers in boldface indicate the best fit for the given index.

There is disagreement between these indices as to the optimal number of clusters. The Calinski–Harabasz index indicates a five cluster solution while the Duda-Hart indices indicate a six cluster solution to be optimal. We proceed with a six cluster model because the sixth cluster has a substantial, distinct meaning for our application.

The six clusters are shown in six rows of figure 2. The graphs in the left hand column show so called sequence index plots where each line represents the partnership sequences of an individual; these individual sequences are stacked on top of each other. The graphs in the right hand column are so called chronograms, which represent the distribution of states at each age providing a more readable summary than sequence index plots (Halpin, 2014).

The first cluster (figure 2, first row) is characterised by late partnership formation, where the first partnership is either direct marriage or cohabitation. Therefore, this cluster is titled '*late, varied partnerships*'. Women who belong to the second cluster form partnerships at a much younger age, than those in the first cluster (figure 2, second row). Most of these partnerships start as direct marriage, and there is a large degree of union instability at later ages. This cluster is, thus, named the *'early, direct, unstable marriage'* cluster. The third cluster (figure 2, third row) is largely characterised by stable cohabitation across the life course, although for some individuals cohabitation

is a second partnership following the dissolution of a marriage. This cluster is named the '*cohabitation'* cluster. Cluster four (figure 2, fourth row) consists of women whose first union is direct marriage starting from about the age of 25. There is little evidence of any other partnership behaviour, and for this reason cluster four is called '*later, direct, and stable marriage'* cluster. Cluster five (figure 2, fifth row) captures women who form cohabiting

unions, which then transition to marriage. We, therefore, call this cluster '*cohabitation followed by marriage'* cluster. Finally, the sixth cluster (figure 2, sixth row) comprises of women who married their partner directly at a (very) young age (before age 25). There is very limited evidence of any other partnership form (< 5%) in this cluster, hence we call this cluster the '*early, direct, and stable marriage'* cluster.
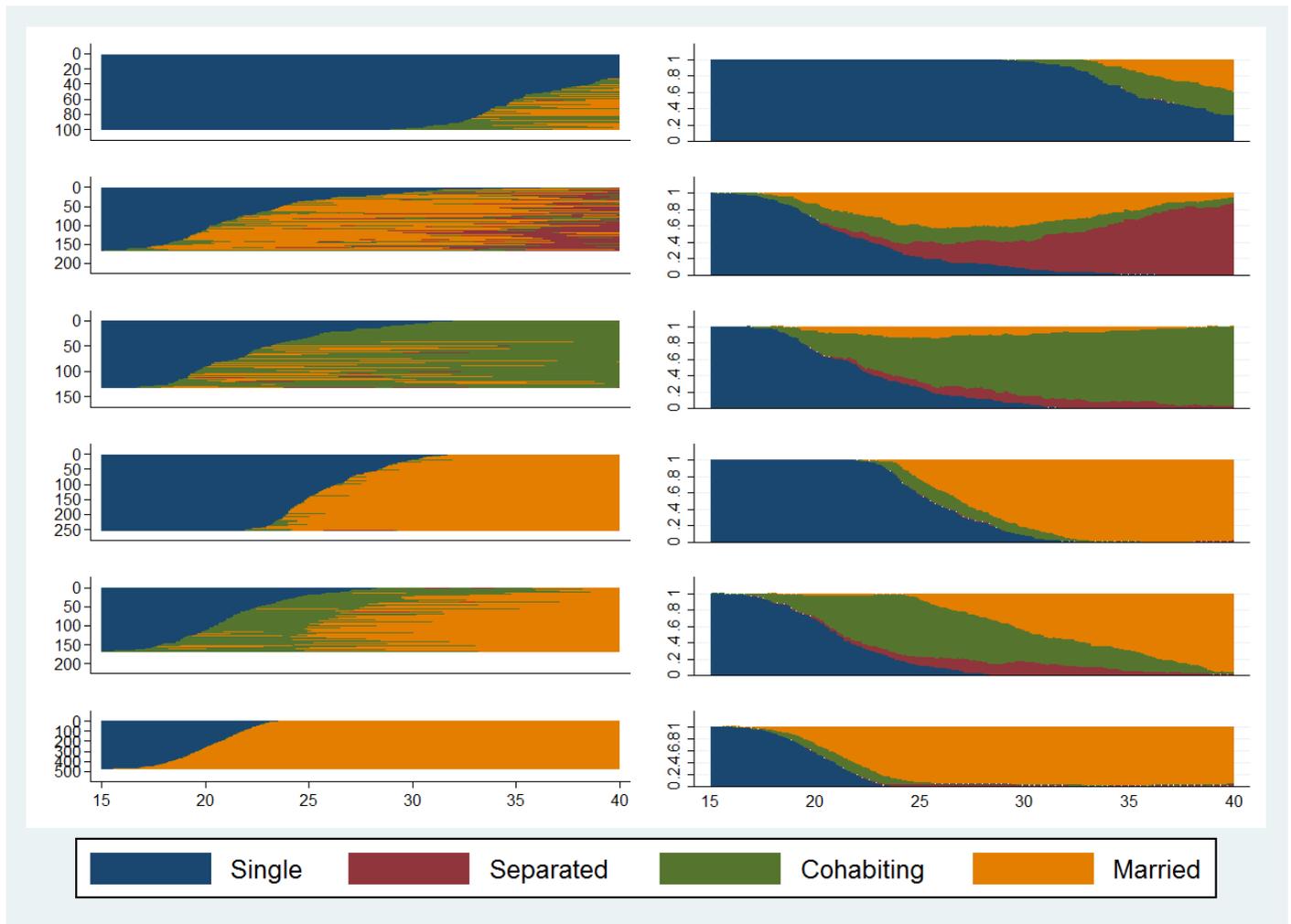
Note: Each row corresponds to a cluster.
Note: Cluster 1: Late, varied partnerships (7.8%); Cluster 2: Early, direct, unstable marriage (12.8%); Cluster 3: Cohabitation (10.2%); Cluster 4: Later, direct, and stable marriage (19.5%); Cluster 5: Cohabitation followed by marriage (13.0%); Cluster 6: Early, direct, and stable marriage (36.8%).

Figure 2. Results of sequence analysis in the form of sequence index plots (left hand column) and chronograms (right hand column)

After having identified these six clusters, we apply multinomial logistic regression to assess how educational attainment influences the likelihood of

women to belong to one of the six clusters (table 2). We compare the likelihood of belonging to each cluster to the likelihood of belonging to cluster six

(early, direct, and stable marriage) because this cluster is the largest. To further facilitate the interpretation of the regression coefficients, predicted probabilities are calculated (table 3). The results show that lower educated women are significantly less likely to belong to the 'late, direct, and stable marriage' cluster (cluster four) and to the 'cohabitation followed by marriage' cluster (cluster five) compared to belonging to the 'early, direct, and stable marriage' cluster than highly educated women.

Table 2. Results of the multinomial logistic regression, regression coefficients

	Membership of cluster 1 vs cluster 6	Membership of cluster 2 vs cluster 6	Membership of cluster 3 vs cluster 6	Membership of cluster 4 vs cluster 6	Membership of cluster 5 vs cluster 6
Education					
High (ref)					
Medium	0.088	-0.151	-0.107	-0.630***	-0.169
Low	-0.162	-0.343	-0.017	-1.220***	-0.405*
Intercept	-1.550***	-0.913***	-1.230***	-0.151	-0.880***

*Note: *p < .05 **p < .01 ***p < .001*

Table 3. Predicted probabilities of cluster membership by educational level

Cluster	Low education	Medium education	High education
1 Late, varied partnerships	0.08	0.21	0.07
2 Early, direct, unstable marriage	0.12	0.16	0.13
3 Cohabitation	0.13	0.17	0.09
4 Late, direct, stable marriage	0.11	0.10	0.27
5 Cohabitation followed by marriage	0.12	0.16	0.13
6 Early, direct, stable marriage	0.44	0.19	0.31

Latent Class Growth Models

Table 4 presents fit statistics for two-, three-, four-, and five class models. The LMR-LRT p-value indicates that the two-class model is an improvement over a one-class model, justifying the LCGM approach. All fit statistics indicate improving model fit with the addition of higher order classes.

From the examined models, the five-class model demonstrated the best model fit based on AIC, BIC and Sample Size BIC (SSBIC) statistics. Although the Lo-Mendell-Rubin Likelihood Ratio Test (LMR-LRT) indicates that a four-class model is adequate, we select a five-class model since this is the optimal number of classes according to all other fit statistics.

Table 4. Fit statistics for two-, three-, four-, and five-class models

Number of classes (J)	AIC	BIC	SSBIC	LMR-LRT (p-value)
2	138352.929	138731.851	138588.841	0.000
3	132500.352	133081.366	132862.085	0.016
4	129273.584	130056.690	129761.137	**0.021**
5	**126725.499**	**127710.697**	**127338.871**	0.174

Note: Numbers in boldface indicate the best fit based on the given statistic.

Figure 3 presents partnership profiles for the five extracted classes by educational attainment. The latent classes are depicted in the rows and educational groups are shown in the columns (e.g. the graph in the first row and first column shows class one for low educated women). Class one captures early and varied partnership forms, with an initial increase in the probability of both cohabitation and marriage for all educational levels. The probability of marriage peaks around age 24 for high and medium educated and at age 22 for low educated, and declines thereafter. The probability of cohabitation rises, plateauing at age 22 for high and medium educated and at age 19 for low educated, before increasing again from around age 31 onwards. These relationships, formed at relatively early ages, are unstable; the probability of separation is high across all partnership forms and educational levels. There is some variation in how the probability of separation changes over age by educational attainment. For women with high or medium education, the probability of separation increases and remains high until age 40. In contrast, for low educated women it reaches its maximum at age 35 and falls thereafter, corresponding to an increase in the probability of post-separation cohabitation.

Class two broadly represents a long-term cohabitation pattern. Most women at all educational levels form cohabiting relationships from their early 20s, with a peak in the probability of cohabitation around the age of 28 for high and medium educated women and at age 25 for low educated women. Thereafter, the probability of being in a cohabiting relationship decreases among women with high and medium education coinciding with an increasing probability of marriage from around age 31 (which reaches 0.45 for highly educated lyomen and 0.39 for women with medium education). In contrast, women with low education continue to exhibit a high probability of

cohabitation (nearly 0.7 at age 40). Consequently, the corresponding increase in the probability of marriage is limited, reaching only 0.2 by age 40. This result indicates that low educated women are less likely to formalise their unions. Additionally, the probability of separation is more than twice as high among low educated women as among their more educated counterparts.

Class three describes a generally early transition to marriage with some pre-marital cohabitation. Women with high and medium education have very similar partnership experiences: partnership formation begins with a small bump in the probability of cohabitation, followed by a transition to marriage with a 50% chance of being married around age 22. The probability of marriage is close to 1 among these women in their late 20s and it remains high with little evidence of separation. The patterns are slightly different for women with low education. Entry into partnership occurs earlier, with a decline in the probability of being single already from age 15. The probability of pre-marital cohabitation is higher among low educated women than among their more educated counterparts (peaking around 0.3 compared to less than 0.2 for both medium and high educated women).

Class four represents the most 'modern' partnership form with a considerably high incidence of cohabitation before marriage, with a peak at age 25, when the probability of cohabitation is roughly 0.4. Thereafter, many unions translate into marriage, the probability of which peaks around age 31. We observe roughly similar patterns of partnership formation for women of all educational levels but there are differences in the timing of different partnership transitions. Women with low education tend to enter partnerships later than the more educated. Among low educated women, the probability of remaining never partnered stays close to 1 until age 21 while among medium and high educated women, this happens around age 18.

Additionally, women with low education are less likely to experience cohabitation before marriage; the peak of the probability of being in a cohabiting relationship is roughly 0.35, compared to 0.40 and 0.47 for women with high and medium education, respectively.

Finally, class five captures a more complex pattern of late partnership formation. Regardless of educational level, the probability of being never partnered does not decline until after age 25 and it never falls below 0.2. After age 25, union forms are varied; the most and the least educated are more likely to form cohabiting unions than marriages at all ages while women with medium education are more likely to be married after age 37. Finally, there is some evidence of union instability in this class at later ages.
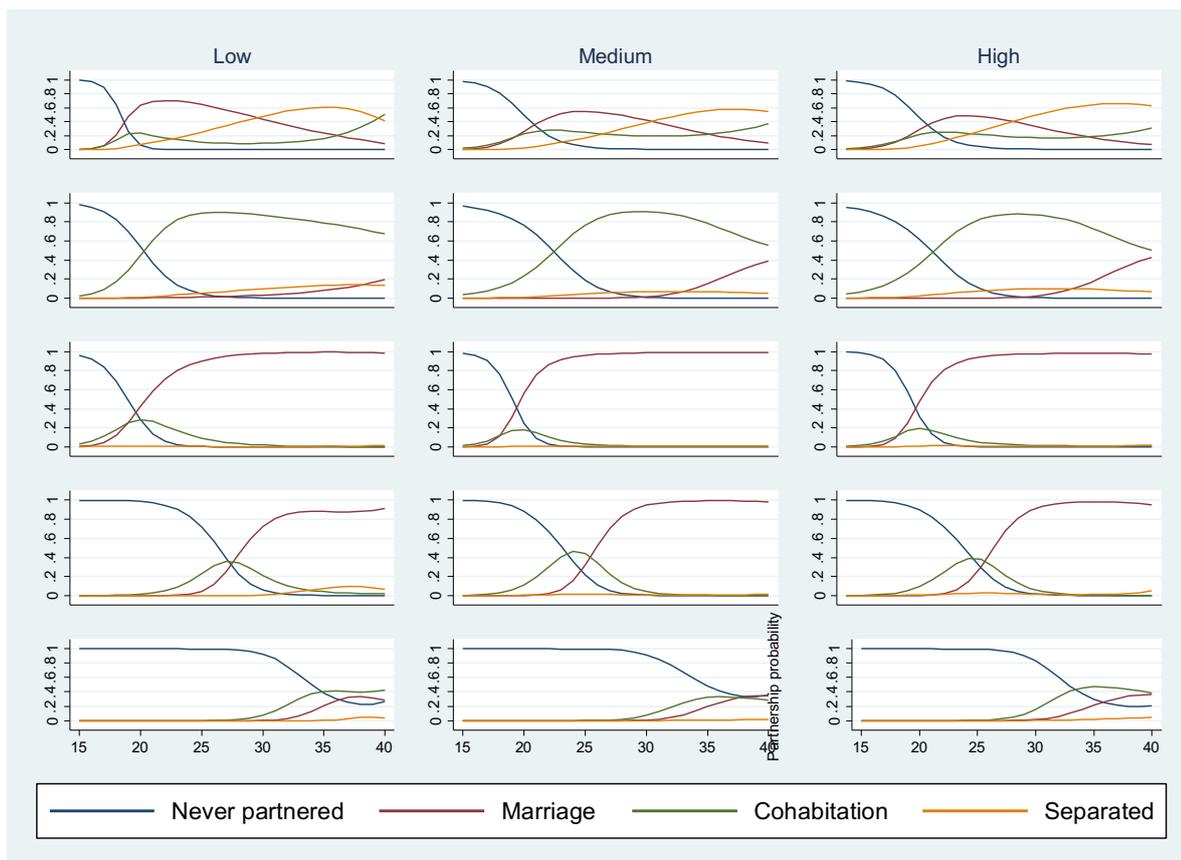
Note: Each row represents a latent class.
Note: Class 1: Early, varied partnerships (30.7%); Class 2: Early cohabitation with late transition to marriage (12.6%); Class 3: Early marriage with some cohabitation (21.5%); Class 4: Marriage preceded by cohabitation (25.0%); Class 5: Late and heterogeneous partnership forms (10.0%).

Figure 3. Results of the five-class Latent Class Growth Models by education (predicted proportion of women in each class)

To further facilitate the interpretation of educational differences across classes, table 5 presents the predicted probabilities of class membership by education. This table indicates that medium and highly educated women have a higher probability to belong to the 'early marriage with some cohabitation' class (class three) and to the 'marriage preceded by cohabitation' class (class four) than their low educated counterparts. Additionally, low educated women have a much higher probability (0.52) to belong to the 'early, varied partnerships class' (class one) compared to

those with medium education or higher (0.15 and 0.11, respectively). Last, we did not find large educational differences in the predicted probability of belonging to class two ('early cohabitation with late transition to marriage') and class five ('late and heterogeneous partnerships'). These results indicate that women with higher educational attainment generally experience more complex partnership patterns. The modal class for women with low education is class one ('early, varied partnerships') while for medium and highly educated women it is class three ('early marriage with some cohabitation') although their probability to belong to class four ('marriage preceded by cohabitation') is also larger than that of the other classes.

Table 5. Predicted probability of class membership by educational level

Class	Educational level		
	Low	Medium	High
1 Early, varied partnerships	0.52	0.15	0.11
2 Early cohabitation with late transition to marriage	0.15	0.12	0.12
3 Early marriage with some cohabitation	0.11	0.34	0.39
4 Marriage preceded by cohabitation	0.13	0.28	0.29
5 Late and heterogeneous partnership forms	0.09	0.11	0.09

Multi-state Event History Model

Table 6 describes the number of women who were at risk of each transition (total entering) and the number and proportion of those who experienced them. In the examined sample, 70% of never partnered women formed a cohabiting union while 28% got married. The remaining 9% of never partnered women remains never partnered until the end of the observation. Over two thirds of cohabiting unions transitioned to marriage while 22% ended in union dissolution. A similar proportion (25-26%) of marriages (both direct marriage and marriage that was preceded by cohabitation) ended with union dissolution. Finally, 75% of women who experienced union dissolution formed a new partnership.

Table 6. Number (and proportion, %) of women who experience each partnership transition

				Destination state					
		S	C	M	CM	D	R	no event	total entering
O r i g i n s t a t e	S	0	908 (70%)	363 (28%)	0	0	0	19 (2%)	1290
	C	0	0	0	621 (68%)	202 (22%)	0	85 (9%)	908
	M	0	0	0	0	91 (25%)	0	272 (75%)	363
	CM	0	0	0	0	163 (26%)	0	458 (74%)	621
	D	0	0	0	0	0	343 (75%)	113 (25%)	456
	R	0	0	0	0	0	0	0	0

The results of the multi-state event history model are summarised in table 7. Higher educated never partnered women have a higher risk of entering cohabitation and direct marriage than medium and low educated. Furthermore, education has a positive gradient on the transition from cohabitation to marriage; low and medium educated cohabiting women are 45% and 32% less likely, respectively, than their highly educated counterparts to marry their cohabiting partner. Following union dissolution, women with low education are less likely to find a new partner compared to highly educated women. Education does not have a significant influence on the dissolution of a cohabiting or a marital union (whether or not it was preceded by cohabitation).

Table 7. Result of the multi-state event history model, hazard ratios

	S → C	S → M	C → CM	C → D	M → D	CM → D	D → R
Education							
Low	0.73*	0.47***	0.55***	1.27	1.13	1.31	0.68*
Medium	0.70**	0.59**	0.68**	1.23	0.98	1.29	0.88
High (ref)							

*Note: *p < .05 **p < .01 ***p < .001*

Conclusion and Discussion

This paper qualitatively compared three methodological approaches (i.e. sequence analysis, latent class growth models, and multi-state event history models) to the analysis of life course data focusing on the influence of education on partnership experiences of Norwegian women born between 1955 and 1964. These methods have several similarities and differences. For example, sequence analysis and latent class growth models establish the relationship between education and the probability of belonging to certain groups (clusters or classes) based on women's partnership experiences. In our application, sequence analysis revealed six clusters based on women's partnership experiences (late, varied partnerships; early, direct, and unstable marriage; cohabitation; late, direct, and stable marriage; cohabitation followed by marriage; and early, direct, and stable marriage), latent class growth models suggested five partnership classes (early, varied partnerships; early cohabitation with late transition to marriage; early marriage with some cohabitation; marriage preceded by cohabitation; and late, heterogeneous partnership forms). Multi-state event history models do not classify individuals but rather examine the influence of education on each partnership transition thereby enabling us to draw conclusions about the changing influence of education over the early family life course.

Overall, the examined methods arrive at similar conclusions with respect to the influence of education on partnership experiences. For example, all three methods found that women with higher education are more likely to marry their cohabiting partner. Moreover, sequence analysis showed that the lower educated are more likely to belong to early union formation clusters than women with high education, who are more likely to form direct, stable marriages at later ages. This is in line with findings of LCGMs, which showed that low educated women are the most likely to belong to the 'early, varied partnerships' class (class one). Additionally, the results of multi-state models revealed that the more educated have a higher risk of experiencing direct marriage than the lower educated. Multi-state models showed that more educated women have a higher risk of finding a new partner following union dissolution than their lower educated counterparts. However, the results of the LCGMs did not suggest significant educational differences in the probability of belonging to the 'late and heterogeneous partnerships' class (class five). Similarly, we found no significant educational differences between the likelihood of belonging to the 'late, varied partnerships' cluster (cluster one) and the 'early, direct, and stable marriage' cluster (cluster six). Additionally, multi-state event history models showed that never partnered highly educated women have higher risks to enter

cohabitation than their lower educated counterparts. At the same time, these women are also more likely to marry their cohabiting partner and thus less likely to remain cohabiting. However, LCGMs did not find educational differences in the probability of belonging to class two ('early cohabitation with late transition to marriage').

The examined methods have different properties and approach studying the life course in a different way. In order to emphasise the strengths of each technique and to accommodate their limitations, the presented analyses could not have been implemented in exactly the same way for the three techniques. For example, the multi-state event history model and sequence analysis were estimated using monthly data while the LCGMs relied on yearly data due to computational issues. This implies that in LCGMs the number of transitions might be underestimated and some variation in life courses might be lost. Additionally, the multi-state event history model incorporated a time-varying education variable while the other two methods investigated the association between the highest level of education at the time of the interview (i.e. a time-constant variable) and partnership formation. Finally, while the multi-state event history model estimated the influence of education on first and higher order partnership transitions separately, the order of union is encoded in the sequences for LCGMs and SA.

These differences in the implementation of the analyses could potentially explain some of the differences in the results of the multi-state event history model and the other two methods but they cannot account for differences between the results of SA and LCGMs. However, it has been shown that SA and LCGMs should not be expected to give the same answer. Using simulated data, Warren, Luo, Halpern-Manners, Raymo, and Palloni (2015) showed that the number of trajectories these methods produce might differ from each other as well as from the true number of trajectories. Moreover, LCGMs and SA might assign the same individuals to different trajectory groups.

Another possible explanation for the differences in the results produced by the three methods may be related to specific choices made by the researcher during different stages of the analyses. For example, in sequence analysis, the researcher has an array of options to calculate the distance matrix, to define insertion, deletion, and substitution costs, as well as to perform the clustering. Each of these decisions might influence the outcome of sequence analysis. Similarly, in case of LCGMs, the method chosen to create the classes may alter the resulting latent classes. In case of the multi-state event history analysis, there is no need to make such arbitrary choices.

Even though the applications are not exactly the same and occasionally they provide somewhat different results, by illustrating the properties and application of the different techniques, we were able to identify similarities and differences between these methods with respect to their ability to address certain desirable aspects of studying the family life course. These are summarised in table 8.

Table 8. Summary of the properties of sequence analysis, latent class growth models, and multi-state event history analysis

	SA	LCGM	Multi-state Event History model
Transition intensities	(✓)	✘	✓
Classifying individuals	✓	✓	✘
Covariate information alters pattern	✘	✓	✓
Computationally simple	✓	✘	✓
Time-varying covariates	✘	✘	✓
Model based	✘	✓	✓
Protection against baseline misspecification	✓	✘	✓
Possibility to incorporate weights	(✓)	✓	✘

Note: The given method is ✓ able to, ✘ not able to or (✓) partially able to deal with this dimension of the family life course.

Based on this table, we can formulate broad recommendations for researchers choosing between different life course methods. First, sequence analysis is best applied to research questions which attempt to describe partnership behaviours of different groups of women and the overall associations of these groups with certain covariates. This can be achieved through the method's ability to classify individuals and allow for covariates to predict women's membership in the different clusters. Overall, fitting the model does not require a lot of computing power and because the procedure is not model based, the user is protected against baseline misspecification (i.e. no baseline needs to be specified). Although not presented in this paper, the method can also calculate transition intensities between the different states. As it is not possible to condition sequences, or more importantly transition probabilities, on covariate information or to allow for the incorporation of changing covariate information over the life course, this method cannot answer research questions relating to the changing influence of a variable over the life course. In other words, whereas LCGMs and multi-state event history models directly involve covariates thereby providing a better assessment of the net effect of a covariate of interest on different family life transitions, SA is not able to directly incorporate covariates. Last, it is not possible to take into account survey weights while computing distances between sequences (by definition this is a one to one comparison) although weights can be incorporated when comparing cluster sizes and in regression models.

Second, latent class growth models have a number of similar properties to sequence analysis. Its main advantage is that it is able to incorporate more complicated structures by, for example, allowing for covariate information to alter the shape of partnership trajectories. Additionally, LCGM is the only one among the three examined methods which allows the researcher to fully incorporate survey weights in the analysis. Unfortunately, the implementation of LCGMs is computationally intense and requires considerable computing power to estimate models for large datasets. Moreover, as LCGMs are model based, a greater degree of robustness is required particularly when estimating the shape of the growth curves.

On the other hand, this also means that a greater variety of fit-statistics is available than in sequence analysis, where the decision of the optimal number of clusters is more arbitrary than in LCGMs. Last, it should be noted that while LCGMs allow for testing the model performance via simulation approaches (e.g. Nylund et al., 2007), such a test is not available for sequence analysis. Thus, LCGMs are most suited to studying complex research topics where the aim is to identify differences in covariate effects between groups of individuals. The present paper has demonstrated this by extracting different classes of partnership behaviour and comparing the effect of educational attainment within these classes.

Finally, although multi-state event history models do not classify individuals in the same way as the previous two techniques, there are a number of distinct advantages to using this method. For example, the estimation of transition intensities allows for examining several transitions over the life course within the same model as well as for estimating the changing influence of covariates over the life course by allowing for the incorporation of time-varying covariates. This is one of the key advantages of multi-state event history models as neither sequence analysis, nor latent class growth models are capable of studying changing covariate effects over the life course. Additionally, the use of a stratified Cox model provides some protection against baseline misspecification. However, currently, it is not possible to incorporate survey weights using the *mstate* package. To conclude, multi-state event history models can best answer research questions specifically related to changing covariate effects over the life course. For example, as this paper has shown, it can estimate the changing influence of education on different partnership transitions over the early family life course.

The analyses presented in this study have some limitations. First, the multi-state event history model assumes that the hazards of the examined transitions for women with different educational level are proportional. This assumption might not be realistic. The multi-state event history model would allow for the incorporation of interaction effects between age and education in order to relax the assumption of proportional hazards. However, LCGMs and SA are unable to explicitly incorporate such interactions.[viii] To keep the models comparable, we refrained from including interactions between age and education in the multi-state event history model. Second, next to education, many factors may influence the timing and sequencing of partnership transitions. For LCGMs and SA, which included a time-constant education variable, the influence of other time-constant covariates on the timing and sequencing of the examined transitions could have been studied. However, in the examined dataset time-varying information, which was used in the multi-state event history model, could only be reconstructed for education. Including more covariates for LCGMs and SA but not for the multi-state event history model would not have facilitated the comparison of the methods and the results they produce. Researchers can build on this simple application and perform more complex analyses. Last, rather than comparing the three methods based on goodness of fit statistics or applying them to the same simulated dataset, we took a more applied approach and qualitatively compared the advantages and disadvantages of the three techniques. This approach was in line with the aim of the paper, namely, to give a qualitative overview of the strengths and weaknesses of these techniques and their applicability to answering different research questions.

Taken together, by comparing sequence analysis, latent class growth models, and multi-state event history models, this paper contributed to the discussion on the applicability of different methods for studying the life course. We showed that latent class growth models and multi-state event history models are a useful addition to life course researchers' methodological toolkit and that these methods can address certain research questions better than the more commonly applied sequence analysis or simple event history analysis. In particular, we have stressed the types of research questions that may be better addressed using these techniques which provide new insights in the field of life course studies.

Acknowledgements

References

Abbott, A., & Tsay, A. (2000). Sequence analysis and optimal matching methods in sociology: Review and prospect. *Sociological Methods & Research, 29*(1), 3-33. https://doi.org/10.1177/0049124100029001001

Andersen, P. K., & Keiding, N. (2002). Multi-state models for event history analysis. *Statistical Methods in Medical Research, 11*(2), 91-115. https://doi.org/10.1191/0962280202SM276ra

Baizán, P., Aassve, A., & Billari, F. C. (2003). Cohabitation, marriage, and first birth: The interrelationship of family formation events in Spain. *European Journal of Population, 19*(2), 147-169. https://doi.org/10.1023/A:1023343001627

Baizán, P., Aassve, A., & Billari, F. C. (2004). The interrelations between cohabitation, marriage and first birth in Germany and Sweden. *Population and Environment, 25*(6), 531-560. https://doi.org/10.1023/B:POEN.0000039064.65655.3b

Barban, N., & Billari, F. C. (2012). Classifying life course trajectories: A comparison of latent class and sequence analysis. *Journal of the Royal Statistical Society Series C Applied Statistics, 61*, 765-784. https://doi.org/10.1111/j.1467-9876.2012.01047.x

Billari, F. C. (2001a). The analysis of early life courses: Complex descriptions of the transition to adulthood. *Journal of Population Research, 18*(2), 119-142. https://doi.org/10.1007/BF03031885

Billari, F. C. (2001b). Sequence analysis in demographic research. *Canadian Studies in Population, 28*(2), 439-458.

Billari, F. C. (2005). Life course analysis: Two (complementary) cultures? Some reflections with examples from the analysis of the transition to adulthood. *Advances in Life Course Research, 10*, 261-281. https://doi.org/10.1016/S1040-2608(05)10010-0

Billari, F. C., & Piccarreta, R. (2001). Life courses as sequences: an experiment in classification via monothetic divisive algorithms. In S. Borra, R. Rocchi, M. Vichi & M. Schader (Eds.), *Advances in Classification and Data Analysis* (pp. 351-358). Berlin and New York: Springer Verlag. https://doi.org/10.1007/978-3-642-59471-7_43

Billari, F. C., & Piccarreta, R. (2005). Analyzing demographic life courses through sequence analysis. *Mathematical Population Studies, 12*(2), 81-106. https://doi.org/10.1080/08898480590932287

Bonetti, M., Piccarreta, R., & Salford, G. (2013). Parametric and nonparametric analysis of life courses: An application to family formation patterns. *Demography, 50*(3), 881-902. https://doi.org/10.1007/s13524-012-0191-z

Brzinsky-Fay, C., & Kohler, U. (2010). New developments in sequence analysis. *Sociological Methods & Research, 38*(3), 359-364. https://doi.org/10.1177/0049124110363371

Brzinsky-Fay, C., Kohler, U., & Luniak, M. (2006). Sequence analysis with Stata. *The Stata Journal, 6*(4), 435-460.

Calinski, R. B., & Harabasz, J. A. (1974). A dendrite method for cluster analysis. *Communications in Statistics-Theory and Methods, 3*(1), 1-27. https://doi.org/10.1080/03610927408827101

de Wreede, L. C., Fiocco, M., & Putter, H. (2011). mstate: An R package for the analysis of competing risks and multi-state models. *Journal of Statistical Software, 38*(7), 1-30. https://doi.org/10.18637/jss.v038.i07

Duda, R. O., & Hart, P. E. (1973). *Pattern Classification and Scene Analysis*. New York: Wiley.

Elzinga, C. H., & Liefbroer, A. C. (2007). De-standardization of family-life trajectories of young adults: A cross-national comparison using sequence analysis. *European Journal of Population, 23*(3-4), 225-250. https://doi.org/10.1007/s10680-007-9133-7

Halpin, B. (2014). SADI: Sequence analysis tools for Stata, Working Paper WP2014-03, Department of Sociology, University of Limerick, http://www.ul.ie/sociology/pubs/wp2014-03.pdf.

Heuveline, P., & Timberlake, J. M. (2004). The role of cohabitation in family formation: The United States in comparative perspective. *Journal of Marriage and Family, 66*, 1214-1230. https://doi.org/10.1111/j.0022-2445.2004.00088.x

Hougaard, P. (1999). Multi-state models: A review. *Lifetime Data Analysis, 5*(3), 239-264. https://doi.org/10.1023/A:1009672031531

Jung, T., & Wickrama, K. A. S. (2008). An introduction to Latent Class Growth Analysis and Growth Mixture Modeling. *Social and Personality Psychology* Compass*, 2*(1), 302-317. https://doi.org/10.1111/j.1751-9004.2007.00054.x

Nylund, K. L., Asparouhav, T., & Muthen, B. O. (2007). Deciding on the number of classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 535-569. https://doi.org/10.1080/10705510701575396

Perelli-Harris, B., & Gerber, T. P. (2011). Nonmarital childbearing in Russia: Second Demographic Transition or Pattern of Disadvantage? *Demography, 48*(1), 317-342. https://doi.org/10.1007/s13524-010-0001-4

Perelli-Harris, B., Kreyenfeld, M., & Kubisch, K. (2010). Harmonized Histories: Manual for the preparation of comparative fertility and union histories. *Rostock, MPIDR Working Paper WP-2010-011.*

Perelli-Harris, B., & Lyons-Amos, M. (2015). Changes in partnership patterns across the life course: An examination of 14 countries in Europe and the United States. *Demographic Research, 33*, 145-178. https://doi.org/10.4054/DemRes.2015.33.6

Perelli-Harris, B., Sigle-Rushton, W., Kreyenfeld, M., Lappegård, T., Keizer, R., & Berghammer, C. (2010). The educational gradient of childbearing within cohabitation in Europe. *Population and Development Review, 36*(4), 775-801. https://doi.org/10.1111/j.1728-4457.2010.00357.x

Piccarreta, R., & Billari, F. C. (2007). Clustering work and family trajectories using a divisive algorithm. *Journal of Royal Statistical Society, A, 170*(4), 1061-1078. https://doi.org/10.1111/j.1467-985X.2007.00495.x

Putter, H., Fiocco, M., & Geskus, R. B. (2007). Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine, 26*(11), 2389-2430. https://doi.org/10.1002/sim.2712

Putter, H., van der Hage, J., de Bock, G. H., Elgalta, R., & van de Velde, C. J. H. (2006). Estimation and prediction in a multi-state model for breast cancer. *Biometrical Journal, 48*(3), 366-380. https://doi.org/10.1002/bimj.200510218

Steele, F., Kallis, C., Goldstein, H., & Joshi, H. (2005). The relationship between childbearing and transitions from marriage and cohabitation in Britain. *Demography, 42*(4), 647-673. https://doi.org/10.1353/dem.2005.0038

Wang, C. P., Hendricks Brown, C., & Bandeen-Roche, K. (2005). Residual diagnostics for growth mixture models: examining the impact of a preventive intervention on multiple trajectories of aggressive behavior. *Journal of the American Statistical Association, 100*(471), 1054-1076. https://doi.org/10.1198/016214505000000501

Warren, J. R., Luo, L., Halpern-Manners, A., Raymo, J., & Palloni, A. (2015). Do different methods for modelling age-graded trajectories yield consistent and valid results? *American Journal of Sociology, 120*(6), 1809-1856. https://doi.org/10.1086/681962

Endnotes

[i] This paper used the version that is available in the Harmonized Histories (Perelli-Harris, Kreyenfeld, & Kubisch, 2010).

[ii] Throughout this study never partnered women are defined as those who have never lived in a co-residential union for at least three months.

[iii] $\frac{1}{\left(\frac{Je(2)}{Je(1)}\right)} = 1 + \frac{T^2}{N_1 + N_2 - 2}$, where N_c denotes the number of observations in cluster *c*.

[iv] If two partnership states are present in the same year, the higher value is selected. This means that short episodes of e.g. cohabitation or being separated before re-partnering will be missed.

[v] We use yearly (instead of monthly) intervals to reduce the size of the dataset and to increase the speed of estimation. Robustness checks for similar analyses have shown that the reduction of information from monthly to yearly intervals do not substantially influence the results (Perelli-Harris & Lyons-Amos, 2015).

[vi] Exploratory analyses revealed that higher order classes tend to be sparsely populated with limited interpretability.

[vii] Although it can be argued that multi-state models predict group membership in terms of state occupation probabilities, in multi-state models individuals move from one state to the next. However, in SA and LCGMs each individual can only belong to one cluster or class.

[viii] It would be possible to build sequences of changes in educational level and examine these sequences together with sequences of partnership states.

TREE (Transitions from Education to Employment): A Swiss Multi-Cohort Survey

Andrés Gomensoro   TREE, University of Bern, Switzerland
andres.gomensoro@soz.unibe.ch
Thomas Meyer   TREE, University of Bern, Switzerland

Abstract

TREE (Transitions from Education to Employment) is a Swiss nationwide longitudinal study that follows two cohorts of compulsory school leavers throughout their transitions from education to employment and middle adulthood. To date, the first cohort survey (initial N=6,343) based on the Swiss PISA 2000 sample has covered a 14-year period from age 15 to 29 (nine follow-up surveys). The second cohort survey started in 2016 (with one follow-up survey in spring 2017 so far; initial N=9,762) and is based on a large national representative sample of students (N=22,378) who sat a mathematics test at the end of the ninth grade (approximately 15 years old). TREE is designed to provide comprehensive data for the analysis of post-compulsory education, employment and other pathways (e.g. family and household situation, income/financial situation, critical life events, social integration and participation, psycho-social personal characteristics, health and wellbeing). As a social science infrastructure project of national and international importance, its data is freely available to the scientific community at large. This paper provides an overview of the TREE study with a specific focus on the latest data release (September 2016) for the first TREE cohort.

Keywords

Education, occupation, pathways, cohort survey, longitudinal research, PISA, Switzerland, life course

Introduction

Transitions from Education to Employment (TREE)[i] is a unique and innovative multi-user, multi-cohort survey in Switzerland that follows two large representative samples (>6,000 respondents each) of compulsory school leavers throughout their transitions from education to employment and middle adulthood.[ii] The TREE survey has an interdisciplinary approach and covers a wide range of topics. The first cohort survey (TREE1) started in 2000 on the basis of the Swiss sample that participated in the Programme for International Student Assessment (PISA) and has since followed individuals from an average age of 15 to 29 (another panel wave is planned at an average age of

34). The second cohort (TREE2) was launched in 2016 and is based on a large sample of ninth graders who underwent a standardised national mathematics test at the end of their compulsory schooling. With this extension to a multi-cohort design, Switzerland will be among the few countries in the world in which comparative inter-cohort analyses can be carried out. As one of Switzerland's major social science data infrastructure projects, TREE is mainly funded by the Swiss National Science Foundation (SNF). The TREE1 data are among the country's most widely used datasets. As of 2014, the project has been located at the Institute of Sociology and the Institute of Educational Science at the University of Bern.

This study profile aims to introduce the main design features of the TREE survey and to provide information on the latest data release (September 2016) for the first TREE cohort (TREE1), which is available free of charge for researchers in Switzerland and abroad (www.forscenter.ch). First, we focus on the survey's origin and its main objectives. Second, we describe the multi-cohort design and sampling procedure, data collection methods and samples, response rates and weighting. Finally, we introduce the survey content and data use of the TREE1 sample.

Origin and objectives of TREE

The TREE survey was launched in response to a recommendation the OECD made for Switzerland after its international review of transitions from initial education to working life (TIEW). Looking at the country report for Switzerland, which revealed substantial research gaps on the subject matter, the OECD expert committee noted:

> *"The opportunity should not be missed to equip Switzerland with a longitudinal survey of transitions at national level [...]. Transition pathways to employment are becoming increasingly complex. To understand young people's decisions and options, and to take them into account in policy decisions, appropriate analytical instruments are needed."* (OECD, 1999:53)

The first TREE cohort survey was the empirical answer to this recommendation. In 2000, Switzerland participated in the first PISA study, which was designed to assess general literacy skills (main focus on reading) among students who have reached the end of compulsory school (age 15/ninth grade). Along with Canada and Denmark, Switzerland has been the only participating country to follow the respondents of its PISA 2000 sample on their pathways through post-compulsory education and training and their transition to the labour market and to adult life.[iii] Since then, the TREE survey has become one of Switzerland's major social science data infrastructure projects (see later sections for detailed information on sampling procedures, data collection methods, response rates, attrition and weighting).

Four main objectives guide the TREE survey. First, TREE aims to precisely describe post-compulsory educational and employment transitions and pathways at the national and regional levels. With the continuation of the survey over the years, TREE1 has been gradually enlarging the perspective and scope of cohort observation into the "middle age" life phase. With the last survey panel carried out in 2014, the overall span of observation had reached almost 15 years, and another survey wave is scheduled in 2019 (at an average age of 34; see figure 1). Later transitions such as those from tertiary education to work or the consolidation of work careers can thus be studied in detail. In addition, the long observation period opens up new opportunities for conducting life-course research on the crossover effects of various life domains; for instance, social relations, education/career, family, psycho-social and physical health, socio-political integration, and commitment and attitudes.

Second, TREE aims to explain transitions and trajectories. Its context-rich data allow researchers to analyse explanatory factors such as socioeconomic and socio-cultural background, skills and academic achievement, personality traits, career ambitions as well as features of the school environment at the end of compulsory education and education/employment careers pursued thereafter. PISA 2000 is an excellent baseline survey for TREE. To begin with, it is a large, nationally representative sample of school leavers in Switzerland. Standardised skills measures as well as school and student characteristics at the end of compulsory school can therefore be linked to later educational outcomes and other life-domain trajectories. Moreover, by extending TREE to a multi-cohort survey (TREE2 started in 2016), Switzerland will be equipped with large-scale representative multi-cohort data on school-to-work and early life-course transitions, thereby allowing the study of specific cohort or systemic effects such as economic cycles, demographic trends or major reforms of the education system.

Third, TREE aims to provide longitudinal datasets to Swiss and international scholars along with a selection of key findings that can serve as a reference base for further in-depth analyses (see, e.g. Bergman, Hupka-Brunner, Keller, Meyer & Stalder 2011; Scharenberg, Hupka-Brunner, Meyer & Bergman 2016; Scharenberg, Rudin, Müller, Meyer & Hupka-Brunner 2014).

Fourth, data provided by TREE also contribute to a better understanding of transitions for policymakers, educational and economic

stakeholders, teaching and counselling staff, civil society actors and so forth.

Overall design

TREE follows two nationally representative cohort samples of compulsory school leavers at an interval of 16 years, both starting at approximately age 15 (see figure 1). The first cohort (TREE1) is based on a sample of 6,343 ninth graders who participated in the PISA test in spring 2000 and left compulsory school shortly thereafter. As mentioned above, the TREE1 sample is defined as a subset of PISA 2000 respondents (PISA 2000 *N*=13,463; see Sacchi 2011, Renaud, 2002 and Adams & Wu, 2002 for detailed information on the PISA 2000 sample in

Switzerland). The TREE1 sample includes all participants of PISA 2000 who

- had attended a regular public school at the lower secondary level at the time of the PISA survey;[iv]
- had completed ninth grade and left compulsory education at the end of the 1999/2000 school year;
- had agreed to participate in the TREE survey.

The TREE1 sample was surveyed nine times between 2001 and 2014, up to 2007 at yearly intervals. A tenth follow-up survey is planned for 2019, by which time the sample will have reached an average age of approximately 34 (see figure 1).

Figure 1: TREE multi-cohort design 2000–2020 (TREE 2016b)
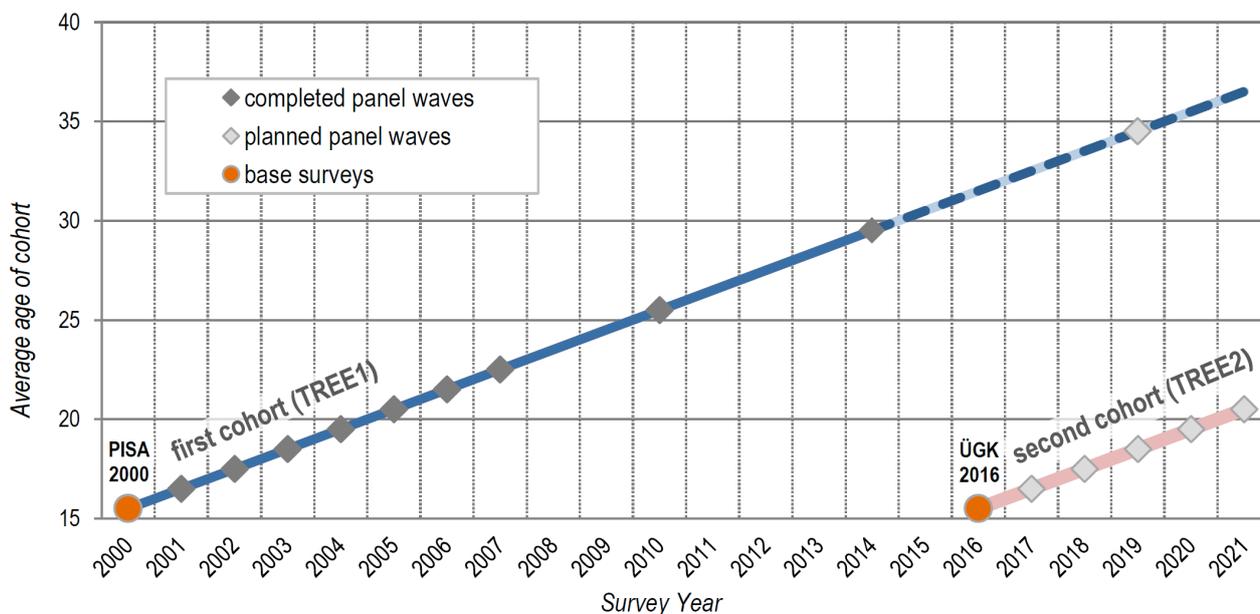
ÜGK 2016 = Acronym for Swiss National Mathematics Skills Assessment 2016

The second cohort survey (TREE2) was launched in 2016. It is based on a national large-scale assessment (LSA) in mathematics[v] carried out among more than one-quarter of all ninth graders in Switzerland (gross sample≈25,000, response rate 92%, average age 15). The sample is representative of both national and cantonal levels.[vi] TREE was granted permission to ask for the tested students' consent to be contacted for the TREE2 follow-up survey, and if so, for them to provide contact information. Close to 14,000 students agreed to be

contacted (over 60% of all 2016 math LSA respondents). The TREE2 survey design essentially follows that of TREE1 for reasons of comparability. As with TREE1, the TREE2 sample is to be surveyed at annual intervals at its initial stage. The initial data will be publicly available by early 2019. Given that TREE2 has only just started and that so far only data on the first cohort is available to researchers, the following sections mainly focus on TREE1.

TREE1 survey design

The TREE1 panel waves were invariably carried out toward the end of the Swiss school year (i.e., between February and June); the first four (2001–2004) were conducted mostly by mail (standardised written questionnaires).[vii] Non-respondents were first reminded by mail and, if necessary, a second time by telephone (see the survey flowchart in figure 2). Respondents unable or unwilling to complete the written questionnaire were offered the opportunity to reply by means of a telephone interview as a secondary method.[viii]

Starting with panel wave five, the survey design shifted to a mixed-mode design that relied on a combination of computer-assisted telephone interviewing (CATI) and complementary written questionnaires tailored to specific education, training or job situations so as to accommodate the growing diversity of individual education and employment careers (see figure 2). The CATI interviews mainly collected episodic data on education and employment situation and history. The complementary questionnaire, sent out immediately after each CATI interview, aimed to gather in-depth descriptions and assessments of the principal activity (usually education or employment) specified in the CATI interview. CATI data were used to determine which type of supplementary written questionnaire would be sent to respondents. Respondents who could not be reached or who refused to answer questions by telephone were sent a written questionnaire as an alternative option. At each stage of the survey, non-responses were followed up by at least one reminder. The duration of the CATI interview ranged from 20 to 30 minutes;[ix] filling out the complementary written questionnaire required another 20 minutes. As a general rule, total survey time per respondent (both interview and self-administered questionnaire) was not to exceed one hour.[x]

Measures to increase panel retention

With regard to panel maintenance and retention, TREE1 has adopted a strategy that relies on respondents' intrinsic motivation (i.e. to generally abstain from offering incentives). Measures to minimise attrition/maximise sample retention under this general strategy have included the following:

- Provision of alternative response modes, including short versions of the questionnaire, in order to lower the threshold for participation.
- Continuous extensive follow-up and update of respondents' contact information.
- From 2003 on, respondents began receiving a newsletter a few weeks prior to each survey that informed them about new results, publications and news related to the project.[xi] These newsletters, carefully customised to the target audience both linguistically and visually (layout), contributed not only to encouraging participation but also to updating respondents' contact data.
- Hotlines were installed for all three survey languages (German, French and Italian), which respondents could contact at any time during the field phase.
- Comprehensive interviewer training, including strategies on how to convince respondents to participate.
- In five of the nine waves, custom TREE ballpoint pens were included with the questionnaires as a small symbolic incentive and for the practical purpose of completing them.
- Regularly updated pages in all survey languages expressly addressing respondents on the project website (www.tree.unibe.ch).
- TREE has regularly made an effort to communicate to the respondents that their participation is not only of great importance but indeed indispensable to the project.
- Carefully individualised and immediate response to suggestions, criticism and questions voiced by the respondents, assuring them that their feedback is taken seriously.

Figure 2: TREE1 survey design, 2000–2014 (TREE 2016b)
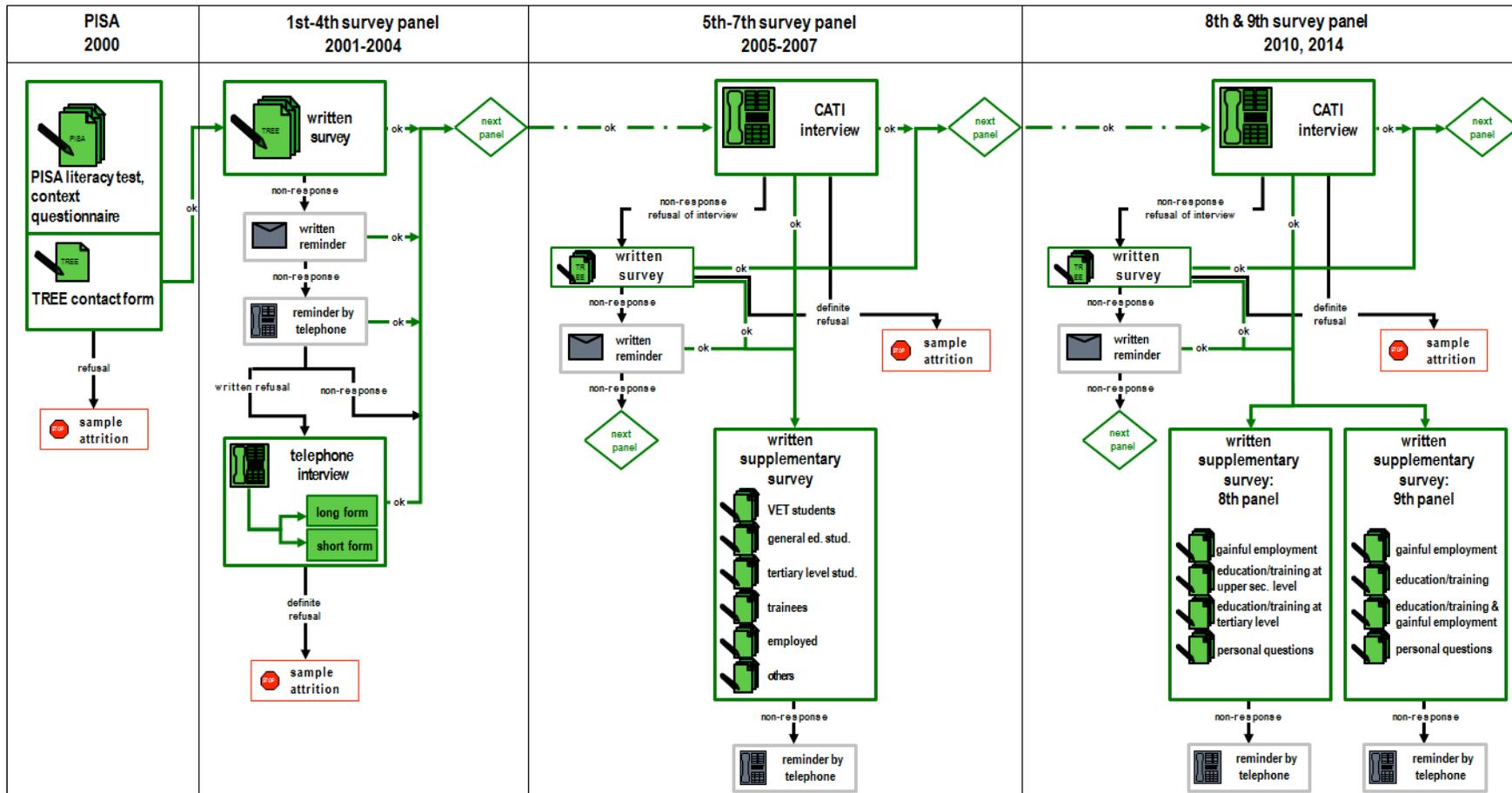
Sample, response rates and weighting

As mentioned above, TREE1 is based on the Swiss PISA 2000 sample. PISA 2000 minimal sampling standards required national samples of 5,000 to 6,000 students at age 15 (Adams & Wu, 2002). Switzerland by far exceeded these standards. The Swiss PISA 2000 sample size exceeded 13,000 (approximately one-sixth of the relevant student population) owing to a substantial oversampling of ninth graders and students from selected cantons (for detail, see Renaud, 2002). As PISA 2000 basically granted anonymity to respondents, TREE had to ask for their consent to be contacted and for their contact data (e.g., mailing address, email address, phone number). Rates of consent varied considerably across survey regions. Thus, whereas in Italian-speaking Switzerland 81% of the PISA sample gave their consent (and their contact data) to participate in TREE, response rates were much

lower in other cantons (e.g., 32% in Zurich; Sacchi, 2011). Overall, this resulted in a relatively low rate of consent of only 54%.

Throughout the first seven panel waves from 2001 to 2007, TREE1 achieved wave-specific response rates of 85–89% (see table 1). These rates are remarkably high considering the heterogeneous sample structure. In waves eight and nine, response was considerably lower (75% and 71%, respectively); however, given the broadening of the survey interval from one to three to four years, the 2010 and 2014 response rates can also be considered satisfactory. Cumulative response did not dramatically decrease between 2010 and 2014 (from 54% to 50%). After nine follow-up surveys, the gross sample size that can be contacted to participate in the next survey panel in 2019 is still approximately 4,400.

Table 1: TREE1 response rates, 2000–2014

Survey year	2000	2001	2002	2003	2004	2005	2006	2007	2010	2014
Samples (N)										
TREE gross sample	11,710 *	6,343	5,944	5,609	5,345	5,060	4,852	4,659	4,571	4,402
Responses	6,343**	5,528	5,206	4,877	4,679	4,506	4,133	3,979	3,423	3,142
Response rates (%)										
% of wave	54%	87%	88%	87%	88%	89%	85%	85%	75%	71%
% of 2001 survey			82%	77%	74%	71%	65%	63%	54%	50%

* Sample base (after exclusion of participants not satisfying TREE sampling selection criteria; see the "Overall design" section).

** PISA respondents consenting to participate in the TREE panel survey (including provision of contact data).

In the first two years, 80% of the respondents completed the written questionnaire (see table 2). The proportion of returned questionnaires dropped to roughly 70% during the third and fourth survey panels, whereas the number of responses by way of short telephone interviews rose sharply. After the changes in survey design described in the "TREE1 survey design" section, the share of the sample that responded to the full survey (CATI or basic written

questionnaire plus supplementary questionnaire) remained at 70%. Between one-fifth and one-sixth participated by telephone only. Thus, TREE achieved an exceptionally high level of overall response by offering potential dropouts alternative and/or streamlined modes of participation. The high overall response rate confirms the importance TREE gave to sample maintenance (see the "Measures to increase panel retention" section above).

Table 2: TREE1 response rates from 2001 to 2014 by type of response

Survey year	2001	2002	2003	2004	2005	2006	2007	2010	2014
Written questionnaire	80%	80%	71%	71%					
Telephone questionnaire, long form	6%	4%	3%	5%					
Telephone questionnaire, short form	2%	3%	13%	12%					
CATI/basic questionnaire and supplementary questionnaire *					71%	69%	68%	60%	59%
CATI/basic questionnaire only*					19%	16%	17%	15%	12%
No response, single wave refusal (remaining in the sample)	7%	8%	10%	8%	9%	13%	13%	23%	23%
No response (sample attrition)	6%	5%	3%	4%	2%	2%	2%	2%	5%
Total *N* (=100%; gross sample)	6,343	5,944	5,609	5,345	5,060	4,852	4,659	4,571	4,402

* 2007/2010/2014 panel survey: basic and supplementary questionnaire were merged into one questionnaire.

As in any panel survey, sample attrition has varied substantially by a number of socio-demographic, socio-economic and academic performance characteristics. For instance, attrition rates among low achievers, young men and migrants are substantially above average (Sacchi 2011:13ff.). As information about both respondents and non-respondents is abundant in our case, such biases can be modelled and corrected in fair approximation. Moreover, it is advantageous that non-response was very low in the baseline survey (PISA 2000: approximately 5%). All information about respondents and survey conditions obtained in the baseline survey are thus available for correcting for the high level of non-response with regard to respondents' consent to join the TREE panel. In addition, all the information from any given TREE survey wave can be used to correct for non-response biases that occur in subsequent waves (see Sacchi 2011 for more detailed information).

Demographic features

To describe some characteristics of the TREE1 sample, table 3 provides the main demographic features in 2000 (baseline survey) and 2014 (ninth panel wave sample). We can see that fourteen years after leaving compulsory school, the sample size still allows for analysis of a wide range of subgroups with regard to education, occupation, family and other situations. Note that weights should be taken into account to ensure representativeness.

Table 3: Demographic features of TREE1 sample in 2000 (~15 years old) and 2014 (~29 years old)

TREE1 baseline survey, 2000		*N* *(unweighted)*	*%* *(unweighted)*
Total		6,343	100
Gender	Female	3,440	54
	Male	2,903	46
Year of birth	1983	757	12
	1984	3,399	54
	1985	2,052	32
	Other years	111	2
Academic track attended at lower secondary level	Extended requirements	4,358	68
	Basic requirements	1,626	26
	No formal grouping	357	6
PISA reading literacy score	Very low	222	4
	Low	587	9
	Medium low	1,315	21
	Medium high	2,144	34
	High	1,555	25
	Very high	514	8
Migration background	Swiss native*	3,776	60
	Second generation**	1,560	25
	Migrant***	900	14
Language region	German	2,970	47
	French	2,540	40
	Italian	833	13
Rural vs. urban area	Rural	2,072	33
	Urban	4,271	67
TREE1 9th panel wave sample, 2014 (~29 years old)			
Total		3,142	100
Educational attainment	Compulsory	87	3
	Secondary	1,161	37
	Tertiary	1,881	60
Gainful occupation	Full-time work (≥90%)	2,074	66
	Part-time work (<90%)	712	23
	Other situation	356	11
Marital status	Single	2,202	70
	Married	893	28
	Other status	46	1
Parenthood	Yes	767	24
	No	2,374	76

- Participants and their parents born in Switzerland.
  ** Participants born in Switzerland with at least one parent born abroad.
  *** Participants born abroad.

Theoretical and conceptual framework

TREE collects data on education, employment and family transitions within an interdisciplinary and general life-course research framework (Baltes, 1990; Blossfeld & von Maurice, 2011; Fend, Berger & Grob, 2009) by taking into account the complex interaction of objective and subjective factors at the individual, organisational/institutional and societal levels. Life course approaches are held to be promising when it comes to disentangling effects of age, cohort and temporal effects (Baltes, 1990; Elder, 1994; Levy, Ghisletta, Le Goff, Spini & Widmer, 2005). Moreover, they build conceptual bridges between sociological and psychological issues as well as between individual development and social structures (Blossfeld & von Maurice, 2011). TREE features a broad variety of concepts, thus making itself an excellent tool for inter-disciplinary research across disciplines such as sociology, educational sciences, psychology, economics, health and life-course studies (see TREE 2016b for further information). The insight gained with regard to the circumstances and mechanisms of transitions from initial education to middle adulthood provide sound empirical evidence for specific educational, labour market and social policy measures.

With regard to education, the concepts and scales adopted in the TREE surveys refer to sociological theories of status reproduction on the one hand (Bourdieu, 1977, 1984; Bowles & Gintis 1976; Willis 1977; for an overview on the current state of research, see Draelants & Ballatore, 2014 or Winkle-Wagner, 2010) and of decision-making in rational choice approaches (Baumert & Schümer, 2002; Blossfeld & Shavit, 1993; Boudon, 1974; Breen & Goldthorpe 1997; Ramseier & Brühwiler, 2003) on the other. These theories model the effects of social origin on educational decisions at each interface of the education system. In the fields of psychology and educational sciences, TREE draws on developmental theories that deal with educational and occupational socialisation (Heinz, 1984; Ulich, 1991) as well as with the ways in which youths cope with, in Bronfenbrenner's (1979) terms, "ecological transitions". Other concepts used to explain educational pathways derive from theories of self-concept (Eccles, Vida & Barber, 2004; Greve, 2000), institutional resources and constraints (which may lead to individual stress; Semmer, 1997), wellbeing (e.g., self-esteem,

depression, attitude towards life, personal values; see Fischer, 2006; Hascher, 2004) and critical life events (e.g., relocation, divorce/separation of parents, illness, death of a relative; see table 4 and Filipp, 1995).

When it comes to analysing labour market entry, TREE relies on concepts and instruments used by classical human capital theory (Becker, 1964), signal(ling) and social capital theories (Coleman, 1988, 1990; Spence, 1973), theories of discrimination (Arrow, 1994; Becker, 1957/1971) and school-to-work transition models based on labour market economics (OECD, 2000; Ryan, 2001).

On the basis of these theoretical frameworks, TREE1 has been designed to collect data on post-compulsory education life trajectories (see table 4). First, it provides highly detailed information on academic skills and achievement at the end of compulsory school (based on the PISA 2000 survey) as well as in-depth information on post-compulsory educational pathways. In addition, data on detailed individual characteristics such as gender, social and migration background and place of residence as well as psychological and psycho-social variables and scales such as personal values with regard to family, education and work, psycho-social resources (self-esteem, positive attitude towards life and coping) and social support are available. This allows researchers to analyse the influence of compulsory educational outcomes, of schooling contexts and of individual characteristics on later educational, occupational and family outcomes and/or trajectories.

Second, TREE1 provides highly detailed longitudinal information on gainful occupation. The new 2016 data release includes a data file on month-by-month episodes of gainful employment that covers an observation period of twelve years (2003 to 2014). Employment data include type of employment (hours of work per week, type of employment contract, wage), characteristics of the job and of the work team, challenges and resources at the workplace, subjective assessment of gender and equal opportunity, social support and job satisfaction (TREE 2016b). It further allows one to link job career development to both macro conditions of employment and subjective assessment of the individual job situation.

Third, TREE offers comprehensive longitudinal information on family life (parenting, household composition, financial situation of the household,

relocations) and civil status that opens promising research opportunities on interdependence between life domains.

The instrumentation of the TREE survey(s) is guided by the principle of comparability to previous waves and between cohorts. Consequently, adjustments of the survey's instruments are only made to explore the sample's progression into new biographical phases or to enhance research opportunities in terms of international comparability (see "TREE1 survey design" and "Call for instruments" sections). In the first phase (from 2001 to 2003), education and employment transition patterns at the interface of compulsory school to upper secondary education (or other activities) were at the centre of attention. During this first stage, the main focus was on the underlying reasons for typical trajectories as well as consequences of irregular or critical educational careers, particularly with regard to premature

dropout (young people who fail to graduate from a post-compulsory education or training programme). During the second stage of TREE (from 2004 to 2007), the study centred on the transition from upper secondary-level education (be it vocational or general) to working life or tertiary education. In the third phase from 2008 to 2014, it was particularly important to capture in detail individual integration into the labour market, education and continuing education in mid-career as well as family transitions.

Table 4 shows the main concepts employed by TREE. A separate document titled Concepts & Scales (TREE 2016a) and the codebooks specify in detail the concepts and items that have been used in the nine panel waves carried out so far. All methodological documents necessary for data use are included in the TREE1 public use dataset package.

Table 4: Concepts and scales, TREE1, 2001–2014

Education/training

Description of education/training situation, certifications and pathways
− Education episodes (type of education/traineeship, starting/end date, change of curriculum)
− Description of education/training context (size of training firms, commuting distance to school/firm, skills/competences of trainers, teachers and supervisors, etc.)
− Certifications (types of certification, date of certification)

Challenges and resources in education and vocational education and training (VET) traineeship
− Caused by education/training conditions, by environment, etc.
− Freedom of action, cooperation, comprehensiveness of tasks, variety of tasks, etc.

Learning styles
− Concentration, willingness to make an effort, etc.

Equal opportunity in schools and VET training firms
− Gender and nationality

Satisfaction with regard to education/training
− In general, with the school/training firm, with the class, etc.

Employment

Description of employment situation and working conditions
− Employment episodes (type of employment, starting/end date)
− Terms of employment (employment status, working hours, salary, position in the hierarchy, type of employment contract, commuting distance, etc.)
− Structural data on employers (firm size, economic sectors, etc. [added from third-party sources])

Challenges and resources at the workplace
− Caused by work organisation, by working environment, etc.
− Freedom of action, cooperation, comprehensiveness of tasks, variety of tasks, etc.

Job search and application strategies
− Job search strategies, application activities, duration of job search, etc.

Equal opportunities of employer
− Gender, nationality, skills, certifications

Satisfaction with regard to work and profession
− With professional career, employing firm, etc.

Other life situations and activities

Description of other life situations
− Not in education, employment or training (NEET) episodes (type of activity, starting/end date)
− Desired activity
− Family situation/status (housing situation, civil status, children, nationalities/naturalisation, financial situation, etc.)

Significant life events
− With regard to education, work and private life (relocation, separation, divorce, illness, death)

Personal characteristics and values
− Persistence, self-efficacy, coping ability, ambition, etc.
− Value orientation (partnership/family, leisure, work)

Health and wellbeing

Social integration and participation

Social relations and social support with regard to education/training
− Privately, at school and in VET training firms

Social relations and social support with regard to work
− Privately and colleagues at work

Relationships

Societal and cultural participation

Call for instruments

On several occasions TREE has launched *calls for instruments*. These invited researchers to propose additional concepts, scales, modules or single items that might enhance the analytic potential and (inter)national comparability of the TREE dataset. This approach ensures that concepts and scales provide answers to current and relevant research questions from the scientific community in a variety of disciplines. Furthermore, involving the scientific community in the instrumentation process ensures intensive use of the collected data (see, e.g. Bertogg & Szydlik, 2016; Gauthier & Gianettoni, 2013; Keller, 2013 for the use of items and scales proposed on the basis of calls for instruments).

The selection of instruments has been made by the project's interdisciplinary scientific team in consultation with the TREE Advisory Board. TREE has prioritised instrument proposals that promise to foster international comparative research and given preference to validated, widely used instruments with a good theoretical foundation and a concrete analytical design. Instrument selection was further based on interdisciplinary relevance, feasibility of implementation and academic cost–benefit analysis. The response to these calls was extremely gratifying. Experienced researchers from a variety of disciplines have submitted a large number of proposals, and 10 to 20% of the instruments used in the ninth TREE1 survey wave (2014) have been improved or added in the process. On the basis of this encouraging experience, TREE will continue to launch calls for instruments in preparation for further survey waves.

Data use

The new 2016 data release of TREE1 (2000–2014) is available free of charge for researchers and students on the Swiss Centre of Expertise in the Social Sciences FORS website (www.forscenter.ch). It includes the PISA 2000 dataset (baseline survey), nine wave-specific datasets (2001–2014) as well as new datasets that cover several waves (for instance, month-by-month job episodes or educational certificates). Extensive data documentation is available in three languages: English, French and German (see TREE 2016b for more detail). The 2016 data release thus offers promising new opportunities for highly detailed analysis of early adulthood transitions (up to age 29) and on the interplay between education and professional careers (or other life domains such as health, family situation and significant life events). In addition, international comparative research could be further developed to study the impact of economic, institutional or social policy as well as other contexts on individual life courses (see, e.g. Buchmann, Kriesi, Koomen, Imdorf & Basler, 2016; Imdorf, Helbling & Inui, forthcoming; Imdorf & Hupka-Brunner, 2015; Kamanzi, Guégnard, Imdorf, Koomen & Murdoch, 2014; Murdoch, Guégnard, Griga, Koomen & Imdorf, 2016; Murdoch et al., 2017; Picot & Hou, 2013). To date, around 270 researchers and students from national and international institutions across a multitude of disciplines have been working with the TREE1 data. Their analyses have led to over 200 publications that address a wide range of issues such as educational inequality (Buchmann et al., 2016), gendered education and labour market pathways (Imdorf & Hupka-Brunner, 2015), the particular role that VET plays for labour market entry in Switzerland (Müller & Schweri, 2015; Helbling & Sacchi, 2014), the impact of discontinuous pathways on later education and labour market outcomes (Sacchi & Meyer, 2016) and the scarring effect of precarious employment at labour market entry on occupational careers of young workers (Buchs & Helbling, 2016; Helbling & Sacchi, 2014; Helbling, 2017).[xii]

Conclusions

The latest TREE1 data release, freely available online (www.forscenter.ch), offers extensive analysis opportunities for life course research on a large Swiss cohort sample over an observation span of roughly 15 years (ages 15–29). TREE employs a longitudinal and comprehensive approach at the intersection of various disciplines, thereby providing extensive opportunities to analyse life transitions in all their complexity at the individual, organisational/institutional and societal levels as well as their interaction. In summary, the following particular features of the TREE dataset enhance its analytic power for the scientific community:

- Given the ever-increasing duration of formal education in knowledge-based societies, TREE's long observation span ensures that *all* school-to-work transitions are covered, including those of students in extended study programmes at the tertiary level.

- Linking the TREE1 panel to the Swiss PISA 2000 sample not only provides the study with an (abundantly contextualised) standardised basic skills measure at its baseline. It also allows for cross-national comparative studies with datasets of other countries that have done the same (for examples of comparative analyses, see Picot & Hou, 2013; Murdoch et al. 2016).[xiii]
- Tight intervals of observation, particularly during the first seven years (yearly panels).
- The detailed, month-by-month episodic data on education and employment renders the data particularly suitable for thorough analysis of school-to-work transition, labour market entry and medium- to long-term career development (for examples for this kind of analysis, see Imdorf & Hupka-Brunner, 2015; Müller & Schweri, 2015).
- The availability of abundant context data allow for complex modelling of factors that influence the observed pathways and transitions.
- Broad multi-disciplinary instrumentation allows for analysis of a particularly wide range of research questions and for interdisciplinary analyses.
- Individual data on youth passing through the highly decentralised, VET-orientated Swiss education system with its high variability of institutional arrangements for school-to-work transitions provide a kind of "life-sized" laboratory for comparative research at both the (sub)national and cross-national levels.

From a critical perspective, it should not be withheld that the TREE data are also subject to some limitations. One of the most substantial among them is that cohort observation starts at a later than ideal time when taking into account some of the crucial features of the Swiss education system. Given the early and pronounced tracking of the country's lower secondary education, cohort observation should begin earlier, ideally before the transition from primary to lower secondary education (i.e. at approximately 12 years of age at the latest). Starting the panel survey as late as the ninth grade means that for a large part of the sample crucial educational decisions in the early phase of respondents' educational careers have already taken place and can only be assessed retrospectively. We have seriously attempted to remedy this limitation by advancing the launch of the second cohort (TREE2) to an earlier age. However, we have not been granted the funds to do so, unfortunately. This flaw in TREE's study design is somewhat mitigated by the fact that in recent years several cohort studies investigating earlier phases of Swiss educational careers have been launched.[xiv]

A second major limitation is that the available data are basically restricted to standardised self-reported student data. Data from other sources (such as teachers or parents) as well as qualitative data or repeated skills measures are not available.

As mentioned earlier, another survey panel for the first TREE cohort TREE1 is planned for 2019. By then, respondents will be 34 years of age on average. Family formation will have initiated for a large part of the sample, and we intend to integrate the respondents' children into the survey, thus extending the scope of the study to the next generation. We expect data from this panel survey to be available by early 2021. After the launch of the second cohort in 2016, TREE will soon provide the first opportunities for cross-cohort research. We expect to release the first waves (baseline and first follow-up panel) of TREE2 data in early 2019.

References

Adams, R. & Wu, M. (2002). *PISA 2000 Technical Report*. OECD.

Arrow, K. J. (1994). The Theory of Discrimination. In D. B. Grusky (ed.), *Social Stratification* (pp. 604–606). Boulder: Westview Press.

Baltes, P. (1990). Entwicklungspsychologie der Lebensspanne: Theoretische Leitsätze. *Psychologische Rundschau*, (41), 1–24.

Baumert, J. & Schümer, G. (2002). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb im nationalen Vergleich. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann & M. Weiß (eds.), *PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich* (pp. 159–202). Opladen: Leske + Budrich.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education.* New York and London: Columbia University Press.

Becker, G. S. (1957/1971). *The Economics of Discrimination*. Chicago: University of Chicago Press.

Bergman, M., Hupka-Brunner, S., Keller, A., Meyer, T. & Stalder, B. E. (eds.). (2011). *Youth Transitions in Switzerland: Results from the TREE Panel Study*. Zurich: Seismo.

Bertogg, A. & Szydlik, M. (2016). The Closeness of Young Adults' Relationships with Their Parents. *Swiss Journal of Sociology,* 42(1), 41–59. https://doi.org/10.1515/sjs-2016-0003

Blossfeld, H.-P. & Shavit, Y. (1993). Persisting Barriers. Changes in Educational Opportunities in Thirteen Countries. In Y. Shavit & H.-P. Blossfeld (eds.), *Persistent Inequality: Changing Educational Attainment in Thirteen Countries* (pp. 1–23). San Francisco: Westview-Press.

Blossfeld, H.-P. & von Maurice, J. (2011). Education as a Lifelong Process. *Zeitschrift für Erziehungswissenschaft*, 14(2), 19–32. https://doi.org/10.1007/s11618-011-0179-2

Boudon, R. (1974). *Education, Opportunity, and Social inequality: Changing Prospects in Western Society*. New York: John Wiley.

Bourdieu, P. (1977). Cultural Reproduction and Social Reproduction. In J. Karabel (ed.), *Power and Ideology in Education* (pp. 487–511). New York: Oxford University Press.

Bourdieu, P. (1984). *Distinction: A Social Critique of the Judgment of Taste.* Cambridge, MA: Harvard University Press.

Bowles, S. & Gintis, H. (1976). *Schooling in Capitalist America. Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.

Breen, R. & Goldthorpe, J. H. (1997). Explaining Educational Differentials: Towards a Formal Rational Action Theory. *Rationality and Society*, 9(3), 275–305. https://doi.org/10.1177/104346397009003002

Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design.* Cambridge, MA: Harvard University Press.

Buchmann, M., Kriesi, I., Koomen, M., Imdorf, C. & Basler, A. (2016). Differentiation in Secondary Education and Inequality in Educational Opportunities: The Case of Switzerland. In H.-P. Blossfeld, S. Buchholz, J. Skopek, & M. Triventi (eds.), *Models of Secondary Education and Social Inequality – An International Comparison* (pp. 111-128). Cheltenham: Edward Elgar. https://doi.org/10.4337/9781785367267.00016

Buchs, H. and Helbling, L.A. (2016). Job Opportunities and School-to-Work Transitions in Occupational Labour Markets. Are Occupational Change and Unskilled Employment after Vocational Education Interrelated? *Empirical Research in Vocational Education and Training*, 8(17). https://doi.org/10.1186/s40461-016-0044-x

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94, S95–S120. https://doi.org/10.1086/228943

Coleman, J. S. (1990). *Foundations of Social Theory*. Cambridge, MA and London: The Belknap Press of Harvard University Press.

Draelants, H. & Ballatore, M. (2014). Capital culturel et reproduction scolaire. Un bilan critique. *Revue française de pédagogie. Recherches en éducation*, (186), 115–142. https://doi.org/10.4000/rfp.4430

Eccles, J. S., Vida, M. N. & Barber, B. (2004). The Relation of Early Adolescents' College Plans and Both Academic Ability and Task-Value Beliefs to Subsequent College Enrollment. *Journal of Early Adolescence,* 24(1), 63–77. https://doi.org/10.1177/0272431603260919

Elder, G. H. J. (1994). Time, Human Agency, and Social Change: Perspectives on the Life Course. *Social Psychology Quarterly,* 57(1), 4–15. https://doi.org/10.2307/2786971

Fend, H., Berger, F. & Grob, U. (eds.). (2009). *Lebensverläufe, Lebensbewältigung, Lebensglück. Ergebnisseder Life-Studie.* Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91547-0

Filipp, S.-H. (ed.). (1995). *Kritische Lebensereignisse*. Weinheim: Psychologie Verlags Union.

Fischer, L. (ed.). (2006). *Arbeitszufriedenheit. Konzepte und empirische Befunde*. Göttingen: Hogrefe.

Gauthier, J.-A. & Gianettoni, L. (2013). Socialisation séquentielle et identité de genre liées à la transition de la formation professionnelle à l'emploi. *Swiss Journal of Sociology*, 39(1), 33–55.

Greve, W. (2000). *Psychologie des Selbst*. Weinheim: Psychologie Verlags Union.

Hascher, T. (2004). *Wohlbefinden in der Schule*. Münster: Waxman.

Heinz, W. R. (1984). *Der Übergang von der Schule in den Beruf als Selbstsozialisation*. Bremen: Universität.

Helbling, L. A. (2017), Fixed-Term Jobs after Vocational Education and Training in Switzerland: Stepping Stone or Impediment? *Swiss Journal of Sociology*, 43(1), 89–113. https://doi.org/10.1515/sjs-2017-0005

Helbling, L. A. & Sacchi, S. (2014). Scarring Effects of Early Unemployment among Young Workers with Vocational Credentials in Switzerland. *Empirical Research in Vocational Education and Training*, 6(12), 1–22. https://doi.org/10.1186/s40461-014-0012-2

Imdorf, C., Helbling, L. A. & Inui, A. (forthcoming). Transition Systems and Non-Standard Employment in Early Career: Comparing Japan and Switzerland. *Journal of Education and Work*, 1–15.

Imdorf, C. & Hupka-Brunner, S. (2015). Gender Differences at Labor Market Entry in Switzerland. In H.-P. Blossfeld et al. (eds.), *Gender, Education and Employment. An International Comparison of School-to-Work Transitions* (pp. 267–286). Cheltenham/Northampton: Edward Elgar. https://doi.org/10.4337/9781784715038.00022

Kamanzi, P. C., Guégnard, C., Imdorf, C., Koomen, M. & Murdoch, J. (2014). Démocratisation de l'enseignement supérieur et emprise des parcours : Comparaison entre la France, la Suisse et le Canada. *Télescope*, 20(2), p. 170–188.

Keller, A. C. (2013). *Development of Job Resources and Personal Resources among Job Newcomers*. Inaugural dissertation submitted to the University of Bern. Bern: Eigenverlag.

Levy, R., Ghisletta, P., Le Goff, J.-M., Spini, D. & Widmer, E. (eds.). (2005). *Towards an Interdisciplinary Perspective on the Life Course*. Amsterdam & Boston: Elsevier.

Müller, B. & Schweri, J. (2015). How Specific Is Apprenticeship Training? Evidence from Inter-Firm and Occupational Mobility after Graduation. *Oxford Economic Papers*, 67(4), 1057–1077. https://doi.org/10.1093/oep/gpv040

Murdoch, J., Guégnard, C., Griga, D., Koomen, M. & Imdorf, C. (2016). How Do Second-Generation Immigrant Students Access Higher Education? The Importance of Vocational Routes to Higher Education in Switzerland, France, and Germany. *Swiss Journal of Sociology*, 42(2), 245–263.

Murdoch, J., Guégnard, C., Koomen, M., Imdorf, C. & Hupka-Brunner, S. (2014). Pathways to Higher Education in France and Switzerland: Do Vocational Tracks Facilitate Access to Higher Education for Immigrant Students? In G. Goastellec & F. Picard (eds.), *Higher Education in Societies. A Multi Scale Perspective* (pp. 149–169). Rotterdam, Boston & Taipei: Sense Publishers. https://doi.org/10.1007/978-94-6209-746-9_10

Murdoch, J., Guégnard, C., Koomen, M., Imdorf, C., Kamanzi, C. & Meyer, T. (2017). Pathways Fostering Mobility to Higher Education for Vulnerable Immigrants in France, Switzerland and Canada. *European Journal of Higher Education*, 7(1), 29–42. https://doi.org/10.1080/21568235.2017.1254918

OECD (1999). *Thematic Review of the Transition from Initial Education to Working Life. Switzerland. Country Note.* Paris: OECD.

OECD (2000). *From Initial Education to Working Life. Making Transitions Work*. Paris: OECD.

Picot, G. & Hou, F. (2013). Why Immigrant Background Matters for University Participation: A Comparison of Switzerland and Canada. *International Migration Review*, 47(3), 612–642. https://doi.org/10.1111/imre.12038

Ramseier, E. & Brühwiler, C. (2003). Herkunft, Leistung und Bildungschancen im gegliederten Bildungssystem: Vertiefte PISA-Analyse unter Einbezug der kognitiven Grundfähigkeiten. *Schweizerische Zeitschrift für Bildungswissenschaften,* 25(1), 23–58.

Renaud, A. (2002). Plans d'échantillonnage pour PISA 2000 en Suisse. Rapport de méthodes PISA. Neuchâtel: Office fédéral de la statistique.

Ryan, P. (2001). The School-to-Work Transition: A Cross-National Perspective. *Journal of Economic Literature*, 39(1), 34–92. https://doi.org/10.1257/jel.39.1.34

Sacchi, S. (2011). *Construction of TREE Panel Weights. Documentation for the Panel Waves from 2000 to 2010*. Basel and Zurich: TREE & cue sozialforschung.

Sacchi, S. & Meyer, T. (2016). Übergangslösungen beim Eintritt in die Schweizer Berufsbildung: Brückenschlag oder Sackgasse? *Swiss Journal of Sociology*, 42(1), 9–39.Scharenberg, K., Hupka-Brunner, S., Meyer, T. & Bergman, M. (eds.). (2016). *Transitions in Youth and Young Adulthood: Results from the Swiss TREE Panel Study. Volume 2*. Zurich: Seismo

Scharenberg, K., Rudin, M., Müller, B., Meyer, T. & Hupka-Brunner, S. (2014). *Education Pathways from Compulsory School to Young Adulthood: The First Ten Years. Results of the Swiss Panel Survey TREE, Part I*. Basel: TREE.

Semmer, N. (1997). Stress. In H. Luczak & W. Volper (eds.), *Handbuch Arbeitswissenschaft*. Stuttgart: Schäfer-Pöschel Verlag.

Spence, M. (1973). Job Market Signalling. *Quarterly Journal of Economics*, 87, 355–379. https://doi.org/10.2307/1882010

TREE (ed.) (2016a). *Concepts and Scales. Survey Waves 1 to 9, 2001–2014*. Bern: TREE.

TREE (ed.) (2016b). *Documentation on the first TREE cohort (TREE1), 2000–2016*. Bern: TREE.

Ulich, K. (1991). Schulische Sozialisation. In K. Hurrelmann & D. Ulich (eds.), *Neues Handbuch der Sozialisationsforschung* (pp. 378–396). Weinheim: Beltz.

Willis, P. (1977). *Learning to Labor. How Working Class Kids Get Working Class Jobs*. New York: Columbia University Press.

Winkle-Wagner, R. (2010). *Cultural Capital: The Promises and Pitfalls in Education Research* (ASHE Higher Education Report No. 36 (1)). Hoboken: ASHE.

Endnotes

[i] Online at http://www.tree.unibe.ch/index_eng.html.

[ii] Past and/or contemporary (multi-)cohort survey projects exist in some other countries, e.g., NLSY in the USA, NEPS in Germany, the British birth cohort studies and "Growing up" studies in Australia, Ireland, Scotland.

[iii] Other PISA follow-up studies started later in Australia (PISA 2003, 2006, 2009), the Czech Republic (PISA 2003), Denmark (PISA 2003), Uruguay (PISA 2003), the USA (PISA 2012) and Hong Kong (PISA 2012).

[iv] In Italian-speaking Switzerland school leavers from private schools were also included in the sample.

[v] Online at http://www.edk.ch/dyn/12928.php.

[vi] See the factsheet titled "Die Überprüfung der Erreichung der Grundkompetenzen (ÜGK) in der Schweiz. Haupterhebung 2016 Mathematik" at

http://grundkompetenzenmathematik.educa.ch/sites/default/files/20151104/uegk_2016_mathematik_factsheet_0.pdf.

[vii] Computer-assisted telephone interviewing (CATI) would have been the method of choice. However, to insufficient funding in the early stage of the project forced TREE to switch to standardized written questionnaires as the less costly means of surveying. For TREE2, CATI was used in the first-wave survey, so we expect to obtain higher participation rates.

[viii] To further minimise attrition, a short form of the telephone interview (containing only few key questions on respondents' education and employment situation) was offered to those unwilling to reply to the long form.

[ix] The duration of CATI participation varies depending on individual situation and on the number of education, employment and other types of activities pursued since the previous panel wave.

[x] See Sacchi (2011) and TREE (2016b) for further information.

[xi] For an example (in German), see

http://www.tree.unibe.ch/ergebnisse/e305140/e305154/files307972/TREE_2014_Newsletter_8_de_ger.pdf.

The hand-written inserts are authentic comments and remarks from the respondents.

[xii] For a complete list of publications, see www.tree.unibe.ch/results/publications/index_eng.html.

[xiii] See endnote iii.

[xiv] E.g., CoCon ("Competence & Context", see http://www.cocon.uzh.ch); Zurich longitudinal skills assessment ("Zürcher Lernstandserhebung", see

http://www.bi.zh.ch/internet/bildungsdirektion/de/unsere_direktion/bildungsplanung/arbeitenundprojekte/lernstand.html).

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 crandall@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 crandall@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

STUDY PROFILE

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