Longitudinal and Life Course Studies:
International Journal

Special section: Transition to young adulthood

Inside this issue

LLCS EDITORIAL BOARD

SUBSCRIPTIONS

INTRODUCTION

SPECIAL SECTION: Transition to young adulthood
Editors Marlis Buchmann, Tina Malti, Heike Solga

PAPERS

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

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)

CALL FOR POSTERS, PAPERS, SYMPOSIA & WORKSHOPS
Deadline for Submissions: Monday 6 March 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, we welcome conference submissions 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).

Proposals are sought for four kinds of conference presentation:
1. A poster presentation for which an abstract of no more than 200 words is required.
2. An individual paper for oral presentation for which an abstract of no more than 300 words is required.*
3. A symposium comprising of 3 – 5 papers. Accepted symposia will be allocated a time slot of between 1 – 1.5 hrs. For each symposium we require a summary abstract of no more than 200 words, plus an abstract of no more than 300 words for each paper.
4. Post-conference workshop. Duration: either half or full day. Half day workshops can take place on the afternoon of Friday 13 October and on Saturday 14 October 2017. Full day workshops can **only** take place on Saturday 14 October 2017. An abstract of no more than 300 words is required.

*Papers which do not fit into regular sessions may be allocated as posters.

ABSTRACT SUBMISSION: Visit http://www.slls.org.uk/untitled-c11p2 to submit your abstract.

All contributors will be notified of the conference committee's decision by **Friday 31 March 2017.**

More information about the conference and recommended accommodation in Stirling will be posted on the Society 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 Journal will be a possible outlet for the publication of conference papers.

We look forward to receiving your submissions!

Editorial: A critical life course stage

John Bynner　Executive Editor

This issue of the journal displays the variety and richness of the submissions the publication receives in the field of longitudinal and life course study.

We start with a *special section* on transitions in the third decade of life – the 20s – guest edited by Marlis Buchman (Zurich), Heiki Solga (Berlin) and Tina Malti (Toronto). The section comprises four papers introduced by Marlis Buchmann and Heike Solga, with authors spanning Canada, USA, Finland, Germany, UK, Switzerland and China. We then move to two *individual papers*. The first is on changes in early transition from school to work in England. The second is a methodological paper on the biasing effects brought about by selection processes in sample attrition in Swedish longitudinal study.

The twin effects of secular change and heterogeneity of experience from age 20 are central to this stage of the life course, as reflected in the developmental transitions that take place. Many, such as from education to work, are completed during this period, if not before, while others such as partnership/marriage, have yet to occur or are just beginning. The key point is that the transition's impact on the direction the life course takes will be felt across the whole of the life span.

The biological processes in maturation set the physical parameters for young adulthood and the stages that follow of middle life through the 30s, 40s and increasingly as life extends, the 50s. They may also signal the beginnings of closure with the winding down of physical prowess and the onset of old age.

I was fascinated in reading a few days ago the ages of the players in UK Premier Division football clubs. The average age for every single one of them was 27 or 28 and those top clubs with 28-year-olds were already identifying them for potential transfer to other less prestigious clubs and then retirement. Seen from this perspective the huge salaries top players earn are less surprising – part of a life course with a highly restricted window of opportunity in the 20s through which to build a secure future or progress potentially to nothing.

But biology is only one part of the story. In the current era, we are witnessing transformation of institutions and experience in such life course domains as employment, with impact on all the others – ranging through partnership and marriage, housing, social life, political engagement, crime, mental and physical health and, of course, massive movement of people.

This is while technological development transforms society at an ever accelerating rate and automation from artificial intelligence to robotics takes over what was once defined as human work. The misconceived idea (in my view) of the 'standardised' life course, not to mention the stratification that goes with it, takes even more of a battering. The occupational and labour market certainties of the past give way to the casualisation of labour, the move to part-time contracts and, most damaging to personal security of all, the zero-hours contract. These employment forms have wider implications reflected by a new sub-stratum cutting across the skills-based socioeconomic statuses of the past – the 'precariat'. For these 20-30 year olds the collective organisation and protection supplied in the past by, for example, trade unions are largely non-existent.

Unlike the experience of past centuries, as the international range of our contributors affirms, these processes are operating on an ever-more global scale, universalising labour market experience and driving ever widening inequality. They are also accompanied by diminishing effectiveness of national governments to control it.

And it is not only employment as we knew it that is disappearing. There is a mediating effect of labour market changes on other life course domains such as family formation and especially housing – where the prospects of security through home ownership, without substantial parental subsidy, are rapidly disappearing. The golden age of affluence and autonomy has been replaced since the banking collapse of 2007/8 by that of austerity and extended dependency, making the rich very much richer and the poor relatively very much poorer.

We hear from Oxfam that the 62 people whose collective wealth, in January 2016, exceeded that of half the world's population has now reduced to 8.

Yet at the same time the ingenuity of such ideas as the 'Citizen Wage' and the passion of the movements to promote it suggests that, as ever, human striving for individual and collective survival will prevail.

The papers in this edition's special section pay particular regard to these challenging aspects of young adulthood in the contemporary world. The role of psycho-social resources is the centre of attention coupled with other resources available to the new generation in making their transitions in particular domains. The focus is first on the role of the moral antecedents of young adults' attitudes to inequality, then on life satisfaction, education and occupation, then socioeconomic resources and individual agency. Finally, a first for the journal completes the picture – an exploration of the consequences of exposure to the criminal justice system for health.

The individual papers in this issue extend the story, focusing first on historical change: cohort shifts in relation to the key life course transition from school to work. Finally, we learn that our ability to understand these effects goes only as far as the representativeness of our sample data and the robustness of our analyses can take us, because of the biasing effects of different types of 'selective' sample loss. *Attrition* – described as the 'Achilles heel' of longitudinal research – takes on even more significance at a time of massive social, economic and technological change.

Congratulations to the authors in drawing our attention to these challenges to the conventional wisdom about the world as it used to be and what it is likely to become. There was never a time when longitudinal and life course study and platforms communicating its findings, such as LLCS, were more needed.

GUEST EDITORIAL

Challenges of the third decade of life: The significance of social and psychological resources

Marlis Buchmann   University of Zurich, Switzerland
Heike Solga   Berlin Social Science Center and Freie Universitaet Berlin, Germany

Status transitions to adulthood have been delayed well into the third decade of life in advanced Western societies over recent years. The third decade has thus mutated into a life period in which young people face more key status transitions and life decisions than ever before. It is therefore of particular interest to learn more about how young people cope with these social and developmental transitions. While psychological and social resources can be assumed to play a key role, research has paid little attention to differences in the social, psychological, and physical resources with which young people start the (challenging) third decade of life. Hence, the question of how resource inequalities experienced in the previous two decades of life are related to young people's opportunities, constraints, and agentic capacities for adopting adult roles has barely been addressed.

This special section aims to partially fill this gap and to provide much needed insight into the antecedents and consequences of inequalities in various kinds of resources among 20-somethings of different social groups and across various societal contexts. The resources targeted in the four papers are physical, psychological, and social. The authors assess how they help to successfully deal with the extended period of transitions into adulthood in the third decade of life and how they affect the development of adulthood commitments and wellbeing. The consequences of young adults' exposure to critical contexts for resource acquisition in the transition to adulthood are also scrutinised. It is highly commendable that all contributions assembled in this Special Section address their particular research questions from a life course perspective. Moreover, they consider how contextual characteristics affect the role of resources and base the analyses on superb longitudinal data.

The collection of the four papers assembled in this special section makes a unique contribution to our understanding of how unequal starting conditions in the third decade of life, related to variations in both resources and contexts, affect health, wellbeing, and early adulthood commitments. Across the four studies, an astonishingly broad range of social and psychological resources has been taken into account, demonstrating the ways they are intertwined with the social and developmental tasks to be accomplished in early adulthood. Malti and colleagues show for Switzerland, a country known for its well-developed direct democratic political system, how the social resource of friendship quality developed in the second decade of life is associated with the internalised acceptance of the principle of equality in early adult life, considered to be a core principle in democratic societies. Positive experiences in close friendships are likely to provide opportunities to connect with the needs and perspectives of others, thus stimulating concern for social equality. This paper also highlights the importance of moral resources accrued in the second decade of life (sympathy, in particular) for instilling sensitivity to social inequality and related values of social justice. The internalisation of these beliefs and values is widely recognised as an important developmental milestone in young adulthood.

How life satisfaction acts as an important personal resource for coping with challenges in the transition to young adulthood is aptly demonstrated in Upadyaya and Salmela-Aro's study. The authors examine the intricate

longitudinal interplay between life satisfaction and study/work engagement over the transition from post-comprehensive education to higher education or work, following young Finns from the age of 17 to 25. While life satisfaction, indicating general wellbeing, predicts young people's study/work engagement across the entire time period under consideration, study/work engagement predicts life satisfaction only *after* the transition to higher education or work at the beginning of the third decade. This finding squarely reminds us that the significance of psychological resources greatly depends on the social context in which they play out, in this case, the structural context of the life course and life course transitions. Along this line of reasoning, the developmental changes in the ways domain-specific engagement (school/work) and general wellbeing relate to each other might be attributed to changes in the stage-environment fit, which is assumed to be higher after the transition to higher education or work.

In the contribution by Schoon and Lyons-Amos, general life satisfaction is the outcome of interest at the brink of the third decade of life. The research focuses on how structural and agentic resources and the ways in which young people manage education and employment transitions between the ages of 16 and 20 years in England affect their general wellbeing. Like the aforementioned Finnish study, Schoon and Lyon-Amos also show how dimensions of agency, such as goal certainty and school engagement, predict young people's life satisfaction at the beginning of the third decade. This is remarkable given the institutional differences in the educational system between Finland and England. The reported effects are upheld even when controlling for the type of education and employment transitions, identified by sequence and cluster analyses. By contrast, transition experiences fully mediate the significance of social resources (measured by a familial cumulative socioeconomic risk index) on young people's life satisfaction at the brink of the third

decade of life. Taken together, these findings underscore the independent influence of young people's agentic capacities on how well they are satisfied with life.

A particularly incisive, non-normative life transition is the experience of incarceration. This might be especially the case when this experience occurs in late adolescence or early adulthood – a time that would normally be devoted to establishing an independent life. Because inmates are kept in a 'total institution,' deprived of conventional social contacts and interactions, and confronted with limited education and work opportunities, this experience is prone to have negative (long-term) consequences on a broad range of life dimensions. Esposito and co-authors have addressed this issue for the United States, a country known for currently having the highest incarceration rate in the world. They focus on health consequences of incarceration in young adulthood as this stage in the life course is a particularly critical period for developing poor health. The great merit of this study is that it successfully parses out the influence the experience of incarceration has on health from the social and economic confounders that are linked to both the criminal justice contact and health. The findings document that young adults' incarceration is an important stressor with immediate consequence for mental and general health.

In sum, these four papers address, from different angles and perspectives, how unequal social resources and individual agency accrued over the first two decades of life, and the exposure to social contexts characterised by disparate opportunities and constraints, shape life course outcomes in early adulthood. Using high quality longitudinal data and applying state-of-the-art analytical strategies, they document the intricate longitudinal interplay between opportunities and agency in the particularly sensitive period of the transition to the third decade of life.

Moral and social antecedents of young adults' attitudes toward social inequality and social justice values

Tina Malti   University of Toronto, Canada
tina.malti@utoronto.ca
Sebastian P. Dys   University of Toronto, Canada
Lixian Cui   New York University Shanghai, China
Marlis Buchmann   University of Zurich, Switzerland

Abstract

In light of growing social stratification, there have been calls to better understand the developmental antecedents of attitudes and values related to social inequality. In this study we predicted attitudes toward social inequality and social justice values from moral and social antecedents in a representative sample of Swiss adolescents (N = 1,258) at 15 (Time 1), 18 (Time 2), and 21 years of age (Time 3). We assessed children's sympathy and morals in the context of individuals' decision-making and anticipation of emotions in moral dilemmas. Social-contextual factors included relationship quality, which was assessed by the quality of one's closest friendship and education level. Adolescents who reported higher friendship quality and sympathy showed stronger attitudes toward social inequality later. Interestingly, adolescents' own education level at age 18 positively predicted attitudes toward social inequality at age 21 above and beyond parent education level, but only marginally at a younger age. Social justice values at age 18 were predicted by sympathy and the anticipation of moral emotions at age 15, and social justice values at age 21 were associated with sympathy at age 18. Results are discussed with respect to the potential significance of morality and social-contextual factors in the development of attitudes toward social inequality and social justice values in early adulthood.

Keywords

Introduction

In light of increasing social and economic stratification across the globe, disciplines across the social sciences have become more interested than ever in issues related to social inequality (Breen & Jonsson, 2005; Gordon & Dew-Becker, 2008; Norton & Ariely, 2011). Though much of this work has examined the impact of social inequalities on outcomes such as educational attainment, income, and health, far less has sought to understand the developmental antecedents of attitudes and values related to these social issues (Killen & Smetana, 2015; see Daniel, Dys, Buchmann, & Malti, 2016). Here, we address this gap, in part, by examining attitudes toward inequality, i.e. the degree to which principles of fairness and equality are important to an individual (Rubin, Malti, & McDonald, 2012) and social justice values, i.e., how much one values justice in the social treatment of individuals (Killen & Smetana, 2010; Marini, 2000). Attitudes toward inequality and social justice values have been posited as a central outcome of productive development in the third decade of life and an important developmental milestone in young adulthood (Larson, Brown, & Mortimer, 2002; Lerner, 2004). Both attitudes toward inequality and social justice values are presumed to reflect dimensions of civic competence and factor into the development of social engagement (Lenzi et al., 2014; Rubin, Malti, & McDonald, 2012; see Brewer, 2003), which may be particularly important for motivating awareness and change around issues of social inequality.

The role of moral and social concepts, such as moral emotions and the quality of close relationships, may yield valuable insight into the formation and development of attitudes toward inequality and social justice values (Flanagan & Christens, 2011). Yet, few longitudinal studies on the moral and social antecedents of attitudes toward inequality and social justice values have been conducted thus far. The present study addressed this research gap by investigating the moral and social antecedents of attitudes toward inequality and social justice values (for a simplified illustration of the conceptual relations, see figure 1). Here, we highlight important psychological dimension of human functioning: morality. We emphasise the role of young people's morality as it may form the basis for other-oriented concern and as such, may motivate a broadened care for all members of society. Here, we examine both the role of sympathy and the anticipation of self-evaluative moral emotions (e.g. guilt, sadness following wrongdoing; see Malti, 2016) as antecedents of attitudes toward inequality and social justice values. Sympathy, an other-oriented moral emotion, is believed to play a central role in motivating individuals to work toward meeting the needs of others (Eisenberg, 2000). This is because sympathy stems from the apprehension of another's emotional state and is expressed in other-oriented concern, which heightens an individual's attention to the needs of others (Eisenberg, 2000; see Piaget, 1932/1997). Consistent with this notion, there is cross-sectional, longitudinal, and qualitative evidence demonstrating that sympathy is associated with social justice values in adults (Daniel, Dys, Buchmann, & Malti, 2014; Daniel et al., 2016; Malin, Tirri, & Liauw, 2015). The anticipation of moral emotions, such as guilt and sadness, in response to hypothetical transgressions reflects the internalisation of moral norms (Malti & Ongley, 2014). Previous research indicates that individuals often attribute self-evaluative emotions, such as guilt and related feelings of sadness, in such contexts. It has been shown that individuals anticipate such self-evaluative emotions due to the perceived acceptance of internalised norms of fairness and justice and an awareness that one has violated one's own internalised moral principles (Malti, 2016; Malti, Gummerum, Keller, & Buchmann, 2009; Malti, Keller, & Buchmann, 2013).

The anticipation of other-oriented and self-evaluative moral emotions, such as sympathy, as well as guilt and sadness following own wrongdoing, has also been empirically linked to prosocial, other-oriented, and fair behaviour (Eisenberg, 2000; Malti & Krettenauer, 2013). As such, we anticipated that the presence of such emotions, which reflect an internalisation of moral norms and are associated with fair behaviour, may longitudinally predict attitudes toward inequality and social justice values.

In addition, our conceptual model suggests that adolescents' moral and social-emotional development is situated within a broader social space. We propose that an individual's cultural capital and experiences in social, cultural, and economic domains may influence

their attitudes and values through community involvement (see Verba, Burns, & Schlozman, 2003). For instance, in the social domain, close relationships act as important basis for social values and can serve as a buffer for risks associated with disengagement and individual maladaptation (Hinde, 1979). The formation of healthy social relationships may also promote the generalisation of other-oriented concerns to broader values. For instance, adolescents who form healthy, secure relations with peers may be more motivated to consider the wellbeing of others (Rubin, Bukowski, & Laursen, 2009). As broader social concerns become more self-relevant, these experiences may in turn motivate an interest in and commitment to issues of social justice (see Wilkinson, 2010; Wray-Lake & Syversten, 2011; Yates & Youniss, 1998).

Furthermore, the model accounts for social inequalities based on cultural capital, such as education. For example, adolescents' education may act as an impetus for social awareness in adolescence and emerging adulthood (Hillygus, 2005; Nie & Hillygus, 2001). Being aware of the need to act upon principles of fairness, social justice, and equality is likely going to promote the development of social justice values and the motivation to act according to such principles. Meanwhile, parents' education appears to have a multilayered form of influence— children can directly learn from their parents' attitudes and values, as well as receive indirect influences from their education, for instance, by being able to achieve a higher level of education themselves (Plutzer, 2002; Verba, Burns, & Scholzman, 2003) or through exposure to broader social networks (Gesthuizen, van der Meer, & Scheepers, 2008) and civic activities (Bekkers, 2007). Thus, we also examined the effects of adolescents' and parents' levels of education on attitudes toward inequality and social justice values.

Though there has been an increasing interest in the development of attitudes and values related to social inequalities, studies that have investigated both moral and relational factors in their development are sparse. We, therefore, sought to answer the following research questions: First, what is the role of adolescents' moral emotions in the development of attitudes toward inequality and social justice values?

Second, what roles does their relationship quality with close friends play in the development of attitudes toward inequality and social justice values in young adulthood? Based on our conceptual model and the extant evidence, we expected the dimensions of morality to significantly predict subsequent attitudes toward inequality and social justice values. We also hypothesized that the perceived quality of close friendships and parents' and adolescents' levels of education would be associated with later attitudes toward inequality and social justice values. In addition, since adolescent gender has been linked to the development of adolescents' moral emotions and social justice values (e.g., Eisenberg, Zhou, & Koller, 2001), we controlled for adolescent gender in all multivariate analyses.

Method

Participants

Data from the first three waves of the representative ongoing Swiss longitudinal survey COCON (ww.cocon.uzh.ch) were utilised to test our hypotheses. A representative random sample from the German- and French-speaking parts of Switzerland was drawn by a two-stage process in which 131 communities—broken down by community type (i.e., urban, suburban, and rural) and size— were selected. The residents of each community were then randomly sampled on the basis of information provided by the community's official register. Here, we used data from a cohort of adolescents at 15 years of age at Time 1 (T1, N = 1,258, 54% female), 18 years of age at Time 2 (T2, N = 952), and 21 years of age at Time 3 (T3, N = 815). The adolescents' primary caregivers (N = 1,056) also participated at Time 1.

Procedure

At the first assessment, in spring 2006, the adolescents were individually interviewed in a quiet room at their home using a computer-assisted personal interview for about 60 minutes. The interview contained questions on the participant's social and moral development. The mothers responded to a questionnaire regarding their adolescent's social and moral development, which was mailed back to the research institute. Both the participant and the primary caregiver supplied written

informed consent for participation. At the second and third time points (spring 2009 and spring 2012, respectively), a computer-assisted personal interview was conducted with the adolescents and young adults, respectively. The interviewers were recruited from a professional research institute specialising in social sciences, and had been trained extensively by the research team on interview techniques.

Measures

Attitudes toward inequality. From T1-T3, attitudes toward inequality were assessed using seven items on a six-point Likert scale from 1 (*fully disagree*) to 6 (*fully agree*). This measure included items such as "Wealth should be distributed more evenly around the world, even if it requires me to give up some consumer good", "In Switzerland, we spend too much money on welfare, i.e., to support the unemployed and single parents to raise their children" (reverse coded), and "Those who earn well should make a major contribution to the community". The Cronbach's α was .63 at T2 and .65 at T3. At T1, the α coefficient was too low to justify inclusion of the scale into the statistical analyses.

Social justice values. From T1-T3, adolescents' social justice values (Gille, Sardei-Biermann, Gaiser, & de Rijke, 2006) were reported using a scale that consisted of three items taken from the German Youth Survey. The DJI is a representative, large-scale survey; the social justice value scale has shown to be reliable and valid, both in the German Youth Survey (Gille et al., 2006) and in our pilot study. The scale asked how important it is "To interact with others in a fair way," "To treat all humans equally," and "To minimize inequalities between humans." The answers were rated on a 10-point Likert scale, ranging from 1 (*not important at all*) to 10 (*extremely important*). Mean scale scores were computed, with higher scores indicating greater importance of social justice values. Cronbach's α for the value scale was .55 at T1, .70 at T2, and .63 at T3.

Sympathy. Sympathy was assessed by self- and caregiver-reports (Zhou, Valiente, & Eisenberg, 2003). Adolescents reported on their own sympathy from T1-T3. The adolescents' scale consisted of five items from Zhou et al. (2003). A sample item is "When I see another child who is hurt or upset, I feel sorry for him

or her." Items were rated using a six-point scale ranging from "*not at all like me*" to "*very much like me.*" Cronbach's α for the self-reported sympathy scale was .72 at T1, .70 at T2, and .73 at T3.

In addition, at T1, the mothers rated their children's sympathy on three items from Zhou et al. (2003). A sample item is "My child usually feels sorry for other adolescents who are being teased." Items were rated using a six-point scale ranging from "*not at all like my child*" to "*very much like my child*." Cronbach's α for the mother-reported sympathy scale was .78. To create a more robust measure of sympathy for the multivariate analyses, we aggregated adolescent- and caregiver-reports of sympathy, which were significantly correlated, *r* = .22, *p* < .001.

Moral emotions. Two moral dilemmas were chosen for the current study, involving the choice between well-known moral rules and personal benefit, which closely resembled those experienced in the everyday lives of adolescents in their nature and structure. In the first story, adolescents were read the following: "Imagine you offered your bike for sale. You want to sell it for 500 Swiss Francs. A young man is interested. He bargains with you and you agree on 420 Swiss Francs. Then he says: 'Sorry, I don't have the money on me; I'll quickly run home to get it. I'll be back in half an hour.' You say: 'Agreed, I'll wait for you.' Shortly after he is gone, another customer shows up who is willing to pay the full price." In the second story, the adolescents were read: "Imagine that you have found a purse with 150 Swiss Francs in it and an identity card of the owner." After reading each story, the adolescents were asked what they would do in order to identify their moral evaluations in each situation (moral decision-making), and what they feel thereafter (moral emotions).

To create a score for moral emotions, we coded both responses to dilemmas (moral decision-making and moral emotions) together, such that they created an overt composite score for the anticipation of self-evaluative emotions (for a more detailed description of the coding procedure, see Malti et al., 2013). More specifically, decisions and emotions were combined to create the different patterns of decision-making and emotions (see Malti & Keller, 2010). The *happy victimizer* pattern applied to participants who based their decision on selfish reasons and attributed

positive emotions to the self (for selfish reasons). The *unhappy victimizer* pattern applied to participants who based their decision on selfish reasons, but who attributed negative emotions to the self for moral reasons. The *happy moralist* pattern applied to participants who based their decision on moral reasons and attributed positive emotions to the self for moral reasons. The *unhappy moralist* pattern applied to participants who based their decision on moral reasons, but who felt unhappy due to moral reasons (e.g., empathy for the new customer because he/she is not getting the bike). Participants who opted for the moral decision, but who felt regret due to selfish reasons, were very rare; thus, this pattern was not considered further.

Responses in the unhappy victimizer, happy moralist and unhappy moralist category were coded as "moral emotions occurred" (1), whereas responses in the "happy victimizer" category were coded as "moral emotions did not occur" (0). This was done because the anticipation of emotions in these categories implied an awareness and internalisation of moral norms. Next, to generate a composite score for our measure, we averaged the two scores across both moral dilemmas.

Friendship quality. Perceived friendship quality with a best friend was assessed by self-reports (Parker & Asher, 1993) from T1-T3 on a 6-point scale, 1 (*fully disagree*) to 6 (*fully agree*). This measure consisted of four items such as, "My friend and I trust each other when we give advice," "My friend and I help each other," and "My friend and I discuss what is important in life." Cronbach's α for the scale was .61 at T1, .67 at T2, and .73 at T3.

Parent education. Fathers' and mothers' education levels were reported on a seven-point scale ranging from low to high levels of education (see Malti et al., 2013), i.e., 1 (*no education*), 2 (*compulsory school*), 3 (*middle school/school section II*), 4 (*apprenticeship/full time vocational school*), 5 (*higher technical and vocational training*), 6 (*higher technical school/college*), and 7 (*university*). A composite score of parent education was created by aggregating father's and mother's education levels for each family. A high score indicates higher levels of parent education.

Adolescent education. At T1, adolescent's education level was reported on a five-point scale (see Malti & Buchmann, 2010), 1 (*without classes/small classes*), 2 (*section C/junior high school*), 3 (*section B+G/unarticulated*), 4 (*section A+E, district school*), and 5 (*high school/secondary school*). At T2, adolescent's education level was reported on a four-point scale, 1 (*no certifying training*), 2 (*upper secondary education*), 3 (*training with professional baccalaureate*), and 4 (*high school*). A higher score indicates higher levels of education. The scoring system differed at T1 and T2 because at T1, adolescents were in lower-secondary education (i.e. general education), and at T2, they were in upper-secondary education, with the majority of young people in VET (vocational training and education) serving an apprenticeship, while the rest being in general education (specialised middle schools or 'Gymnasium').

Analytical Approach

We first ran descriptive statistics on attitudes toward inequality, social justice values, and all other focal variables, followed by a series of multivariate regression models predicting attitudes toward inequality and social justice values at T2 and T3, respectively using M*plus* 7.2 (Muthén & Muthén, 1998-2012). Adolescent gender was controlled for in all the multivariate analyses.

The retention rates for the current study were 76% from T1 to T2 and 86% from T2 to T3. Missing data analyses of all the models we ran suggested that although the missing patterns were not Missing Completely at Random (MCAR, Little, 1988), the missingness was predicted by the variables included in the model, such that adolescents with lower morality, and education level were more likely to be missing on attitudes toward inequality and social justice values at both T2 and T3. Therefore, Missing at Random (MAR) was assumed and the maximum likelihood with robust error (MLR) was used for parameter estimation in all analyses to account for missing values in M*plus* 7.2 (Muthén & Muthén, 1998-2012).

Results

Descriptive statistics

Table 1 displays the means and standard

deviations of all continuous study variables by assessment point. As shown in table 2, attitudes toward inequality at T2 and T3 were associated with most independent variables at both T1 and T2. Similarly, social justice values at T1-T3 were related to most independent variables in the expected direction but not significantly associated with parent and adolescent education levels.

Prediction of attitudes toward inequality and social justice values by moral emotions and social factors

To test how moral development and social factors predicted attitudes toward inequality and social justice values at both T2 and T3, we ran linear regression analyses using T1 independent variables to predict T2 attitudes toward inequality and social justice values, and using T2 independent variables to predict T3 attitudes toward inequality and social justice values, respectively (see table 3). These two variables were treated separately because they are conceptualised as distinct constructs. However, we predicted them in the same model in a multivariate way. We found that T1 perceived friendship quality, T1 sympathy (self- and parent-reports combined), and parent education level significantly predicted T2 attitudes toward inequality (table 3). We also found that T1 sympathy (self- and parent-reports combined) and T1 moral emotions significantly predicted T2 social justice values, after controlling for T1 social justice values. The correlation between attitudes toward inequality and social justice values at T2 was in the medium range, $r = .41$, $p < .001$. Total variances explained by the independent variables were 11% for attitudes toward inequality and 12% for social justice values.

Next, we predicted attitudes toward inequality and social justice values at T3 using T2 independent variables. The findings showed significant effect of T2 sympathy (self-reports), as well as T2 adolescent education levels and parent education levels in predicting T3 attitudes toward inequality, after controlling for T2 attitudes toward inequality. For the prediction of T3 social justice values, T2 sympathy showed a significant effect, after controlling for T2 social justice values. As expected, attitudes toward inequality and social justice values at T3 were significantly correlated in a small range, $r = .28$, $p < .001$. Total variances explained by the independent

variables were 26% for attitudes toward inequality and 27% for social justice values.

Discussion

One of the major challenges during the transition into adulthood involves engaging in, and committing to, democratic principles of justice and fairness (Flanagan & Gallay, 1995). Our study aimed at investigating central moral-developmental and social antecedents of attitudes toward inequality and social justice values. In particular, we focused on the influence of moral emotions, friendship quality, and socioeconomic status in predicting these attitudes and values.

Our findings suggest that moral emotions, i.e. sympathy, play a central role in the subsequent development of attitudes toward inequality and social justice values. This is in line with previous studies highlighting the role of sympathy in promoting related outcomes, such as prosocial orientations and behaviours (Eisenberg, Spinrad, & Knafo-Noam, 2015). Similarly, the findings support the notion that individuals who feel concern for the distress of victims of injustice are more likely to feel committed to such principles (Hoffman, 2000). Moreover, this is consistent with, and extends previous studies documenting a longitudinal link between sympathy and social justice values trajectories (e.g. Daniel et al., 2014). There was also some support for the role of self-evaluative moral emotions at T1 on subsequent social justice values. This may support the idea that self-evaluative moral emotions may function to promote other-oriented behaviour and values by evoking unpleasant feelings in response to violating one's personal moral code (see Carlo, McGinley, Davis, & Streit, 2012; Ongley & Malti, 2014). In short, it appears that young adults' moral sentiments may promote a sense for inequality and related values of social justice (Montada & Schneider, 1989; see van Goethem et al., 2012).

In line with our developmental-relational model, we also found some support for the role of friendship quality in subsequent attitudes toward inequality, which resonates with theorizing (Lewis, MacGregor, & Putnam, 2013) and US longitudinal research (Obradovic & Masten, 2007). Positive experiences in close friendships may help young

people appreciate the need for social cohesion and interpersonal responsibility, and experience advancing these causes as meaningful and satisfying. Moreover, high quality friendships are likely to be indicative of other social competencies (not directly measured here), which are central to attitudes toward justice and equality. For instance, active listening and perspective-taking are often applied to achieve social understanding and unity (Flanagan & Faison, 2001; Obradovic & Masten, 2007). By providing concrete opportunities to connect with the needs and perspectives of others, close friendships may stimulate adolescents' concern for social equality.

Consistent with previous literature (e.g., Verba et al., 2003), we found parental education to predict attitudes toward inequality. Interestingly, adolescents' education at T2 also predicted attitudes toward inequality at T3, beyond the influence of parents' education. One potential reason is that educational programming in late adolescence may be much better at helping young people understand how society operates, and how they can influence that operation (see Hillygus, 2005). Alternatively, adolescents may not develop extensive social networks of their own until their twenties, resulting in a more pronounced effect of their parents' social networks on adolescents' attitudes toward inequality at an earlier age (see Gesthuizen et al., 2008). Thus, consistent with our model, it appears that early social experiences and differences in educational attainment may colour young peoples' perceptions and expectations of social interactions and society, and in so doing influence their degree of concern for broader social issues.

While the present study addresses an important void in research on attitudes toward inequality and social justice values in adulthood, it is not without its limitations. First, due to the large-scale approach to our data collection, our measure of social justice values was limited to three items and as such

relied on relatively simple general statements. However, this measure has been employed and validated in previous large-scale studies (Gille et al., 2006). Second, many of our measures relied on self-reports, which may be susceptible to social desirability. Still, self-reports are necessary and invaluable for the study of morality and values, as these are typically challenging to measure through other informants or behavioural measures. Moreover, while social desirability biases tend to positively relate to valuing harmony, they are unrelated to universalism values (e.g. social justice values; Schwartz, Verkasalo, Antonovsky, & Sagiv, 1997). Third, the validity of our measures was somewhat limited by the internal consistency of our measures. Although some of our measures (e.g. social justice values, friendship quality) had, at times, relatively low reliability coefficients, they are still considered acceptable as lower scores are typical of scales with few items (Schmitt, 1996). Given that one of our scales, attitudes toward inequality, had low reliability at the first assessment point, we were unable to control for it in the longitudinal prediction of outcomes. Future studies are therefore warranted to replicate finding when controlling for Time 1 attitudes toward inequality.

Despite these limitations, the present study provides a valuable contribution to our understanding of attitudes and social values in the third decade. To conclude, we aimed, in part, to address the need for research connecting social and moral concepts in adolescence to attitudes toward inequality and social justice values in early adulthood (Killen & Smetana, 2010). Both moral and social factors in adolescence were longitudinally predictive of attitudes toward injustice and social justice values. Taken together, these findings support the view that both moral emotions and close relationships are likely to impact young adults' attitudes and social values.

Acknowledgments

References

Bekkers, R. (2007). Intergenerational transmission of volunteering. *Acta Sociologica, 50*(2), 99-114. https://doi.org/10.1177/0001699307077653

Breen, R., & Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology*, *31*, 225-243. https://doi.org/10.1146/annurev.soc.31.041304.122232

Brewer, G.A. (2003). Building social capital: Civic attitudes and behavior of public servants. *Journal of Public Administration Research and Theory, 13*(1), 5-26. https://doi.org/10.1093/jopart/mug011

Carlo, G., McGinley, M., Davis, A., & Streit, C. (2012). Behaving badly or goodly: Is it because I feel guilty, shameful, or sympathetic? Or is it a matter of what I think? *New Directions for Youth Development*, *136,* 75-93. https://doi.org/10.1002/yd.20040

Daniel, E., Dys, S.P., Buchmann, M., & Malti, T. (2014). Developmental relations between sympathy, moral emotion attributions, moral reasoning, and social justice values from childhood to early adolescence*. Journal of Adolescence, 37*(7), 1201-1214*.* https://doi.org/10.1016/j.adolescence.2014.08.009

Daniel, E., Dys, S.P., Buchmann, M., & Malti, T. (2016). Developmental trajectories of social justice values in adolescence: Relations with sympathy and friendship quality. *Social Development, 25,* 548-564*.* https://doi.org/10.1111/sode.12146

Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology, 51*, 665–697. https://doi.org/10.1146/annurev.psych.51.1.665

Eisenberg, N., Spinrad, T.L., & Knafo-Noam, A. (2015). Prosocial development. In M. E. Lamb (Vol. Ed.) and R.M. Lerner (Series Ed.), *Handbook of child psychology and developmental science, Vol. 3: Social, emotional and personality development (7th edition, pp. 610-656).* New York: Wiley. https://doi.org/10.1002/9781118963418.childpsy315

Eisenberg, N., Zhou, Q., & Koller, S. (2001). Brazilian adolescents' prosocial moral judgment and behavior: relations to sympathy, perspective taking, gender-role orientation, and demographic characteristics. *Child Development*, *72,* 518-534. https://doi.org/10.1111/1467-8624.00294

Flanagan, C. A., & Faison, N. (2001). Youth civic development: Implications of research for social policy and programs. *Social Policy Report*, Vol. XV (1). Ann Arbor, MI: Society for Research in Child Development.

Flanagan, C. A. & Gallay, L. S. (1995). Reframing the meaning of "political" in research with adolescents. *Perspectives on Political Science, 24*, 34-41. https://doi.org/10.1080/10457097.1995.9941863

Flanagan, C. A., Levine, P. (2010). Civic engagement and the transition to adulthood. *The Future of Children*, *20*(1), 159-179. https://doi.org/10.1353/foc.0.0043

Gesthuizen, M., van der Meer, T., & Scheepers, P. (2008). Education and dimensions of social capital: Do educational effects differ due to educational expansion and social security expenditure? *European Sociological Review, 24*(5), 617-632. https://doi.org/10.1093/esr/jcn021

Gille, M., Sardei-Biermann, S., Gaiser, W., & de Rijke, J. (2006). *Jugendliche und junge Erwachsene in Deutschland [adolescents and young adults in Germany]*. Wiesbaden, Germany: Verlag fuer Sozialwissenschaften.

Gordon, R. J., & Dew-Becker, I. (2008). Controversies about the rise of American inequality: A survey. NBER Working Papers 13982*, National Bureau of Economic Research*. https://doi.org/10.3386/w13982

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

Hinde, R. A. (1979). *Towards understanding relationships.* London: Academic Press.

Hitlin, S., & Vaisey, S. (2013). The new sociology of morality. *Annual Reviews of Sociology*, *39*, 51-68. https://doi.org/10.1146/annurev-soc-071312-145628

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice.* New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511805851

Killen, M., & Smetana, J. G. (2010). Future directions: social development in the context of social justice. *Social Development*, *19*(3), 642-657. https://doi.org/10.1111/j.1467-9507.2009.00548.x

Killen, M., & Smetana, J.G. (2015). Origins and development of morality. In M. E. Lamb (Ed.), *Handbook of child psychology and developmental science, Vol. 3, 7th edition* (pp. 701-749). Editor-in-Chief, R. M. Lerner. NY: Wiley-Blackwell. https://doi.org/10.1002/9781118963418.childpsy317

Jennings, M. K., & Stoker, L. (2004). Social trust and civic engagement across time and generations. *Acta Politica*, *39*(4), 342-379. https://doi.org/10.1057/palgrave.ap.5500077

Larson, R., Brown, B. B., & Mortimer, J. (2002). Adolescents' preparation for the future: Perils and promise. *Journal of Research on Adolescence, 12*, 1-166. https://doi.org/10.1111/1532-7795.00023

Lenzi, M., Vieno, A., Sharkley, J., Mayworm, A., Scacchi, L., Pastore, M., & Santinello, M. (2014). How school can teach civic engagement besides civic education: The role of democratic school climate. *American Journal of Community Psychology, 54*(3), 251-261. https://doi.org/10.1007/s10464-014-9669-8

Lerner, R. (2004). *Liberty: Thriving and civic engagement among America's youth.* Thousand Oaks, CA: Sage.

Lewis, V. A., MacGregor, C. A., Putnam, R. D. (2013). Religion, networks, and neighborliness: The impact of religious social networks on civic engagement. *Social Science Research, 42*(2), 331-346. https://doi.org/10.1016/j.ssresearch.2012.09.011

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83,* 1198-1202. https://doi.org/10.1080/01621459.1988.10478722

Malti, T. (2016). Toward an integrated clinical-developmental model of guilt. *Developmental Review, 39,* 16-36. https://doi.org/10.1016/j.dr.2015.11.001

Malti, T., & Buchmann, M. (2010). Socialization and individual antecedents of adolescents' and young adults' moral motivation. *Journal of Youth and Adolescence, 39(2),* 138-149. https://doi.org/10.1007/s10964-009-9400-5

Malti, T., Gummerum, M., Keller, M., & Buchmann, M. (2009). Children's moral motivation, sympathy, and prosocial behavior. *Child Development, 80,* 442-460. https://doi.org/10.1111/j.1467-8624.2009.01271.x

Malti, T., & Keller, M. (2010). Development of moral emotions in cultural context. In W. Arsenio & E. Lemerise (Eds.), *Emotions, aggression, and morality in children: Bridging development and psychopathology* (pp. 177-198)*.* Washington, DC: American Psychological Association. https://doi.org/10.1037/12129-009

Malti, T., Keller, M., & Buchmann, M. (2013). Do moral choices make us feel good? The development of adolescent's emotions following moral decision-making. *Journal of Research on Adolescence, 23*(2), 389-397. https://doi.org/10.1111/jora.12005

Malti, T., & Krettenauer, T. (2013). The relation of moral emotion attributions to prosocial and antisocial behavior: A meta-analysis. *Child Development*, *84*(2), 397-412. https://doi.org/10.1111/j.1467-8624.2012.01851.x

Malti, T., & Ongley, S. F. (2014). The development of moral emotions and moral reasoning. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (2nd edition) (pp. 163-183). New York: Psychology Press. https://doi.org/10.4324/9780203581957.ch8

Marini, M. M. (2000). Social values and norms. In E. F. Borgatta, & R. J. V. Montgomery (Eds.), *Encyclopedia of sociology* (2nd ed., pp. 2828-2840). Detroit: Macmillan Reference.

Montada, L., & Schneider, A. (1989). Justice and emotional reactions to the disadvantaged. *Social Justice Research, 3*(4), 313-344. https://doi.org/10.1007/BF01048081

Muthén, L. K., & Muthén, B. O. (1998-2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Norton, M. I., & Ariely, D. (2011). Building a better America—One wealth quintile at a time. *Psychological Science*, *6*, 9-12. https://doi.org/10.1177/1745691610393524

Nie, N. H., & Hillygus, D.S. (2001). Education and Democratic Citizenship: Explorations into the Effects of What Happens in the Pursuit of the Baccalaureate. In *Education and Civil Society*, by D. Ravitch & J. Viteritti (Eds.). New Haven, CT: Yale University Press.

Obradovic, J., & Masten, A. S. (2007). Developmental antecedents of young adult civic engagement. *Applied Developmental Science, 11*, 2-19. https://doi.org/10.1207/s1532480xads1101_1

Ongley, S. F., & Malti, T. (2014). The role of moral emotions in the development of children's sharing behavior. *Developmental Psychology*, *50*(4), 1148-1159. https://doi.org/10.1037/a0035191

Parker, J. G., & Asher, S. R. (1993). Friendship and friendship quality in middle childhood: Links with peer group acceptance and feelings of loneliness and social dissatisfaction. *Developmental Psychology, 29,* 611-621. https://doi.org/10.1037/0012-1649.29.4.611

Piaget J. (1997). *The Moral Judgment of the Child.* New York: The Free Press (Original work published 1932).

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

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster. https://doi.org/10.1145/358916.361990

Rubin, K. H., Bukowski , W., & Laursen, B. (Eds.) (2009), *Handbook of peer interactions, relationships, and groups.* New York: Guilford.

Rubin, K. H., Malti, T., & McDonald, K. L. (2012). Civic development in relational perspective. In G. Trommsdorff & X. Chen (Eds.), *Values, religion, and culture in adolescent development* (pp. 188-208). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139013659.011

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*(4), 350-353. https://doi.org/10.1037/1040-3590.8.4.350

Schwartz, S. H., Verkasalo, M., Antonovsky, A., & Sagiv, L. (1997). Value priorities and social desirability: Much substance, some style. *British Journal of Social Psychology, 36*(1), 3-18. https://doi.org/10.1111/j.2044-8309.1997.tb01115.x

van Goethem, A. A. J., van Hoof, A., van Aken, M. A. G., Raaijmakers, A. W., Boom, J., & de Castro, B. O., (2012). The role of adolescents' morality and identity in volunteering. Age and gender differences in a process model. *Journal of Adolescence*, *35*, 509-520. https://doi.org/10.1016/j.adolescence.2011.08.012

Verba, S., Burns, N., & Scholzman, K. L. (2003). Unequal at the starting line: Creating participatory inequalities across generations and among groups. *The American Sociologist,* 34(1-2), 45-69. https://doi.org/10.1007/s12108-003-1005-y

Wilkinson, J. (2010). Personal communities: Responsible individualism or another fall for public [man]? *Sociology, 44(*4), 453-470. https://doi.org/10.1177/0038038510362484

Wray-Lake, L., & Syvertsen, A. K. (2011). The developmental roots of social responsibility in childhood and adolescence. *New Directions for Child and Adolescent Development*, *134*, 11-25. https://doi.org/10.1002/cd.308

Youniss, J., & Yates, M. (1999). Youth service and moral-civic identity: A case for everyday morality. *Educational Psychology Review*, *11*(4), 361-376. https://doi.org/10.1023/A:1022009400250

Zhou, Q., Valiente, C., & Eisenberg, N. (2003). Empathy and its measurement. In S. J. Lopez, & C. R. Snyder (Eds.), *Positive psychological assessment: A handbook of models and measures* (pp. 269-284). Washington, DC, US: American Psychological Association. https://doi.org/10.1037/10612-017

Table 1. Means and standard deviations of the main study variable by assessment point

s	Range	T1			T2			T3		
		N	Mean	SD	N	Mean	SD	N	Mean	SD
Attitudes toward inequality	1-6	--[a]	--[a]	--[a]	952	4.79	0.62	814	4.80	0.61
Social justice values	1-10	1257	8.66	1.17	952	8.69	1.18	815	8.65	1.09
Parent education	1-7	1241	4.16	1.28	--[b]	--[b]	--[b]	--[b]	--[b]	--[b]
Adolescent education	1-5	1245	3.32	1.01	950	3.63	1.00	--[b]	--[b]	--[b]
Friendship quality	1-6	1255	5.42	0.69	946	5.69	0.40	811	5.70	0.43
Sympathy (S)	1-6	1257	4.81	0.76	952	4.81	0.72	815	4.84	0.70
Sympathy (P)	1-6	1048	4.92	0.95	--[b]	--[b]	--[b]	--[b]	--[b]	--[b]
Moral emotions	0-1	1227	0.81	0.26	929	0.84	0.27	743	0.89	0.22

Note. [a] The measure was not included due to low reliability. [b] No scale available at assessment point. S = Self-report. P = Primary caregiver.   T1 = Time 1. T2 = Time 2. T3 = Time 3.

Table 2. Bivariate correlations between study variables

	1	2	3	4	5	6	7	8	9	10
1. Attitudes toward inequality T3	--									
2. Attitudes toward inequality T2	.50***	--								
3. Social justice values T3	.41***	.35***	--							
4. Social justice values T2	.30***	.45***	.52***	--						
5. Social justice values T1	.17***	.23***	.28***	.29***	--					
6. Gender	-.10**	-.13***	-.22***	-.24***	-.22***	--				
7. Parent education	.15***	.11**	-.06	-.06	.02	.06*	--			
8. Adolescent education T2	.20***	.11**	.01	-.02	.12***	-.03	.30***	--		
9. Adolescent education T1	.16***	.13***	-.02	-.04	.07*	-.06*	.34***	.57***	--	
10. Friendship quality T2	.10**	.20***	.19***	.25***	.10**	-.23***	-.05	.001	.03	--
11. Friendship quality T1	.04	.14***	.12**	.18***	.22***	-.41***	-.06	-.06	-.01	.34***
12. Sympathy (S) T2	.28***	.40***	.31***	.36***	.21***	-.24***	.02	.02	.06	.24***
13. Sympathy (S) T1	.18***	.25***	.23***	.24***	.40***	-.26***	-.02	.06	.05	.15***
14. Sympathy (P) T1	.15***	.21***	.18***	.16***	.19***	-.20***	.01	.06	.06	.08*
15. Moral emotions T2	.14***	.18***	.18***	.19***	.14***	-.17***	.02	.07*	.07*	.11**
16. Moral emotions T1	.05	.12***	.06	.15***	.13***	-.12***	.01	.04	.02	.07*

Table 2 (Con't)

	11	12	13	14	15
11. Friendship quality T1	--				
12. Sympathy (S) T2	.17***	--			
13. Sympathy (S) T1	.30***	.41***	--		
14. Sympathy (P) T1	.12***	.19***	.22***	--	
15. Moral emotions T2	.10**	.22***	.14***	.16***	--
16. Moral emotions T1	.04	.18***	.19***	.17***	.27***

Note. Gender: female = 0, male = 1. S = Self-report; P = Parent-report. T1 = Time 1. T2 = Time 2. T3 = Time 3.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3. Multivariate regression analyses predicting attitudes toward Inequality and social justice values at T2 and T3

Model 1	Attitudes Toward Inequality T2				Social Justice Values T2		
	β	SE	R^2		β	SE	R^2
	--	--		Social justice values T1	0.15***	0.04	
Gender	-0.02	0.03		Gender	-0.13***	0.03	
Parent education	0.10**	0.03		Parent education	-0.03	0.03	
Adolescent education T1	0.06†	0.03		Adolescent education T1	-0.06†	0.03	
Friendship quality T1	0.07*	0.03		Friendship quality T1	0.05	0.04	
Sympathy (S&P) T1	0.24***	0.03		Sympathy (S&P) T1	0.14***	0.04	
Moral emotions T1	0.04	0.04		Moral emotions T1	0.08*	0.04	
Total R^2			.11***	Total R^2			.12***
N			922	N			922

Model 2	Attitudes Toward Inequality T3				Social Justice Values T3		
	β	SE	R^2		β	SE	R^2
Attitudes toward Inequality T2	0.41***	0.04		Social justice values T2	0.39***	0.05	
Gender	-0.02	0.03		Gender	-0.07*	0.03	
Parent education	0.10**	0.04		Parent education	-0.03	0.04	
Adolescent education T2	0.09*	0.04		Adolescent education T2	-0.02	0.04	
Friendship quality T2	0.04	0.04		Friendship quality T2	0.05	0.04	
Sympathy (S) T2	0.09*	0.04		Sympathy (S) T2	0.14***	0.04	
Moral emotions T2	0.03	0.04		Moral emotions T2	0.07†	0.04	
Total R^2			.26***	Total R^2			.27***
N			766	N			766

Note. Gender: female = 0, male = 1.   S = Self-report; S&P = Self- and parent-reports combined; T1 = Time 1. T2 = Time 2. T3 = Time 3. † $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.
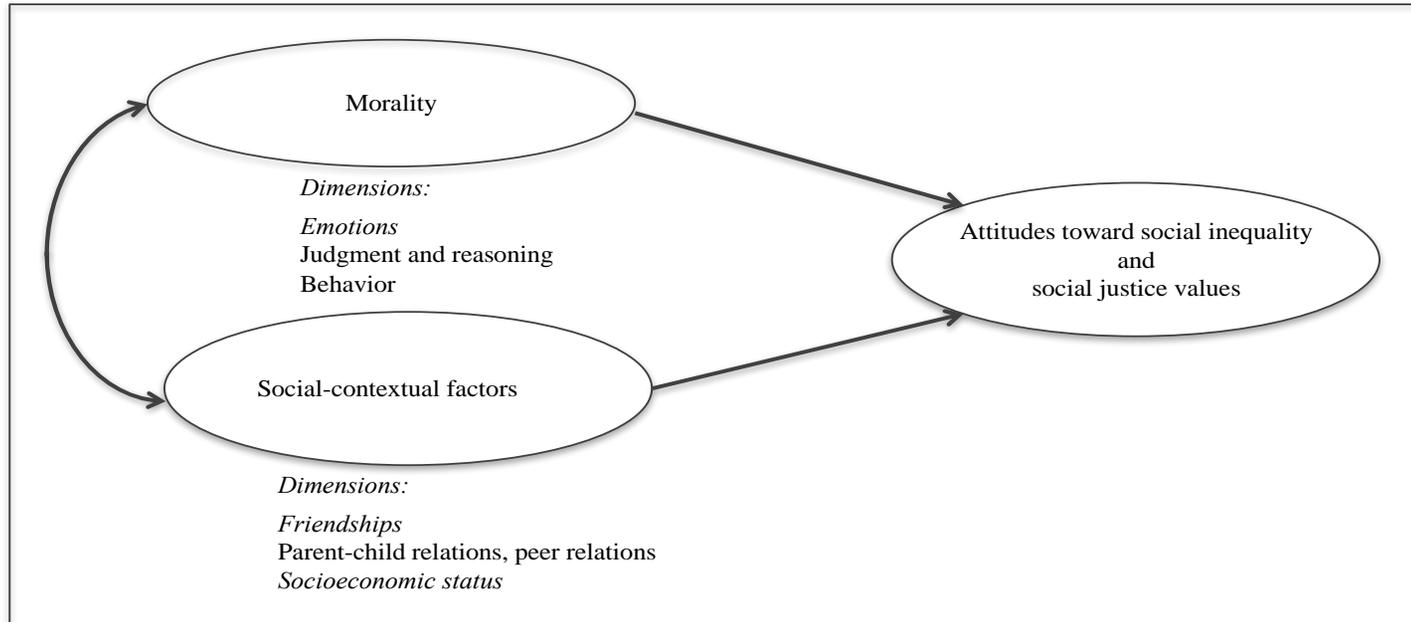
Note: Italicized dimensions of constructs are the focus of the present study.

Figure 1. Simplified conceptual model of attitudes toward social inequality and social justice values and their moral and social-contextual antecedents.

Developmental dynamics between young adults' life satisfaction and engagement with studies and work

Katja Upadyaya   University of Jyväskylä, Finland, University of Helsinki, Finland
katja.upadyaya@gmail.com
Katariina Salmela-Aro   University of Jyväskylä, Finland, University of Helsinki, Finland

Abstract

The present five-wave longitudinal study investigated the cross-lagged associations between young adults' life satisfaction and study/work engagement over the transition from post-comprehensive studies to higher education or work during the second and third decades of life. Gender, educational track, academic performance and family socioeconomic status were also examined. The study is part of the longitudinal Finnish Educational Transitions (FinEdu) study, and used data from secondary education onwards, following 821 participants from age 17 to 25. The developmental dynamics showed that, in particular, young adults' life satisfaction predicted their study/work engagement both during their post-comprehensive education and after the transition to higher education or work. Moreover, study/work engagement positively predicted young adults' life satisfaction during their third decade of life. In addition, high initial life satisfaction was more typical among males. However, no differences related to gender or academic track were observed in the developmental dynamics of life satisfaction and study/work engagement. These results suggested that general wellbeing spills over to study/work domain-specific characteristics of wellbeing and promotes positive personal development and adjustment to study/work transitions during the third decade of life.

Keywords

Life satisfaction, study/work engagement, study/work transition, young adults, longitudinal studies

Introduction

The third decade of life is a period during which individuals are faced with more role transitions and life decisions than at any other life stage (Caspi, 2002). Rindfuss (1991) describes the period between the ages of 18 and 30 as demographically dense, as many key role transitions take place then, including moving from education to employment (Elder & Shanahan, 2006). Moreover, at the moment the global economic crisis is affecting Finland and young people need to navigate to adulthood in challenging times. Thus, during the third decade personal (e.g., life satisfaction) and study/work-related wellbeing (e.g., engagement) become increasingly important.

Engagement with studies and work is characterised by a committed study- and work-related mindset (Fredricks, Blumenfeld, & Paris, 2004; Li & Lerner, 2011; Schaufeli, Salanova, Gonzalez-Roma, & Bakker, 2002), which predicts many long-term positive outcomes, such as higher education, academic success (Annunziata, Hogue, Faw, & Liddle, 2006), better job prospects (Upaydyaya & Salmela-Aro, 2013b), and positive self-perception (Linnakylä & Malin, 2008). Life satisfaction reflects one's general happiness and wellbeing (Lewis, Huebner, Malone & Valois, 2011), whereas study/work engagement, similar to flow (Csikszentmihalyi, 1990), describes one's domain-specific experiences and wellbeing (Hakanen & Schaufeli, 2012; Upadyaya & Salmela-Aro, 2013a). Despite these differences, satisfaction in life is often associated with study/work engagement during different life stages. For example, high engagement in studies is positively associated with life satisfaction in middle school (Lewis et al., 2011), and also manifests as an increase in life satisfaction after finishing vocational/high school education (Salmela-Aro & Upadyaya, 2014). Work engagement, in turn, predicts one's subsequent life satisfaction over several years (Hakanen & Schaufeli, 2012). These previous results also suggest that various positive gain spirals and developmental cascades occur between study/work engagement and life satisfaction across different life stages (see also Masten, Desjardins, McCormick, Kuo, & Long, 2010). However, to the authors' knowledge, life satisfaction and study and work engagement have not been examined longitudinally in the same study, focusing on young adults' transition to higher education/work and taking into account developmental cascades across life stages. The present study seeks to fill this gap by following young adults from their second to their third decade of life.

Life satisfaction

Increasing interest in life satisfaction reflects the growing attention to positive psychology, which focuses on human strengths and optimal functioning rather than on weaknesses and malfunctioning (Hakanen & Schaufeli, 2012; Seligman & Csikszentmihalyi, 2000). Life satisfaction is an important indicator of positive psychological wellbeing (Huebner, Valois, Paxton, & Drane, 2005), as it describes one's general happiness and cognitive assessments of life quality based on one's own standards (Erdogan, Bauer, Truxillo, & Mansfield, 2012; Lewis et al., 2011; Pavot & Diener, 1993). At the beginning of their third decade Finnish young adults often face several transitions simultaneously (e.g., moving out from the childhood home, transitioning to new educational institutions/work, changes in social networks). They face increasing environmental and personal demands (e.g., adjusting to the place of study/work and to more independent life, attaining competence in studies/work) and decreases in the social support they received previously (see also Suldo & Huebner, 2006). Thus, satisfaction with life may serve as a personal resource that promotes adjustment (Cohn, Fredrickson, Brown, Mikels, & Conway, 2009) and becomes increasingly important at the beginning of the third decade when young adults are starting their careers and adjusting to their new work environments (see also Arnett, 2000). However, despite the positive associations between life satisfaction and academic (e.g., achievement, competence, school satisfaction; Suldo, Riley, & Shaffer, 2006) and work outcomes (e.g., job performance, commitment), only a few studies have examined the associations between life satisfaction and academic- (Lewis et al., 2011) and work-related wellbeing (Erdogan et al., 2012). These studies have shown that most adolescents and young adults report positive global life satisfaction, although variation exists across domains; middle-school and high-school students typically report higher dissatisfaction with their school experiences compared to other life domains (Huebner, Drane, & Valois, 2000; Huebner et

al., 2005). Moreover, young adults who have received the high school diploma report higher levels of life satisfaction than their peers who obtained only general educational development tests or dropped out of school (Ou, 2008). Among adult workers, life satisfaction is positively associated with several work-related variables, such as organisational commitment and personal growth (Erdogan et al., 2012); some evidence suggests that life satisfaction is more strongly related to job performance than job satisfaction (Jones, 2006).

Moreover, gender differences may exist in life satisfaction. Although some studies have reported higher study satisfaction among females than males, gender differences in global life satisfaction have been mixed, with some studies reporting similar levels of life satisfaction among females and males (Elmore & Huebner, 2010; Huebner et al., 2000), and other studies reporting higher life satisfaction among males than females (Diseth, Danielsen, & Samdal, 2012). Moreover, a notable amount of spillover exists between different variables related to work (e.g., job involvement, stress, satisfaction) and personal life (e.g., family conflict, stress, satisfaction) (Ford, Heinen, & Langkamer, 2007), and thus, feelings of study/work engagement may be reciprocally related to overall experiences of life satisfaction. Further, typically, satisfaction with life is relatively stable over time, although the antecedents and outcomes of life satisfaction may vary at different ages and life stages (Myers & Diener, 1995). For example, the influence of family socioeconomic status (SES) on life satisfaction may be more pronounced among children (Seligson, Huebner, & Valois, 2003) than adults (Louis & Zhao, 2002).

Study and work engagement

Recent research suggests that the underlying construct of engagement is the same for students and employees (Wefald & Downey, 2009). Studies drawing on the work engagement literature (Hakanen, Bakker, & Schaufeli, 2006; Salanova, Agut, & Peiro, 2005; Schaufeli, Bakker, & Salanova, 2006) have examined students' and workers' psychological engagement resembling flow (Csikszentmihalyi, 1990) in detail, including the dimensions of energy, dedication, and absorption while studying or working (Salmela-Aro & Upadyaya, 2012; Schaufeli, Salanova,

Gonzalez-Roma, & Bakker, 2002; Upadyaya & Salmela-Aro, 2014). Flow refers to an experience which itself is so enjoyable that people tend to pursue it, even at great cost, purely for the sake of the experience (Csikszentmihalyi, 1990). However, the main difference between the concepts of flow and engagement is that flow refers to a short-term (however, often recurrent) peak experience that occurs infrequently in studies/work, whereas engagement is a more general and persistent state of mind. *Energy* refers to high mental resilience and affects while studying, a willingness to invest effort in one's schoolwork, and a positive approach. *Dedication,* in turn, is characterised by a sense of significance, enthusiasm, pride, and inspiration regarding school, as well as perceiving schoolwork as meaningful. *Absorption* is characterised by behavioural accomplishments, fully concentrating and being happily engrossed in one's studying so that time passes quickly. These three dimensions are distinct analytic constructs of study and work engagement, although they correlate highly with each other (Salmela-Aro & Upadyaya, 2014; Schaufeli et al., 2002). They may be considered as a one-dimensional or as a three-dimensional construct, depending on the research purpose (Seppälä et al., 2009). Either approach elucidates the development of engagement during the school-to-work transition in association with general wellbeing (see also Dietrich, Parker, & Salmela-Aro, 2012 for phase specific engagement; Haase, Heckhausem, & Köller, 2008).

Moreover, differences related to gender and socioeconomic status may influence study/work engagement: females typically experience higher levels of overall study engagement than males (Marks, 2000; Salmela-Aro & Upadyaya, 2014), probably because female students tend to perform better academically (Pomerantz, Altermatt, & Saxon, 2002) and attribute greater importance to academic achievement than males (Berndt & Miller, 1990). Similarly, males and students from lower income families are more likely than females and students from higher socioeconomic backgrounds to experience rapid decrements in engagement and to follow unstable school engagement trajectories which may sometimes lead to school dropout (Archambault, Janosz, Morizot, & Pagani, 2009; Li & Lerner, 2011).

Moreover, in some countries men score higher in their engagement with work whereas in other countries the reverse is true (Schaufeli et al., 2006). The fact that not all studies report gender differences in student (Simons-Morton & Crump, 2003) or employee engagement (Langelaan, Bakker, Doornen, & Schaufeli, 2006) highlights the need for further research in different societal contexts.

High academic performance also promotes high engagement in studies (Li & Lerner, 2011; Salmela-Aro & Upadyaya, 2012) and in work (Upadyaya & Salmela-Aro, 2013b). Further, compared to their peers following a vocational track, young adults on an academic track often feel more exhausted and less engaged in their studies (Salmela-Aro & Upadyaya, 2012), which may also manifest in the associations between study/work engagement and life satisfaction.

Study/work engagement and life satisfaction

According to the demands and resources model (Demerouti, Bakker, Nachreiner, & Schaufeli, 2001), high engagement protects against malaise and burnout symptoms, and leads to wellbeing and high life satisfaction (Salmela-Aro & Upadyaya, 2014). Similarly, spillover exists between study- and work-domain specific wellbeing/ill-health and between general wellbeing/ill-health (Hakanen & Schaufeli, 2012; Salmela-Aro & Upadyaya, 2014; Upadyaya & Salmela-Aro, 2016; in press). Thus, positive development in one life domain (e.g., general wellbeing and life satisfaction) may influence adaptation in other life domains and promote domain-specific wellbeing, reflecting developmental cascades that occur during the second and third decades and earlier (see also Masten et al., 2010). In addition, the 'broaden-and-build' theory suggests that students who are satisfied with their studies and life, and experience frequent positive emotions, will exhibit adaptive coping behaviours, feel more engaged, and gain more resources which, in turn, will promote positive upward spirals of success at studies/work (see also Fredrickson, 2001). These positive cycles may become especially important during the third decade of life when decreases occur in the social support young adults typically receive (see also Suldo & Huebner, 2006) as they simultaneously face various personal, educational,

and work transitions and prepare for adult work roles (Arnett, 2000). At the same time young adults' personal agency becomes more important (Evans, 2007) and their lives are less structured than during the second decade. From the phase adequate engagement perspective (Dietrich et al., 2012), general wellbeing and study/work engagement are closely tied together especially after the transition to higher education or work when obtaining a good study/work status is one of young adults' most important goals. Life satisfaction, in particular, may serve as a cognitive resource, derived from frequent positive emotions, which helps in facing a wide range of life's challenges and supports positive personal growth (Cohn et al., 2009). Thus, high life satisfaction may promote study/work engagement, which, in turn, may lead to an increase in subsequent life satisfaction. Empirical studies have shown that engagement and life satisfaction are typically highly correlated with each other (Harter, Smidth, & Hayes, 2002; Lewis et al., 2011). However, only a few studies have investigated the cross-lagged associations between engagement and life satisfaction (Lewis et al., 2011) and longitudinal studies are needed to examine these associations further during different developmental stages. Consequently, the present study aims at investigating the cross-lagged associations between young adults' life satisfaction and study/work engagement over a period of eight years, following young adults from their second to their third decade of life.

Variation in the associations between engagement and life satisfaction across the transition from post-comprehensive studies to higher education/work could reflect the changing fit between person and environment (Eccles & Roeser, 2009) across grade levels and age (Fredricks et al., 2004). During their second decade and before the transition to higher education/work, young adults typically have various sources of support (e.g., parents, teachers, peers) that may promote their life satisfaction (Suldo & Huebner, 2006), which spills over to domain-specific engagements. During the transition, when young adults are adjusting to their new study/work environments and finding their place in their new educational/work settings, associations between life satisfaction and study/work engagement may

momentarily diminish. However, after the transition and at the beginning of their third decade when their new study/work environments become more salient, study/work engagement may increasingly predict life satisfaction. Consequently, to better understand the dynamics of life satisfaction and study/work engagement, the present study investigates the cross-lagged associations between life satisfaction and study/work engagement as young adults make the transition from post-comprehensive studies to higher education studies/work.

Schooling in Finland

Compulsory comprehensive education in Finland lasts for nine years until the students are 16 years old. After that, approximately 50% of adolescents enter senior high schools and approximately 41% go to vocational schools (*School Statistics,* Central Statistical Office of Finland, 2010). Average academic achievement in the ninth grade is the minimum requirement for admission to senior high school. Both senior high schools and vocational education take three to four years to complete, after which students may apply to institutes of higher education. Approximately 39% of high school graduates start studying, 44% begin working, and 25% are studying and working one year after finishing high school, whereas 8% of young adults with a degree from a vocational school are studying, 69% are working, and 10 % are both studying and working one year after their graduation.

Hypotheses

The aim of the present five-wave longitudinal study was to investigate the cross-lagged associations between life satisfaction and study/work engagement during and after the transition from post-comprehensive studies to higher education or work. In addition, the role of gender, academic performance, family socioeconomic status and academic track (e.g., vocational and academic track) in these associations was examined. Based on previous studies of developmental cascades (Masten et al., 2010)), phase adequate engagement (Dietrich et al., 2012), and study- and work domain-specific and general wellbeing (Hakanen & Schaufeli, 2012; Salmela-Aro & Upadyaya, 2014; Upadyaya & Salmela-Aro, 2015), we expected that spillover would occur

between life satisfaction and study/work engagement. That is, these psychological constructs would be positively associated with one another both during post-comprehensive studies and after transition to higher education/work at the beginning of the participants' third decade. More specifically, and following the broaden-and-build theory (Fredrickson, 2001) we expected that life satisfaction predicts study/work engagement during the participants' second and third decade (H1), whereas study/work engagement predicts life satisfaction more prominently after the transition to higher education/work and during the participants' third decade (H2). In addition, we expected that males and students with high academic performance would experience higher levels of life satisfaction (H3) (Diseth et al., 2012; Ou, 2008). Finally, we expected that females and students following a vocational track or with high academic performance would experience higher levels of study/work engagement (H4) (Salmela-Aro & Upadyaya, 2012; 2014).

Method

The data were drawn from the Finnish Educational Transitions (FinEdu) study, which recruited all the ninth-grade students in a medium-sized town (population = 88,000) in Central Finland. The present study used the data from post-comprehensive education onwards. Students from 13 post-comprehensive schools (six high schools and seven vocational schools) participated in the study. The first two years of measurements were carried out during each of the two post-comprehensive years: the first in 2005 at half a year after the transition to post-comprehensive school (Time 1, age = 17, N = 818) and the second one year later in 2006 (Time 2, age = 18, N = 749). The third measurement was carried out two and a half years after Time 2 in 2009 (Time 3, age = 21, N = 611), when most of the students had already finished their post-comprehensive education. The fourth and the fifth measurements were carried out two (2011) and four years (2013) after Time 3 (Time 4, age = 23, N = 561; Time 5, age = 25, N = 533). Thus, the data collection spanned the periods before, during and in the aftermath of the Great Recession.

Reflecting the general population in Central Finland (Kuopion Lukiokoulutus, 2009), the majority of the

participants (99%) were Finnish-speaking. The students most often lived with both parents (62%), or with their mother or father, either as a lone parent (25%), or living with her/his new spouse (11%), or with somebody else (1%). A total of 23% of the adolescents had siblings. The occupational distribution of the parents was as follows: 27% of the fathers and 20% of the mothers worked in higher white-collar occupations (e.g., medical doctors), 16.4% and 49% worked in lower white-collar occupations (e.g., teachers), 36% and 17% had blue-collar occupations (e.g., cooks, bus drivers), 11% and 4% were private entrepreneurs, 1% and 2% were students, 3% and 2% were retired, and 5% and 6% had some other status (e.g., unemployed). The occupational distribution is comparable to the general population in Eastern Finland. The questionnaires (Time 1- 2) were group-administered to the students in their classrooms during regular school hours, and at Time 3-5 were mailed to the participants.

Measures

Study and work engagement were measured using the abbreviated student version of the Utrecht Work Engagement Scale, UWES-S (Salmela-Aro & Upadyaya, 2012; for validity and reliability of the measure; Schaufeli et al., 2002; Schaufeli et al., 2006; Seppälä et al., 2009 for construct validity). At Times 1 and 2 the inventory concerned engagement in studies, and at Time 3-5 it covered engagement with studies or work, depending on the participant's current situation. The scale consists of nine items measuring energy (e.g., *When I study/work, I feel that I am bursting with energy*), dedication (e.g., *I am enthusiastic about my studies/work*), and absorption (e.g., *Time flies when I'm studying/working*) in relation to studies and work. The responses were rated on a seven-point scale (0 = not at all; 6 = daily), and an additive score was calculated to represent the participants' overall engagement in studies/work at each measurement time. To examine attrition between the measurements, young adults who participated in the study at each wave (N = 370) were compared with those who had missing data (N = 474). The results indicated that those young adults who participated in the study at each measurement showed higher engagement at Time 1 (M= 3.73, SD = 0.07), Time 2 (M= 3.65, SD = 0.7), and Time 3 (M =

4.40, SD = 0.06) than those who did not (M = 3.45, SD = 0.07, $t(-2.97)$ =, $p <.01$; M = 3.32, SD = 0.08, $t(-3.16)$ = $p <.01$; M = 4.17, SD = 0.10, $t(-2.16)$= $p <.05$). At Times 4 and 5 no differences emerged in young adults' engagement due to attrition.

Life satisfaction was assessed using the five-item Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin, 1985). The items (e.g., *I am satisfied with my life*) were rated on a five-point scale ranging from 1 (*I totally disagree*) to 5 (*I totally agree*). An additive score was calculated from all five items at each wave. The attrition analyses showed that those young adults who participated in the study at each wave showed higher life satisfaction at Time 4 (M = 4.97, SD = 0.06) and Time 5 (M = 4.63, SD = 0.07) than those who did not (M = 4.67, SD = 0.09, $t(-2.70)$ = $p < .01$; M = 4.20, SD = 0.10, $t(-3.59)$ = $p < .001$). No differences emerged at Times 1-3 in young adults' life satisfaction due to attrition.

Academic performance was measured in accordance with the Grade Point Average (GPA) of the final comprehensive school report on a scale ranging from 4 (lowest) to 10 (highest).

Gender was coded 1 = female, 2 = male.

School track (Time 2) was coded as 1 = high school student (*N* = 529); 0 = vocational school student (*N* = 226);

Family SES was coded as 1 = blue collar, 2 = white collar.

Attrition analyses showed that females (M = 1.41, SD = 0.49) and students following an academic track (M= 1.79, SD = 0.41) more often than males (M= 1.63, SD = 0.48, $t(7.43)$= $p <.001$) and students following a vocational track (M= 1.59, SD = 0.49, $t(-5.86)$= $p <.001$) participated in the study at all waves.

Analysis strategy

The aim of this study was to analyse the associations between study/work engagement and life satisfaction using cross-lagged path models. In these models, all the endogenous variables were allowed to covary. The tested model included stability coefficients for study/work engagement and life satisfaction, as well as cross-lagged paths between all the engagement and life satisfaction variables. The participants' gender, family SES, academic track, and performance were added in the model as

antecedents of study/work engagement and life satisfaction (Time 1). In the final model, all of the statistically non-significant paths were set to zero. Moreover, in order to examine whether the models would show the same fit for males and females and for students on the academic and those on the vocational track, all the analyses were also carried out using the Mplus multigroup procedure (Muthén & Muthén, 1998-2016).

The statistical analyses were performed using the Mplus statistical package (Version 6; Muthén & Muthén, 1998-2016) with the missing data method. This missing data method uses all the data that are available in order to estimate the model without imputing the data. Because the distributions of the variables were skewed, the model parameters were estimated using the MLR estimator (Muthén & Muthén, 1998-2016). Goodness-of-fit was evaluated using various indicators: χ^2 test, Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR). Because the χ^2 test is sensitive to sample size, a relative goodness-of-fit index was also used to evaluate the model fit: the Comparative Fit Index (CFI). According to Hu and Bentler (1999), cutoff values close to 0.95 for the CFI, cutoff values close or below to 0.06 for the RMSEA, and a cutoff value close to or below 0.08 for the SRMR can be considered as indicating a good fit between the hypothesised model and the observed data.

Results

Means, variances, correlations, and Cronbach's alpha reliabilities are presented in Table 1. A path model was constructed to examine the cross-lagged associations between the young adults' life satisfaction and engagement in their studies/work. After all the non-significant paths were set to zero, the final model fitted the data well (χ^2(63, N = 1538) = 161.77, p = ns, CFI = .93, RMSEA = .03, SRMR = .07; Figure 1). The results showed that, as expected, despite the high stability of the dependent variables, several cross-lagged paths were identified between study/work engagement and life satisfaction. Life satisfaction, in particular, positively predicted young adults' subsequent study/work engagement in studies (Time 2; $s.e.$ = .10, p <.05) and higher education/work

(Times 3-5) at the beginning of the participants' third decade (H1). In addition, young adults' engagement in higher education studies/work positively predicted their subsequent life satisfaction (Times 4-5) at the beginning of their third decade (H2). Further, as expected (H3 and H4), the results showed that academic performance positively predicted young adults' life satisfaction ($s.e.$ = .26, p <.001) and study/work engagement (Time 1; $s.e.$ = .16, p <.001), and that young males showed a higher level of life satisfaction (Time 1; $s.e.$ = .13, p <.001) than their female counterparts.

In order to examine whether the model would show the same fit for both males and females, and for both students on the academic and those on the vocational track, the analyses were also carried out using the multigroup procedure (Muthén & Muthén, 1998-2016). In these analyses, the data were divided into two samples: first into males and females, and then into students following an academic and those following a vocational track. Path analyses were then carried out, assuming that all parameters would be equal for these two groups. If the fit of the model was good and no significant modification indices were found, the model would be assumed to fit the groups equally. The results indicated moderately good fit for both models (χ^2(124, N = 856) = 229.33, p = ns, CFI = .93, RMSEA = .05, SRMR = .08 for the gender multigroup model and χ^2(123, N = 755) = 255.86, p = ns, CFI = .91, RMSEA = .05, SRMR = .09 for the track multigroup model). No differences related to gender or study track were found in these multigroup analyses.

Discussion

To summarise, the present study examined the cross-lagged paths between life satisfaction and study and work engagement by following young adults from age 17 to 25 over their transition from post-comprehensive education to higher education or work, and until the beginning of their third decade. The developmental dynamics between life satisfaction and engagement showed that life satisfaction in particular positively predicted young adults' subsequent engagement in their post-comprehensive studies and in higher education or work. Study/work engagement, in turn, predicted young adults' life

satisfaction after their transition to higher education/work. High initial life satisfaction was more typical among males than females. However, no differences related to gender or academic track were found in the associations between life satisfaction and study/work engagement.

Cross-lagged associations between life satisfaction and study/work engagement

The first aim of the present study was to investigate the cross-lagged associations between young adults' life satisfaction and engagement during and after their post-comprehensive studies. The results showed that, as expected (H1), positive spillover occurred between life satisfaction and engagement (Masten et al., 2010). Confirming our hypotheses, life satisfaction, in particular, positively predicted young adults' subsequent study/work engagement during their post-comprehensive studies and after the transition to higher education/work (H1). Young adults' study/work engagement, in turn, positively predicted their subsequent life satisfaction only later on after the transition to higher education or work at the beginning of the third decade (H2). These results are similar to research on flow (Csikszentmihalyi, 1990), an experience which study/work engagement resembles (Upadyaya & Salmela-Aro, 2013a). Life satisfaction and flow are strongly associated because optimal challenge and being immersed in meaningful activities add to satisfaction (Myers & Diener, 1995). In our investigation, study or work engagement predicted life satisfaction during the two last measurement points at the beginning of the third decade of life. By this time the young adults had already transitioned to higher education or work, and had probably better adapted to their new educational/occupational environments and experienced stronger stage-environment fit (Eccles & Roeser, 2009) than earlier during the transition. Moreover, our results showed that changes occur in these associations over time. Throughout the period of study, when young adults' life satisfaction is high, he/she often has experiences similar to flow and feels more engaged in his/her studies and work. It is also possible that high life satisfaction leads to higher job satisfaction, which in turn increases job engagement (see also Erdogan et al., 2012; Upadyaya & Salmela-Aro, in press).

Furthermore, these results suggested that study/work engagement becomes increasingly important at the beginning of the third decade when young adults are preparing for adult work roles (Arnett, 2000): wellbeing in one area of life spills over to life in general as a part of developmental cascades (Masten et al., 2010) promoting young adults' adjustment to their adult work roles.

These results thus suggested that aspects of overall wellbeing spill over to domain-specific experiences, which in our study represented flow-like experiences in studies/work. Previous research on burnout symptoms and malaise have shown that the contrary is usually the case. Domain-specific job burnout symptoms spill over to general wellbeing and show as a decrease in one's overall life satisfaction (Hakanen & Schaufeli, 2012; Salmela-Aro, Savolainen, & Holopainen, 2009), possibly because job burnout also affects psychological and physiological health (Maslach, Schaufeli, & Leiter, 2001). There are at least two possible explanations for these differences: first, it is possible that when one's wellbeing is high, the spillover effects between general and domain-specific experiences are positive and are transmitted in the direction from general wellbeing to domain-specific, whereas when one's wellbeing is low, the spillover effects between general and domain- specific experiences are negative and are transmitted in the opposite direction. However, the results of this study suggest a developmental trend. During the transition to higher education/work (at the beginning of the third decade), when the young adults are adapting to a new study/work environment, study/work-specific experiences might be more affected by one's overall wellbeing, whereas later on while facing the new study/work demands and challenges, study/work-specific experiences increasingly spill over to general wellbeing.

Antecedents of life satisfaction and study/work engagement

In this study, the antecedents of life satisfaction and study/work engagement were similar to those of several previous studies. Young males experienced a higher level of life satisfaction than young females (Marks, 2000; Diseth et al., 2012), and academic performance positively predicted young adults' initial life satisfaction (Suldo, Riley, & Shaffer, 2006) and

study/work engagement (Li & Lerner, 2011; Salmela-Aro & Upadyaya, 2012), as expected (H3 and H4). However, no differences related to gender or academic track emerged in the developmental dynamics of life satisfaction and study/work engagement. These results suggest that the longitudinal associations between life satisfaction and study/work engagement are similar for males and females and for young adults following either a vocational or academic track (see also Huebner et al., 2000).

Limitations and future directions

Several limitations should be taken into consideration when generalising the results of the present study. First, this study focused on young adults during their post-comprehensive education and over their transition to higher education/work. Thus, the results can be generalised only to the same age group. It is possible that the associations between life satisfaction and study/work engagement would have been different had different age groups been studied (see also Fredricks et al., 2004). More studies are needed to examine these associations among a range of age groups. Second, some of the parameters reported in the present study had relatively small values, probably due to the high stability of life satisfaction and study/work engagement. Future studies are needed to confirm these results. Third, young adults who participated in the study at each measurement time showed higher engagement at the beginning of this study and higher life satisfaction towards the end of the study. Fourth, our results only described the associations between overall life satisfaction and study/work engagement, whereas it is possible that the associations between life satisfaction and the separate engagement dimensions would have been more pronounced. Fifth, it is also possible that various latent trajectory groups exist that reflect different levels and development of study/work engagement and life satisfaction. For example, some young adults may experience an overall high level of engagement and life satisfaction whereas among other young adults, engagement and life satisfaction may change across educational/work transitions. Research utilising a longitudinal design and varying study/work contexts is needed to

examine these associations further. Finally, causal inferences are limited owing to the possibility that other, unmeasured variables, such as support from parents (Simons-Morton & Crump, 2003; Upadyaya & Salmela-Aro, 2013a, 2013b) and teachers (Murray, 2009), may have contributed to the results. The data presented in this study was also gathered before, during, and after the Great Recession, which might have heightened young adults' concern about adjusting to the new educational/work environments and increased the importance of engagement for life satisfaction.

Conclusions

The results of this study showed that young adults' life satisfaction predicted their study/work engagement across the transition from post-comprehensive studies to higher education/work and at the beginning of the participants' third decade, whereas study/work engagement predicted young adults' life satisfaction only after the transition to higher education or work at the beginning of their third decade. These results showed that third decade is an important period when wellbeing in one area of life easily spills over to another area of life. Moreover, during the third decade spillover occurs between general and domain-specific characteristics of wellbeing in both directions, whereas earlier during the second decade spillover only occurs from life satisfaction to study/work engagement.. These developmental changes may be due to the changes in stage-environment-fit (Eccles & Roeser, 2009) and reflect young adults' phase-adequate engagement (Dietrich et al., 2012). In addition, these findings supported the ideas of the broaden-and-build theory according to which students who are satisfied with their studies and life, and experience frequent positive emotions, also exhibit adaptive coping behaviours, feel more engaged, and gain more resources, which, in turn, promote positive upward spirals of success in studies/work (see also Fredrickson, 2001). Thus, in future research and education it would be important to take into the account the role of life satisfaction, as it is important in maintaining one's psychological wellbeing and in shaping one's engagement and other experiences in studies/work, and in supporting one's further

development across the life course. The indicators of general wellbeing (e.g., life satisfaction, happiness) may also spill over to academic/work-related engagement, help in adapting to new study/work environments (see also Masten et al., 2010), increase stage-environment fit (Eccles & Roeser, 2009), and support the young person's life course development (Haase et al., 2008).

Acknowledgements

References

Annunziata, D., Hogue, A., Faw, L., & Liddle, H. A. (2006). Family functioning and school success in at-risk, inner-city adolescents. *Journal of Youth and Adolescence, 35,* 105-113. https://doi.org/10.1007/s10964-005-9016-3

Archambault, I., Janosz, M., Morizot, J., & Pagani, L. (2009). Adolescent behavioral, affective, and cognitive engagement in school: Relations to dropout. *Journal of School Health, 79,* 408-415. https://doi.org/10.1111/j.1746-1561.2009.00428.x

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American psychologist*, *55*(5), 469. https://doi.org/10.1037/0003-066X.55.5.469

Babin, B. J., & Boles, J. S. (1998). Employee behavior in a service environment: a model and test of potential differences between men and women. *The Journal of Marketing*, 77-91. https://doi.org/10.2307/1252162

Berndt, T. J., & Miller, K. E. (1990). Expectancies, values, and achievement in junior high school. *Journal of Educational Psychology, 82,* 319-326. https://doi.org/10.1037/0022-0663.82.2.319

Caspi, A. (2002). 11 Social selection, social causation, and developmental pathways: empirical strategies for better understanding how individuals and environments are linked across the life-course. *Paths to successful development: Personality in the life course*, 281.

Cohn, M. A., Fredrickson, B. L., Brown, S. L., Mikels, J. A., & Conway, A. M. (2009). Happiness unpacked: positive emotions increase life satisfaction by building resilience. *Emotion*, *9*(3), 361. https://doi.org/10.1037/a0015952

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper and Row: New York.

Demerouti, E., Bakker, A., Nachreiner, F. & Schaufeli, W. (2001). The job demands resources model of burnout. *Journal of Applied Psychology, 86*, 499-512. https://doi.org/10.1037/0021-9010.86.3.499

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of personality assessment*, *49*(1), 71-75. https://doi.org/10.1207/s15327752jpa4901_13

Dietrich, J., Parker, P., & Salmela-Aro, K. (2012). Phase-adequate engagement at the post-school transition. *Developmental psychology*, *48*(6), 1575. https://doi.org/10.1037/a0030188

Diseth, Å., Danielsen, A. G., Samdal, O. (2012). A path analysis of basic need support, self-efficacy, achievement goals, life satisfaction and academic achievement level among secondary school students. *Educational Psychology: An International Journal of Experimental Educational Psychology, 32,* 335-354. https://doi.org/10.1080/01443410.2012.657159

Eccles, J.S. & Roeser, R.W. (2009). Schools, Academic Motivation, and Stage-Environment Fit. In R.M. Lerner and L. Steinberg (Eds.) *Handbook of Adolescent Psychology* (3rd ed.)(pp. 404-434). Hoboken, N.J.: John Wiley & Sons. https://doi.org/10.1002/9780470479193.adlpsy001013

Elder, G. H. & Shanahan, M. J. (2006). The life course and human development. In R. M. Lerner (Ed.). *Handbook of Child Psychology,* Vol. *1: Theoretical models of human development* (pp. 665-715). Editors-in-chief: W. Damon & R. M. Lerner. Hoboken, NJ: John Wiley & Sons.

Elmore, G. M. & Huebner, S. E. (2010). Adolescents' satisfaction with school experiences: Relationships with demographics, attachment relationships, and school engagement behavior. *Psychology in the Schools, 47,* 525-537. https://doi.org/10.1002/pits.20488

Erdogan, B., Bauer, T. N., Truxillo, D. M., & Mansfield, L. R. (2012). Whistle While You Work: A Review of the Life Satisfaction Literature. *Journal of Management*, *38*(4), 1038-1083. https://doi.org/10.1177/0149206311429379

Evans, K. (2007). Concepts of bounded agency in education, work, and the personal lives of young adults. *International journal of psychology*, *42*(2), 85-93. https://doi.org/10.1080/00207590600991237

Ford, M. T., Heinen, B. A., & Langkamer, K. L. (2007). Work and family satisfaction and conflict: A meta-analysis of cross-domain relations. *Journal of Applied Psychology, 92.* https://doi.org/10.1037/0021-9010.92.1.57

Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59-109. https://doi.org/10.3102/00346543074001059

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American psychologist*, *56*(3), 218. https://doi.org/10.1037/0003-066X.56.3.218

Haase, C. M., Heckhausen, J., & Köller, O. (2008). Goal engagement during the school–work transition: Beneficial for all, particularly for girls. *Journal of Research on Adolescence*, *18*(4), 671-698. https://doi.org/10.1111/j.1532-7795.2008.00576.x

Hakanen, J. J., Bakker, A. B., & Schaufeli, W. B. (2006). Burnout and work engagement among teachers. *Journal of school psychology*, *43*(6), 495-513. https://doi.org/10.1016/j.jsp.2005.11.001

Hakanen, J. J. & Schaufeli, W. B. (2012). Do burnout and work engagement predict depressive symptoms and life satisfaction? A three-wave seven year prospective study. *Journal of affective disorders*, *141*(2), 415-424. https://doi.org/10.1016/j.jad.2012.02.043

Harter, J. K., Schmidt, F. L., & Hayes, T. L. (2002). Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *Journal of Applied Psychology, 87,* 268-279. https://doi.org/10.1037/0021-9010.87.2.268

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Huebner, E. S., Drane, W., & Valois, R. F. (2000). Levels and demographic correlates of adolescent life satisfaction reports. *School Psychology International, 21,* 281-292. https://doi.org/10.1177/0143034300213005

Huebner, E. S., Valois, R. F., Paxton, R., & Drane, W. (2005). Middle school students' perceptions of quality of life. *Journal of Happiness Studies, 6,* 15-24. https://doi.org/10.1007/s10902-004-1170-x

Jones, M. D. (2006). Which is a better predictor of job performance: Job satisfaction or life satisfaction. *Journal of Behavioral and Applied Management*, *8*(1), 20-42.

Kuopion Lukiokoulutus (2009). *Tietoja Kuopion kaupungin lukioista LV. 2009 – 2010*. Retrieved from http://www.kuopio.fi/attachments.nsf/Files/031209121333543/$FILE/Tilastot%20Kuopion%20kaupungin%20lukioistalv2009_2010._versio3_lopullinen_ilman%20esitetta.pdf

Langelaan, S., Bakker, A. B., Van Doornen, L. J., & Schaufeli, W. B. (2006). Burnout and work engagement: Do individual differences make a difference? *Personality and Individual Differences*, *40*(3), 521-532. https://doi.org/10.1016/j.paid.2005.07.009

Lewis, A. D., Huebner, E. S., Malone, P. S., & Valois, R. F. (2011). Life satisfaction and student engagement in adolescents. *Journal of Youth and Adolescence, 40,* 249-262. https://doi.org/10.1007/s10964-010-9517-6

Li, Y. & Lerner, R. M. (2011). Trajectories of school engagement during adolescence: Implications for grades, depression, delinquency, and substance use. *Developmental Pscyhology, 47,* 233-247. https://doi.org/10.1037/a0021307

Linnakylä, P. & Malin, A. (2008). Finnish students' school engagement profiles in the light of PISA 2003. *Scandinavian Journal of Educational Research, 52,* 583-602. https://doi.org/10.1080/00313830802497174

Louis, V. V., & Zhao, S. (2002). Effects of family structure, family SES, and adulthood experiences on life satisfaction. *Journal of Family Issues*, *23*(8), 986-1005. https://doi.org/10.1177/019251302237300

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal, 37,* 153-184. https://doi.org/10.3102/00028312037001153

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*(1), 397-422. https://doi.org/10.1146/annurev.psych.52.1.397

Masten, A. S., Desjardins, C. D., McCormick, C. M., Kuo, Sally I-Chun, & Long, J. D. (2010). The significance of childhood competence and problems for adult success in work: A developmental cascade analysis *Development and Psychopatology, 22,* 679-694.

Moazami-Goodarzi, A., Nurmi, J.-E., Mauno, S. & Rantanen, J. (2014). Cross-lagged relations between work-family enrichment, vigor at work, and core self-evaluations: A three-wave study. *Journal of Business and Psychology,* 1-10.

Murray, C. (2009). Parent and teacher relationships as predictors of school engagement and functioning among low-income urban youth. *Journal of Early Adolescence, 29,* 376-404. https://doi.org/10.1177/0272431608322940

Muthén, L., & Muthén, B. O. (1998-2016af). *Mplus. User's guide.* Los Angeles, CA: Muthén & Muthén.

Myers, D. G., & Diener, E. (1995). Who is happy? *Psychological science*, *6*(1), 10-19. https://doi.org/10.1111/j.1467-9280.1995.tb00298.x

Ou, S. R. (2008). Do GED recipients differ from graduates and school dropouts? Findings from an inner-city cohort. *Urban Education*, *43*(1), 83-117. https://doi.org/10.1177/0042085907305187

Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological assessment*, *5*(2), 164. https://doi.org/10.1037/1040-3590.5.2.164

Perdue, N. H., Manzeske, D. P., & Estell, D. B. (2009). Early predictors of school engagement: Exploring the role of peer relationships. *Psychology in the Schools, 46,* 1084-1097. https://doi.org/10.1002/pits.20446

Pomerantz, E. M., Altermatt, E. R., & Saxon, J. L. (2002). Making the grade but feeling distressed: Gender differences in academic performance and internal distress. *Journal of Educational Psychology, 94,* 396-404. https://doi.org/10.1037/0022-0663.94.2.396

Rindfuss, R. R. (1991). The young adult years: Diversity, structural change, and fertility. *Demography*, *28*(4), 493-512. https://doi.org/10.2307/2061419

Salanova, M., Agut, S., & Peiro, J. M. (2005). Linking organizational resources and work engagement to employee performance and customer loyalty: The mediation of service climate. *Journal of Applied Psychology, 90,* 1217-1227. https://doi.org/10.1037/0021-9010.90.6.1217

Salmela-Aro, K., Kiuru, N., & Nurmi, J.-E. (2008). The role of educational track in adolescents' school burnout: A longitudinal study. *British Journal of Educational Psychology, 78,*   663-689. https://doi.org/10.1348/000709908X281628

Salmela-Aro, K., Savolainen, H., & Holopainen, L. (2009). Depressive symptoms and school burnout during adolescence: Evidence from two cross-lagged longitudinal studies. *Journal of Youth and Adolescence*, *38*(10), 1316-1327. https://doi.org/10.1007/s10964-008-9334-3

Salmela-Aro, K. & Upadyaya, K. (2012). The Schoolwork Engagement Inventory: Energy, Dedication and Absorption (EDA). *European Journal of Psychological Assessment, 28,* 60-67. https://doi.org/10.1027/1015-5759/a000091

Salmela-Aro, K. & Upadyaya, K. (2014). School Burnout and Engagement in the Context of the Demands-Resources Model. *British Journal of Educational Psychology, 84,* 137-151. https://doi.org/10.1111/bjep.12018

Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006). The measurement of work engagement with a short questionnaire: A cross-national study. *Educational and Psychological Measurement, 66,* 701-716. https://doi.org/10.1177/0013164405282471

Schaufeli, W. B., Martinez, I., Pinto, A. M., Salanova, M., & Bakker, A. (2002). Burnout and engagement in university students: A cross-national study. *Journal of Cross-Cultural Psychology, 33,* 5, 464-481. https://doi.org/10.1177/0022022102033005003

Schaufeli, W. B., Salanova, M., Gonzalez-Roma, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies, 3,* 71-92. https://doi.org/10.1023/A:1015630930326

School Statistics. (2010). http://www.stat.fi/til/khak/2010/khak_2010_2011-12-13_tau_001_fi.html. Helsinki: Central Statistical Office of Finland.

Seligman, M. E. P. & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55,* 5-14. https://doi.org/10.1037/0003-066X.55.1.5

Seligson, J. L., Huebner, E. S., & Valois, R. F. (2003). Preliminary validation of the brief multidimensional students' life satisfaction scale (BMSLSS). *Social Indicators Research*, *61*(2), 121-145. https://doi.org/10.1023/A:1021326822957

Seppälä, P., Mauno, S., Feldt, T., Hakanen, J., Kinnunen, U., Tolvanen, A., & Schaufeli, W. (2009). The construct validity of the Utrecht Work Engagement Scale: Multisample and longitudinal evidence. *Journal of Happiness Studies*, *10*(4), 459-481. https://doi.org/10.1007/s10902-008-9100-y

Simons-Morton, B. G. & Crump, A. D. (2003). Association of parental involvement and social competence with school adjustment and engagement among sixth graders. *Journal of School Health, 73,* 121-126. https://doi.org/10.1111/j.1746-1561.2003.tb03586.x

Suldo, S. M., & Huebner, E. S. (2006). Is extremely high life satisfaction during adolescence advantageous?. *Social indicators research*, *78*(2), 179-203. https://doi.org/10.1007/s11205-005-8208-2

Suldo, S. M., Riley, K. N., & Shaffer, E. J. (2006). Academic correlates of children and adolescents' life satisfaction. *School Psychology International*, *27*(5), 567-582. https://doi.org/10.1177/0143034306073411

Upadyaya, K. Salmela-Aro, K. (2013a). Development of school engagement in association with academic success and well-being in varying social contexts: A review of empirical research. *European Psychologist, 18,* 136-147*.* https://doi.org/10.1027/1016-9040/a000143

Upadyaya, K. Salmela-Aro, K. (2013b). Engagement with studies and work: Trajectories from post-comprehensive school education to higher education and work. *Emerging Adulthood, 1*, 247-257*.* https://doi.org/10.1177/2167696813484299

Upadyaya, K., & Salmela-Aro, K. (2014). Cross-lagged associations between study and work engagement dimensions during young adulthood. *The Journal of Positive Psychology*, (ahead-of-print), 1-13.

Upadyaya, K., Vartiainen, M. & Salmela-Aro, K. (2015). From servant leadership to work engagement, life satisfaction, and occupational health: Job demands and resources. *Burnout Research*, *3*(4), 101-108. https://doi.org/10.1016/j.burn.2016.10.001

Upadyaya, K., & Salmela-Aro, K. (*in press*). Development of early vocational behavior: Parallel associations between career engagement and satisfaction. *Journal of Vocational Behavior.* https://doi.org/10.1016/j.jvb.2015.07.008

Wefald, A. J., & Downey, R. G. (2009). Construct dimensionality of engagement and its relation with satisfaction. *The Journal of Psychology*, *143*(1), 91-112. https://doi.org/10.3200/JRLP.143.1.91-112

Table 1. Means, variances, Pearson Correlation Coefficients, and Cronbach's Alpha Reliabilities.

	SWE1	SWE 2	SWE 3	SWE 4	SWE 5	Life 1	Life 2	Life 3	Life 4	Life 5	Gender	GPA	Track	SES
SWE 1														
SWE 2	.58***													
SWE 3	.29***	.36***												
SWE 4	.31***	.30***	.45***											
SWE 5	.28***	.28***	.34***	.45***										
Life 1	.38***	.30***	.19***	.26***	.20***									
Life 2	.27***	.33***	.21***	.24***	.18***	.64***								
Life 3	.14**	.16**	.30***	.27***	.23***	.40***	.52***							
Life 4	.17***	.19***	.25***	.43***	.29***	.40***	.47***	.61***						
Life 5	.24***	.22***	.25***	.34***	.36***	.33***	.35***	.42***	.57***					
Gender	.03	-.02	-.05	.02	-.01	.14***	.14***	.04	.03	-.08**				
GPA	.23***	.25***	0.01	.05	.04	.28***	.24***	.18***	.17**	.09	-.30***			
Track	-.11	-.21*	-.11*	.03	.07	-.05	.01	.03	.04	.02	-.05	.60***		
SES	-.06	-.01	-.05	-.02	-.05	-.02	-.02	.08	.05	-.02	-.02	.20***	.22***	
M	3.56	3.46	4.33	4.21	4.38	4.77	4.80	4.61	4.75	4.39	1.54	7.89	0.55	1.70
Var	1.74	1.74	1.26	1.44	1.40	1.62	1.56	1.63	1.60	1.82	0.27	0.75	0.26	0.15
Cronbach's α	.94	.94	.93	.94	.94	.89	.89	.87	.87	.89	-	-	-	-

Note. ***$p < .001$; **$p < .01$; *$p < .05$

Note. SWE = study/work engagement; Life = life satisfaction; 1 = Time 1 (age = 17, $N = 818$); 2 = Time 2 (age = 18, $N = 749$); 3 = Time 3 (age = 21, $N = 611$); 4 = Time 4 (age = 23, $N = 561$); 5 = Time 5 (age = 25, $N = 533$).
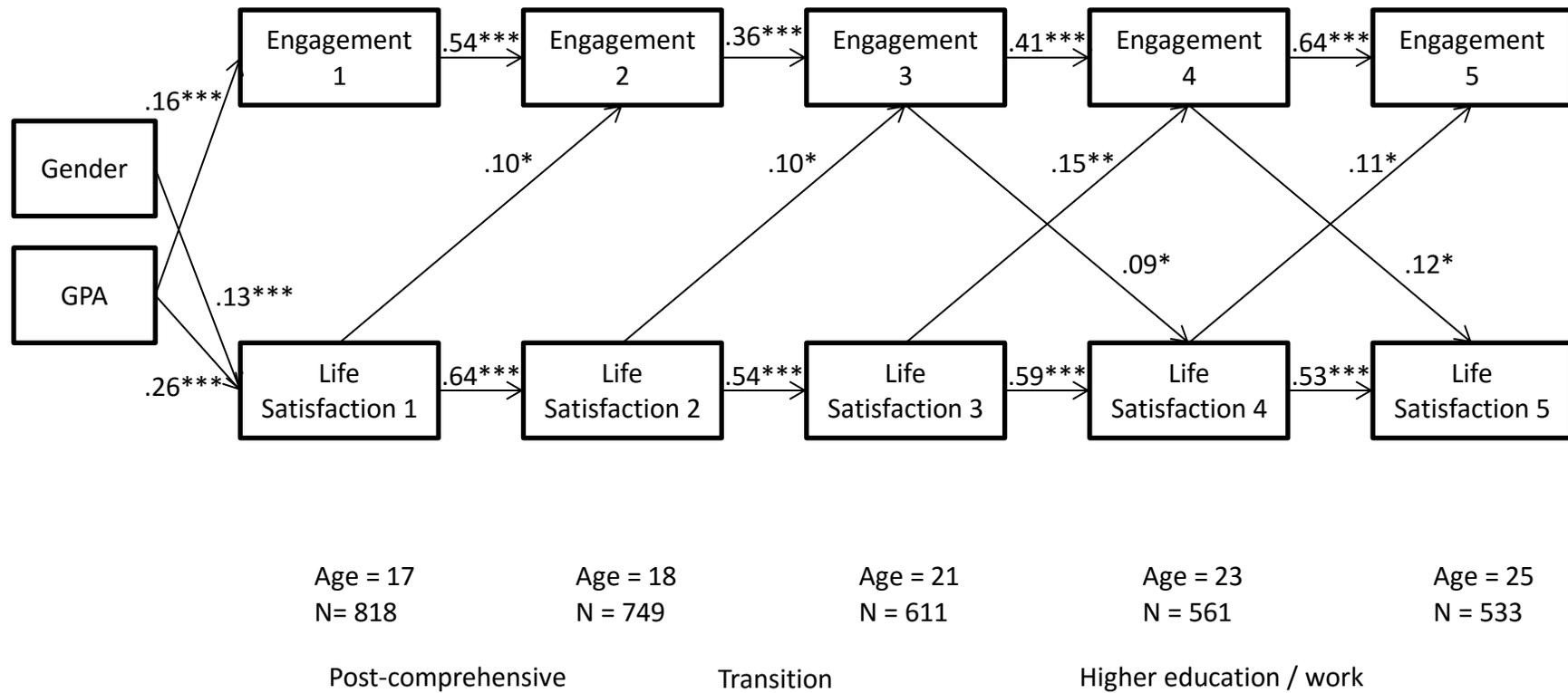
Figure 1. Cross-lagged associations between young adults' life satisfaction and engagement in studies/work (Standardised estimates). Note. ***p < .001; **p < .01; *p <.05; 1 = Time 1, 2 = Time 2; 3 = Time 3; 4 = Time 4; 5 = Time 5.

A socio-ecological model of agency: The role of structure and agency in shaping education and employment transitions in England

Ingrid Schoon   UCL Institute of Education, UK, Wissenschaftszentrum, Berlin, Germany
i.schoon@ucl.ac.uk
Mark Lyons-Amos   London School of Hygiene and Tropical Medicine, UK

Abstract

This study examines the role of structural and agentic resources in shaping school-to-work transitions in England. We ask to what extent are young people able to steer the course of their lives despite the constraining forces of social structure, and how satisfied are they with their lives following the completion of compulsory schooling. Drawing on data from the Longitudinal Study of Young People in England we use sequence analysis of monthly activity data to identify differences in the timing and sequencing of education and employment transitions. We identified six distinct pathways, differentiating between an academic track, three pathways involving further education and training, as well as a work-focused transition and a group of young people who were over a long period not in education or training (NEET). The findings suggest that not all young people are inclined to follow an academic track and instead select into pathways involving vocational training or further education, enabling them to experience competence and life satisfaction. For others (about one in 10), however, the lack of socioeconomic and psycho-social resources is too overwhelming and they encounter long-term experience of NEET or are not able to transform their educational credentials into employment opportunities. The findings highlight that in addition to considering structural constraints it is important to conceptualise the role of the agent for a better understanding of variations in youth transitions

.

Introduction

The transition from school to work is an important developmental task for young people entering the third decade of life, and ranks very high in terms of complexity and relevance for later life outcomes (Buchmann & Kriesi, 2011; Schulenberg & Schoon, 2012). There is however still a lack of understanding regarding the distinct pathways young people take when making the transition, the timing and sequencing of events, and the resources available to young people at the start of the 3rd decade. Recent debates have focused on the notion of 'emerging adulthood' (Arnett, 2000) used to describe the generally extended period of education participation and delay in starting paid employment (often until the mid or late 20s). Although, on aggregate, the transition to independent adulthood has been delayed across many Western countries since the 1970s, not all young people participate in higher education. For example, in 2010 the OECD average percentage of 25-34 year olds to have completed tertiary education was just below 40% (OECD, 2012), and in 2012 the participation rate in higher education among 17 to 30 year olds in the UK was 49% (BIS, 2013). A large group of young people thus do not go to University. How do they navigate the transition to employment? What is the role of structural and agentic factors in influencing variations in transition pathways? And how do the young people themselves evaluate their lives?

Previous research has shown that the transition to adulthood has been prolonged for those with the socioeconomic resources to invest in their education, leading to a polarisation of life chances (Jones, 2002; McLanahan, 2004). Yet, the assumption of a polarised life course does not take into account the experiences of a 'forgotten middle' (Roberts, 2011), those who do not necessarily go to university and who negotiate their lives by balancing the socioeconomic and psycho-social resources available to them (Schoon, 2015a; Schoon & Lyons-Amos, 2016). Empirical studies testing the complex interplay between individual and structural forces in shaping youth transitions are, however, scarce.

In this study we introduce a socio-ecological model of agency to address the issue of person-environment interactions. It is argued that to some extent young people are able to steer the course of their lives despite constraining forces of social structure. The socio-ecological approach aims to understand how individuals and social ecologies define each other (Oishi, 2014). It is informed by life course theory with its emphasis on the multiple sources of influence on individual development, ranging from the micro- to the macro context (Elder, 1998; Elder & Shanahan, 2007), Bandura's social cognitive theory of human action (Bandura, 2001, 2006), and Eccles et al.'s (1993) person-environment fit theory – examing interactions between social structure and agency.

We define social structure as an external, objective force that can influence individual feelings and actions (McLeod & Lively, 2003), focusing in particular on family socioeconomic resources associated with different social structural positions[i]. We also take into account the wider social context, in particular area characteristics that are understood to shape education and employment opportunities. Individual agency is defined as the capacity to transcend the immediate constraints in one's environment and to shape one's life circumstances and the courses that one's lives take (Bandura, 2001). It is conceptualised as a multi-dimensional construct involving indicators of intentionality, forethought, self-directedness, and self-efficacy. Our socio-ecological model of agency includes the measurement of a. agency across multiple dimensions; b. social structures that affect goal pursuit; c. the wider social contexts that shape transition pathways; and d. overall subjective evaluation of one's life. We adopt an explicit developmental approach, taking into account that any point in the life course has to be understood as the consequence of past experiences, and the launch pad of subsequent experiences. We use the model to predict variations in the timing and sequencing of education and employment transitions, following a nationally representative sample of the Longitudinal Study of Young People in England (LSYPE) born in 1989/90 from age 13 to 20. In our analysis we control for differences in academic ability to focus on the role of agentic processes by which differentiation in transition outcomes occur.

A socio-ecological model of agency

Life course theory provides a comprehensive framework to integrate assumptions about social structure and individual agency. In life course theory it is argued that all life choices are contingent on the opportunities and constraints of social structure and culture (Elder & Shananhan, 2007). Moreover, life course theory considers the constraints on human

development posed by social norms and institutions. Pathways through life are embedded within a larger socio-historical and cultural context, and are shaped by complex interdependent relationships, including links to one's family of origin (Elder, 1998). Some individuals are able to select the path they follow, a phenomenon described as human agency, but these choices are not made in a social vacuum. Individual agency is shaped by opportunity structures, social networks and institutions, and it is argued that individuals often unconsciously reproduce their social structural milieu (Bourdieu & Passeron, 1977; Giddens, 1991; Hitlin & Elder, 2007).

Although the concept of agency is a central term in life course theory (Elder, 1998) it has remained an elusive ('slippery') and underspecified theoretical concept (Hitlin & Elder, 2007). As a non-structural factor agency is not universally accepted or valued in sociological theory (Fuchs, 2001; Loyal & Barnes, 2001). Yet, there is a growing interest within sociology to render the mainly theoretical notion of agency into a tangible concept for empirical research. In psychology, on the other hand, agency is a central empirical construct in motivational theories, yet mechanisms linking structural factors to action are left largely unexplained (Bandura, 2006).

Recent attempts to reconceptualise agency within life course research draw on the psychological literature of agency (Hitlin & Elder, 2007; Hitlin & Johnson, 2015), in particular Bandura's social cognitive theory (Bandura, 2001). Social cognitive theory rejects the duality between individual agency and social structure, and assumes that people create social systems, and these systems, in turn, organise and influence individual lives. Likewise in sociology the reciprocal interactions between social structure and the individual are recognised (Emirbayer & Mische, 1998). Yet while in psychology interactions between person and social structure are often used to explain the human capacity to shape one's life circumstances and the course of one's life (Eccles, 2008; Heckhausen & Chang, 2009; Lerner, 1996; Schoon, 2007), in sociology the focus lies on explaining stratification of life course outcomes (Kerckhoff, 2001; Sewell & Hauser, 1993). There is however overlap in the research interests and a growing body of evidence regarding the interplay between structure and agency from across disciplines. A socio-ecological approach to agency enables us to investigate how objective socioeconomic conditions affect individual thinking, feeling and behaviour, and how different aspects of agency might shape the selection of distinct transition pathways, which can be understood as ecological niches.

Structural influences on agency

Young people from less privileged backgrounds are leaving education earlier and are less likely to continue in higher education than their more privileged peers, reflecting persistent social inequalities in life chances and opportunities (Erikson & Jonsson, 1996; Furlong & Cartmel, 1997; Schoon & Silbereisen, 2009; Settersten, Furstenberg, & Rumbaut, 2005). There is also evidence to suggest that children born into less privileged families (characterised by low levels of parental education, low income, unemployment, single or early parenthood, poor housing conditions) show, in general, lower levels of educational attainment (Breen & Goldthorpe, 2001; Bukodi & Goldthorpe, 2013; Shavit, Arum, & Gamoran, 2007), self-confidence and educational achievement motivation (Duckworth & Schoon, 2012; Eccles, 2008; Mortimer, 2003; Schoon, 2014). Explanations of these associations refer to cumulative risk effects (DiPrete & Eirich, 2006), the lack of financial resources, time or energy of parents to invest in the education of their children (Guo & Harris, 2000), familiarity with the dominant culture, social networks and connections, or access to warm and supportive parenting (Conger, Conger, & Martin, 2010).

Moving beyond the proximal family context, there is also evidence that the wider social context, indicated for example by area characteristics, can play a significant role in shaping young people's lives. Living in a disadvantaged neighbourhood (characterised by high levels of unemployment, crime, and lack of community resources), especially in urban areas, has been associated with lower levels of achievement orientation and academic attainment (Ainsworth, 2002; Brooks-Gunn, Duncan, & Aber, 1997; Murry, Berkel, Gaylord-Harden, Copeland-Linder, & Nation, 2011) , as well as a higher risk of precarious employment transitions, such as prolonged experiences of not being in employment, education or training (NEET) (Schoon, 2014). Explanations of these associations point to the role of institutional resources, security and community services, as well as collective socialisation models, i.e. the monitoring function that adults adopt to

control and monitor behaviour (Ainsworth, 2002; Ioannides & Loury, 2004; Jencks & Mayer, 1990).

Socioeconomic risks rarely occur in isolation, and are interlinked. For example, family poverty is increasingly concentrated in certain subgroups of the population and in relatively disadvantaged areas (Duncan & Brooks-Gunn, 1997; Gregg & Wadsworth, 2001). Moreover, the relationship between any single risk factor and subsequent outcomes tends to be weak, and serious risk emanates from the accumulation of risk factors (Dannefer, 2003; Rutter, 1981; Schoon, 2006). In an increasingly competitive labour market, it is those with multiple disadvantages who are likely to face the greatest difficulties in establishing themselves (Scarpetta, Sonnet, & Manfredi, 2010). We thus consider the impact of multiple socioeconomic risk factors in order to a. provide a more comprehensive description of the living situation of young people today, and b. to more accurately predict and understand developmental outcomes. We focus in particular on the role of parental education and employment status, housing conditions, as well as family structure – all of which reflect the socioeconomic resources that shape everyday experiences (Moore, Vandivere, & Redd, 2006; Schoon, 2006), and which have shown independent associations with youth transitions (Duckworth & Schoon, 2012).

Conceptualising agency as a multi-dimensional construct

Individuals are not passively exposed to influences from the immediate or wider social context. According to the agency principle they are able to shape the context, which in turn shapes them, and agency can potentially compensate for family disadvantage (Elder & Shanahan, 2007). Agency is conceptualised as a multi-dimensional construct, involving orientations to the past, the present and the future (Bandura, 2001; Emirbayer & Mische, 1998; Hitlin & Elder, 2007). It is informed by past behaviour and experience, reflections about one's capabilities within given constraints and opportunities, and orientations to the future (i.e. expression of future projects, goals, and intentions), which imply expectations that actions taken will be successful. For example, Bandura (2001, p. 1) specifies that the capacity to shape the course of one's life depends on "the temporal extension of agency through intentionality and forethought, self-regulation by self-reactive influence, and self-

reflectiveness about one's capabilities, quality of functioning and the meaning and purpose of one's life pursuits". Emirbayer and Mische (1998, p. 962) define agency as "informed by the past (in its "irrational" or habitual aspect) but also oriented towards the future (as a "projective" capacity to imagine alternative possibilities) and toward the present (as a "practical-evaluative" capacity to contextualize the past habits and future projects within the contingencies of the moment)."

The notions of forethought and self-efficacy (or mastery) have been used in a study by Hitlin and Johnson (2015) to empirically operationalise agency within a life course approach. They showed that a general measure of future orientation (perceived life chances) and self-reflection, i.e. perception of own capabilities, predict a range of outcomes in young adulthood, i.e. income, financial problems, self-rated health and depressive effect, independent of academic attainment and a range of socioeconomic background factors (parental education, social class, income, family structure, gender and ethnicity). Moreover, they demonstrate that future orientation and mastery are distinct factors showing independent contributions, and that both general (perceived life chances) and specific future orientations (education aspirations) play a role in shaping developmental outcomes, in particular regarding socioeconomic attainment (earnings and financial problems).

The study by Hitlin and Johnson is one of the few to operationalise agency as a multi-dimensional construct, yet it does not take into account the role of intention and self-directedness, which will be included in our model. Intentions reflect plans of action that imply a pro-active commitment to a given purpose or goal. Self-directedness refers to formative influences of agentic processes based on individual preferences and values. Emirbayer and Mische (1998) argue that both the projective and practical-evaluative aspects of agency are grounded in habitual (often unconscious) patterns of action, which are sometimes also referred to as dispositions, competencies, and preferences. These preferences give stability to response tendencies and help to sustain identity. According to Bandura (2001) plans for the future are rooted in a value system and a sense of personal identity (see also (Eccles & Wigfield, 2002). People do things that give them self-satisfaction and refrain from actions that give rise to self-dissatisfaction. They regulate their behaviour

through a set of self-referenced subfunctions, which include self-monitoring and self-guidance via personal standards.

There is ample empirical evidence to confirm the importance of intentions or goal directedness, self-efficacy, forethought, and preferences or values as predictors across a range of outcomes, including socioeconomic attainment (i.e. occupational and social status or income), psychological wellbeing, and health, even after controlling for academic attainment and structural constraints (e.g. Ashby & Schoon, 2010; Bandura, Barbaranelli, Caprara, & Pastorelli, 2001; Eccles, 2008; Hitlin & Johnson, 2015; Schoon, 2008). The role of these psycho-social resources as predictors of transition experiences in combination with socioeconomic constraints is however less well studied. Most previous studies have established positive associations between distinct agentic aspects and later outcomes, yet there is a lack of studies that have comprehensively examined the complex interplay of multiple agentic and structural factors in shaping the transition from school to work, and there is little understanding of the mechanisms linking agency to variations in transition experiences.

The transition from school to work

The transition from school to work is a rather busy juncture of the life course that involves multiple and inter-related social role changes, including completion of full-time education and entry into paid employment, and making the step into family formation and parenthood (Settersten, 2007). Each of these role transitions brings with it new challenges and opportunities that rank very high in terms of their importance, complexity and relevance for later outcomes (Buchmann & Kriesi, 2011; Schulenberg & Schoon, 2012; Shanahan, 2000).

Life course theory emphasises the importance of timing and sequencing of events in determining their meaning and implications (Elder, 1998). It recognises that transitions can have different meanings, antecedents, and consequences depending on when they occur in the life course and how they fit into larger sequences or trajectories. Within a given society, the timing and sequencing of transitions are governed by a set of institutionalised, normative timetables (Elder, 1998) or 'scripts' of life (Buchmann, 1989) that provide models for both role behaviour as well as informal expectations regarding the age and timing of major transitions. Normative,

or 'on-time transitions' are 'culturally prepared' by socialisation and institutional arrangements and are understood to be psychologically salutary. Those who are 'off-time,' i.e. too early or too late, are thought to be the target of negative social sanctions and to experience psychological strain (Heckhausen, 1999; Salmela-Aro, 2009). For example, early transitions (such as early school leaving) and problems in establishing oneself in the labour market, have been associated with lower levels of life satisfaction and health (Kins & Beyers, 2010; Schulenberg, Bryant, & O'Malley, 2004).

Developmental match and wellbeing

Yet, each transition can demarcate a turning point that is associated with change for the better or worse. A number of recent studies across different countries found that there is not one normative way to negotiate a successful transition to adulthood (Schoon, 2015a; Schoon & Lyons-Amos, 2016; Schulenberg & Schoon, 2012). There is heterogeneity in transition experiences: early transitions do not necessarily have a negative outcome, and protracted pathways to adulthood are not necessarily optimal. According to 'developmental match/mismatch models' (Eccles et al., 1993; Schulenberg et al., 2004), transitions that provide a progressive increase in developmentally appropriate challenges through which young people can experience competence, enable the individual to successfully master the transition. Building on person-environment fit theory (Eccles et al., 1993), the developing individual is conceptualised as being embedded in changing ecological niches, where the match between individual developmental needs and opportunities provided by the context is itself a dynamic process. Individuals actively choose or create opportunities (within given constraints) that provide a better fit to their preferences, self-perceptions and intentions. For example, for a young person not enjoying school or academic study, entry into employment can provide the opportunity to feel valued, to belong, and to make a contribution. If, however, the demands of the developmental transitions are not matched to the capabilities of the individual or if they amplify previous difficulties, then there can be a negative effect on mental health and wellbeing. This can, for example, be the case if a young person from a less privileged background with the desire to make a living without obtaining higher qualifications is not able to establish herself

in the labour market, or experiences long-term unemployment.

We thus have to know more about how young people themselves evaluate their transition experiences, how this evaluation is influenced through the available socioeconomic and psycho-social resources, and their match to subsequent transition experiences. We therefore assess individuals' subjective evaluation of their lives by age 19/20, and how this is predicted by a. socioeconomic resources and area characteristics; b. individual agency factors; and c. variations in transition experiences. Although there is stability in reports of subjective life evaluations over time and across situation, there can be changes as circumstances in life are changing (Diener, Inglehart, & Tay, 2013). This is especially relevant during the transition from school to work. We thus examine subjective life evaluations before the transitions were made as an additional control variable. In this way we take into account possible selection effects.

Research questions

This study examines the role of structure and agency in shaping youth transitions. We test assumptions based on the notion of cumulative disadvantage and developmental-person-environment matching, taking into account main effects and interactions between indicators of structure and agency.

1. According to assumptions of cumulative disadvantage we expect diversity in transition patterns, reflecting structural constraints on life chances and opportunities starting early in life. For example, we expect young people with fewer socioeconomic resources to have lower levels of agency and to leave school earlier than their more privileged peers.

2. According to a socio-ecological model of agency, we expect that individuals select into a distinct developmental niche that corresponds to their intentions, self-perceptions and preferences within given structural constraints. In particular, we test three distinct processes by which agency processes interact with structural constraints. First, following the assumption of cumulative risk we expect that young people from less privileged background report lower level agency and that agency has a stronger impact among relatively privileged young people. Second, according to an independent effects model we

assume that agency indicators predict transition outcomes independent of structural constraints, i.e. there is no interaction and both structure and agency have a unique contribution to transition experiences. Third, according to a compensatory, or interactive effect model, we expect that agency has a more beneficial effect for young people with fewer socioeconomic resources. This would imply, for example, that young people from less privileged backgrounds who feel confident about their academic capabilities and who want to stay on in education are more likely to continue in education, at least for a couple years, than their peers in similar socioeconomic circumstances who do not enjoy school, and who do not want to continue in education.

3. How do young people evaluate their lives at the beginning of their third decade? We expect that the closer the match between the realised transition and earlier self-perceptions, preferences and intentions, the higher the level of life satisfaction. According to assumptions of developmental person-environment matching we expect lower levels of life satisfaction where transitions do not meet one's intentions (i.e. the anticipation to attend university is not realised), or where transition experiences amplify previous adversities (i.e. low socioeconomic and low agentic resources).

Method
Data

The study is based on data collected for the Longitudinal Study of Young People in England (LSYPE). LSYPE is a panel study of 15,770 young people born between 1st September 1989 and 31st August 1990. Sample members were all young people in school year nine (age 13/14) or equivalent, in England in February 2004 (for more details see https://www.education.gov.uk/ilsype/workspaces/public/wiki/Welcome).

Annual face-to-face interviews have been conducted with young people and their parents between 2004 and 2010, and linkage is available to other administrative data, such as the National Pupil Database (NPD), which includes national assessments for all children in England. From LSYPE, information from wave 1 to wave 7 of the dataset was used, covering ages 13/14 to 19/20 years. From NPD, a national assessment given at age 11 is used

as an indicator of previous academic performance, which is understood to shape subsequent indicators of agency and developmental outcomes.

The LSYPE was sampled using a probability proportional to size method, using schools as the primary sampling unit. It was additionally stratified on deprivation levels of those schools, oversampling more deprived schools and oversampling pupils from minority ethnic groups. The initial sample size was 15,770 partial responses (data from young person) and 13,914 full responses (young person and parent) although not all young people provided information for all waves of the survey. The wave 7 sample consisted of all young people who had been interviewed at previous waves and who agreed to be re-contacted. In total 9,791 cases were contacted at wave 7 in 2010.

Analytic sample

The analytic sample used for this study comprises individuals with information on their family background at age 13/14 and who participated in the last wave 7 at age 19/20, comprising 9,558 individuals (4,825 males and 4,733 females). The sample is largely representative of the original sample, although there is some greater socioeconomic disadvantage among young people who did not continue in the study. Special sample weights, which are calculated and available from the LSYPE website, were applied to account for the study design, differential selection probabilities and non-response bias.

Measures

Agency

Since our focus is on education to work transitions we used four domain specific indicators, all assessed at age 13/14 (wave 1).

Academic expectations (intention). The young people were asked how likely it is that they will ever apply to go to university to do a degree. Responses were coded on a five-point scale with response options 5=very likely, 4=likely, 3=do not know, 2=not likely, and 1 =not at all likely.

Goal certainty (forethought): The young people were also asked how likely they think it is that if they do apply to go to university that they will get in. Responses were coded on a 5-point scale with response options 5=very likely, 4=likely, 3=do not know, 2=not likely, and 1 =not at all likely.

Academic self-concepts (self-efficacy). Perceived efficacy to master different academic subjects was measured by asking the young people how good they would say they are in math, English, science and Information/Communication/Technology (ICT). Responses were coded on a four-point scale with response options 4=very good, 3=fairly good, 2=not very good, and 1 =no good at all. The items were summed up to create an index of academic self-efficacy (Bandura et al., 2001). The scale score was z-standardised. A high score indicates high and a low score low levels of efficacy.

School engagement (self-directedness). We used indicators of emotional school engagement as a marker of student's attitudes and values reflecting self-directedness (Fredricks, Blumenfeld, & Paris, 2004). A scale score was created based on summed answers to five attitudinal questions: I am happy at school; school work is worth doing; I work as hard as I can at school; I am bored in lessons; on the whole I like being at school. The items were measured on a four point Likert scale ranging from strongly agree to strongly disagree. The scale has good internal consistency (alpha= .73). The summary scale score was z-standardised, and a high score indicates positive school motivation and a low score school disengagement.

Structural factors

In our assessment of structural factors we focus on a range of socioeconomic resources available to the family at wave 1:

Parental education. Information on mother's and father's highest educational level were gathered at wave 1 using the National Vocational Qualification (NVQ) levels. For our analysis we identified the highest level of either parent, using the dominance approach (Erikson, 1984). We differentiated parents with relatively low levels of education (0= no qualifications or qualifications up to level 2, equivalent to seven GCSEs at grades A to C) and those with higher level of education (1= qualifications at level 3 which enables access to University and higher, i.e. degree level qualifications). *Gross household income* was reported by the main parent. The banded information was dichotomised to differentiate between those in the lowest income group (1=less than £10,400 per annum) against others (0). *Parental worklessness* was assessed at the household level (not the individual level). This variable was coded as 1 if no parent living in the household was working at

the time the family was interviewed (comprising those who were looking for work, as well as those who were economically inactive, not looking for work because of health problems, disability, or looking after the family) and 0 if at least one parent was working. Single *parent family.* This variable is coded as 1 if the young person lives in a single parent family and 0 if two parents are present.   *Teen parent*. This variable is coded 1 if the cohort member was born to a teen mother and 0 otherwise. *Home ownership* in wave 1 is coded as 1 if the family owns their own home and 0 if they are renting. We created a cumulative socioeconomic risk index by adding up the six dichotomised indicators.

Area characteristics

The *Index of Multiple Deprivation (IMD)* was measured at wave 1 to provide a relative measure of deprivation at the small area level across England. The IMD is made up of seven constituent domains comprising income, employment, education, crime, health deprivation and disability; barriers to housing and services deprivation; and living environment deprivation (for more details see http://data.gov.uk/dataset/index-of-multiple-deprivation). Areas are ranked from least deprived to most deprived, on an overall composite measure of multiple deprivation. Another source of geographic information is the urban/rural classification in LSYPE, a measure developed by the Department for Environment Food and Rural Affairs (ONS, 2013). Rurality of an area was coded as 0, contrasting it to urban areas or towns coded as 1. Information on rurality and multiple area deprivation provide important contextual information regarding the communities that study members are growing up in.

School-work-transitions

We used monthly activity history data collected as a routine part of survey between ages 16 to 20 (October 2006 and May 2010) comprising information on being in full-time education, employed (part- or full-time), in an apprenticeship or government training, or being out of the labour force (not in education, employment or training (NEET)).

Subjective life evaluations

At wave 2 (age 14/15) students were asked whether they have been feeling reasonably happy, all things considered. And at wave 7 (age 19/20) they were asked whether they were satisfied with their life so far. Although the items are differently worded, they both tap into general evaluations of one's life. Both items are coded on a five-point response scale indicating low (1) versus high (5) levels of subjective life satisfaction or happiness.

Controls

Because both agency and transition experiences are shaped by socio-demographic factors and prior academic attainment, we include additional control variables in our model to ensure that the estimated effects of agency do not simply reflect spurious relationships.

The adolescents reported their *gender* (0=male) (1=female) and *ethnicity*. Ethnicity was coded as (0) white, versus (1) other ethnic groups. Given the ethnic diversity in England, the different ethnic groups were too numerous and the number of each group was too small to examine differences among the groups individually in our model. *Academic performance at age 11* was measured using a latent variable comprising maths, english and science scores in national curriculum tests given at the end of Key Stage 2 (i.e., age 11) [ii].

Statistical analysis

All analyses were carried out using the software package STATA14. We first provide descriptive statistics and correlation tables. To identify patterns in the timing and sequencing of education and employment transitions between 2006 and 2010 (ages 16 to 21) we used sequence analysis. We apply the sequence analysis for Stata ado (Brinzksy-Fay et al. 2006), using the Needleman-Wunsch algorithm to perform optimised matching, with InDel costs set to 0.7, substitution costs set to 0.5 and identical between states to create a distance matrix to indicate the differences between pairs of sequences. We then use the distance matrix to perform a cluster analysis using hierarchical clustering (Ward's distance clustering) with the number of clusters determined by the Duda-Hart statistic. Multinomial logistic regression analysis is used to determine the influence of structural and agentic factors as predictors of cluster membership. Finally we use stepwise OLS regression to predict life satisfaction at age 19/20. We used multiple imputation (mi impute command in STATA) to check robustness of findings. Given the consistency in findings we report coefficients derived from the full sample without imputation.

Table 1: Bivariate correlations between predictor variables included in model with means and std (n=9558)

	1	2	3	4	5	6	7	8	9	10	11	12	Mean	Std
Socioeconomic resources														
1.Family resources	1.00												1.00	1.32
2. Area deprivation (IMD)	.44*	1.00											24.38	17.67
3. Urban/rural	.12*	.21*	1.00										.89	.31
Agency indicators														
4. Expectation to go to university	-.08*	-.04*	-.01	1.00									3.73	1.34
5. Goal certainty	-.08*	-.01	.01	.59*	1.00								3.87	.93
6. Academic self-concept	-.02*	-.03#	.03#	.30*	.36*	1.00							0.00	1.00
7. School engagement	-.04*	-.01	.00	.19*	.19*	.25*	1.00						0.00	1.00
Control variables														
8. Female	.02*	.01	-.01	.07*	.02	-.11*	.03#	1.00					.48	.50
9. Non-white	.26*	.34*	.18*	.21*	.16*	.15*	.10*	.01	1.00				.30	.46
10. Academic attainment at 11	-.30*	-.26*	.07*	.38*	.32*	.31*	.17*	.06*	-.11*	1.00			0.00	1.00
Life satisfaction														
11. Life satisfaction age 14/15 (wave 2)	.00	.02*	.01*	.03*	.06*	.07*	.11*	-.12*	.02	-.00	1.00		3.97	.90
12. Life satisfaction age 20/21 (wave 7)	-.09*	.02*	.01*	.07*	.09*	.07*	.10*	.03#	-.04*	.06*	.11*	1.00	3.97	.93

Note: * p<.001; # p<.01

Data source: Longitudinal Study of Young People in England (n=9558)

Results

Table 1 gives the bivariate correlations between the indicators of structural and agentic variables used in the model. Family resources and area characteristics are positively correlated suggesting cumulative risk effects, i.e. experiencing one risk factor is likely to bring with it exposure to other risks as well. However, the correlation of r = .44 suggests a considerable degree of independence. There are positive correlations between the four indicators of agency, meaning that a young person with higher-level agency in one dimension is likely to also score high on another dimension. The highest correlation is between educational expectations and goal certainty (r = .59) while none of the other correlations is above .38, which again indicates a considerable degree of independence. There are negative associations between indicators of socio-economic risks and individual agency factors,

suggesting that young people from less privileged background are reporting lower levels of agency. However, the associations are rather small, ranging from r = -.02 for academic self-concepts to r = -.08 for goal certainty. Associations between family resources and academic attainment are considerably higher (r = -.30).

1. Transition pathways of a current cohort of young people

Using sequence analysis we identified six distinct patterns of education and employment transitions between ages 16 to 20 years. Table 2 presents the Duda-Hart stopping rule statistics, where lower T square values indicate a better fit. The optimum fit in terms of number of clusters is determined by maximising the Je(2)/Je(1), and the lowest Pseudo T-square statistic. According to these criteria the six-cluster solution has the best fit to the data (highlighted in bold).

Table 2: Stopping rules for cluster analysis of LSYPE based on Duda-Hart stopping rule

Number of clusters	Je(2)/Je(1)	Pseudo T-Square
2	0.9012	271.08
3	0.9031	84.13
4	0.4606	1974.18
5	0.8938	171.13
6	**0.9996**	**0.36**
7	0.5986	565.89

Data source: Longitudinal Study of Young People in England (n=9558)

Cluster compositions are shown in figure 1. Each cluster is presented in panel a-f, where the overall proportion of the cluster in a particular state is denoted on the ordinate, and the progression between October 2006 to May 2010 on the abscissa (45 months). *Cluster a* represents a pattern where respondents spend extended periods in full-time

education beyond compulsory school age. This cluster, which we termed '*higher education'*, comprises 45.2% of the respondents. Educational enrolment in this cluster is consistently above 80%, with little evidence of other consistent labour market activity.

Figure 1: Cluster compositions extracted from LSYPE

a) Higher education (45.2%)   b) Vocational education (6.5%)   c) Employment after some further education
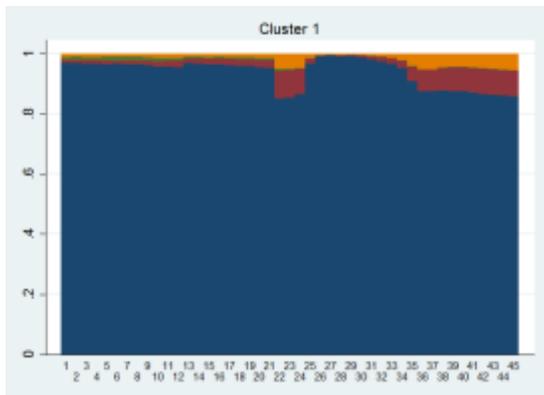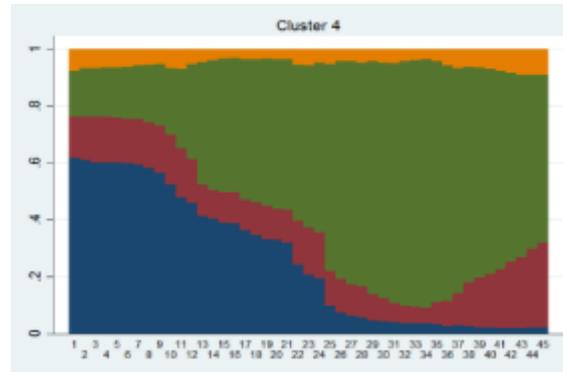
(14.5%)
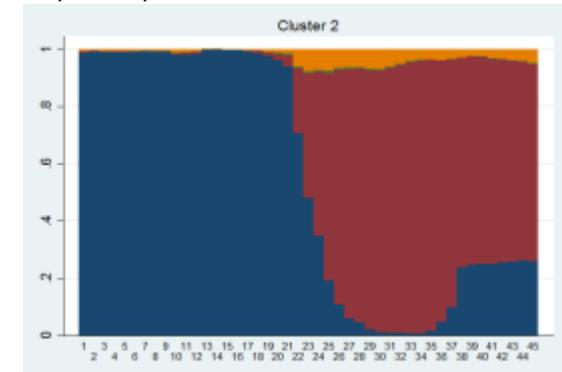
(d) Early work orientation (21.1%)   e) NEET after some further education (7.1%)   f) Long-term NEET (5.6%)
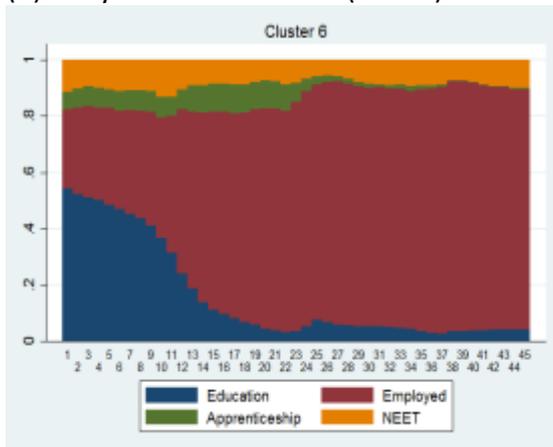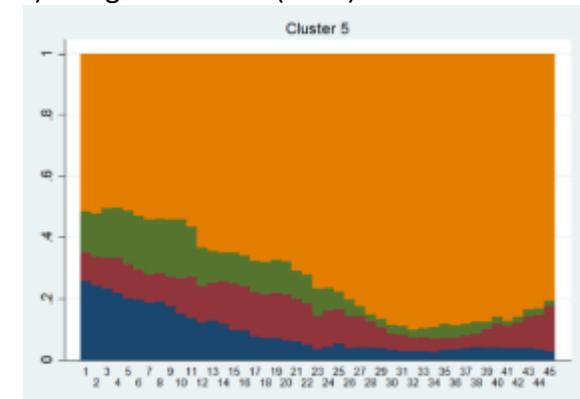
Note: Weighted data. The overall proportion of the cluster in a particular state is denoted on the ordinate, and the progression between October 2006 to May 2010 on the abscissa (45 months). Data source: Longitudinal Study of Young People in England (n=9558)

We identified *Cluster b* to represent *'vocational training'*, comprising about 6.5% of the respondents. Respondents in this cluster gradually transition away from full-time education into labour market orientated activities, including some employment, but largely show an increasing prevalence of vocational training which becomes the modal state, with roughly 90% of respondents engaged in this activity at age 18. After this there is a gradual upswing in the proportion of respondents in employment as youths transition from training into a full-time job.

Cluster c represents a transition to 'full-time *employment after some further education'*, involving 14.5% of the cohort. Respondents in this cluster typically remain in education for a limited time before a rapid rise in the proportion of respondents in employment to in excess of 90%. There is a slight increase in the proportion in higher education by age 19, which could however be due to differential censoring in the LSYPE (respondents in higher education are less likely to drop out).

Cluster d represents an early transition into the labour market, or *'early work orientation'*, capturing 21.1% of the cohort. In this cluster the proportion of respondents in full-time education falls early (and from a relatively low level) to be replaced by full time employment, which remains the modal state with in excess of 85% of respondents employed from age 18 onward. There is a relatively small proportion of NEET within this cluster (consistently around 7%) indicating the relative instability of employment in this cluster (many respondents have short spells out of the labour force).

Cluster e represents a more precarious transition to employment, involving prolonged periods of 'NEET *after some further education'*. This cluster comprises 7.1% of the cohort. Around age 18, participation in full-time education falls rapidly (from a relatively high level) among respondents in this cluster. However, in this cluster evidence of full-time employment is rare, and the proportion of respondents being NEET increases rapidly post-education accounting for more than 90% of activity at age 18, with a recovery thereafter.

Cluster f represents an extended *'NEET'* experience, comprising 5.6% of the respondents. NEET comprises more than 50% of activity at all ages within this cluster, with a movement away from participation in either education or full time work to disengagement. NEET respondents comprise more than 80% of cluster activity by age 18/19.

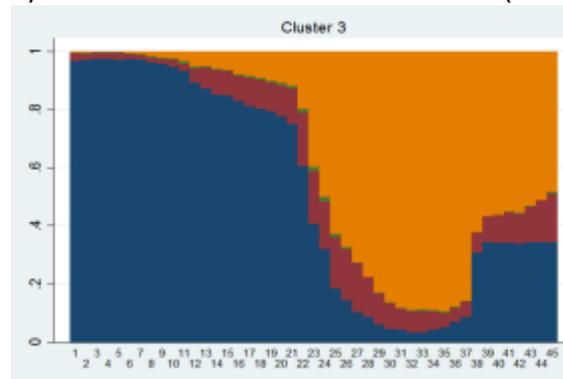**2. *Structure and agency as influences on transition experiences***

To assess the association between resources and transition experiences, the cluster membership was treated as a nominal variable. Multinomial logistic (MNL) regression was applied, using the largest cluster (*higher education)* as the reference group. We also tested interaction effects between socio-economic risk and the four agency dimensions.

Table 3 gives the estimated relative risk ratios and standard errors of the predictor variables and controls. The clusters used in the multi-nominal regression have very different sample sizes. In relatively small clusters, even large effect sizes might not reach statistical significance. However, even the smallest cluster (the long-term NEET) includes 369 cases. We find, that compared to respondents participating in higher education, young people growing up in families experiencing cumulative socioeconomic risk or who were living in relatively disadvantaged areas were more likely to show an early work orientation, were unemployed after some further education, or NEET. There are also significant associations between area deprivation and participation in vocational training instead of higher education.

Taking into account socioeconomic risk and the control variables included in the model, we find that indicators of agency also play a significant role in predicting transition pathways, suggesting independent effects. Compared to respondents participating in higher education, young people with higher academic aspirations were less likely to be in vocational education, were less likely to be employed after some education, or showed an early work orientation. Those with higher goal certainty were less likely to be in vocational training, NEET after some further education or long-term NEET. Those with high levels of academic self-efficacy were less likely to be in vocational training, employed after some education, showing an early work focus or NEET after some education. Those with high levels of school engagement were less likely to show an early work focus or were long-term NEET.

Table 3: Predicting transition patterns: Multinominal Regression using Relative Risk Ratios and Std. Error
(Reference Group= Higher Education [n=5084])

	Vocational Training		Employed after some education		Work-focus (employed since 16)		Unemployed after some education		NEET	
Socio-economic resources										
Family resources	.946	.076	1.019	.050	**1.120#**	.060	**1.186****	.071	**1.485*****	.122
IMD	**1.015****	.005	.996	.004	**1.008#**	.003	**1.013***	.005	**1.020****	.006
Urban	1.243	.335	.911	.134	1.411	.247	1.367	.325	1.700	.931
Agency										
Likely to apply to uni	**.719*****	.062	**.828*****	.047	**.672*****	.039	1.031	.088	.873	.109
Goal certainty	**.786#**	.078	.928	.063	.950	.065	**.798#**	.078	**.687***	.097
Self efficacy	**.710*****	.066	**.842****	.048	**.782*****	.047	**.834#**	.072	1.012	.155
School engagement	.945	.074	.947	.053	**.862***	.049	.933	.0671	**.655*****	.084
Controls										
Female	**.346*****	.056	.990	.100	**.670*****	.075	**.700****	.102	.956	.242
Non-white	**.252*****	.063	**.460*****	.063	**.183*****	.033	**.575****	.101	**.227*****	.072
Academic attainment at age 11	**.682*****	.064	.973	.064	**.565*****	.037	.854	.092	**.460*****	.062
Life satisfaction at wave 2	.885	.0730	.950	.055	**.861****	.046	.922	.068	.856	.098
N (unweighted)	530		1370		1515		690		369	

*** p <.oo1; ** p <.005; * p<.01; # p<.05
Data source: Longitudinal Study of Young People in England (n=9558)

Regarding the control variables we find significant associations between being female and reduced likelihood of entering vocational education, showing an early work orientation, being NEET after some education or long-term NEET (compared to those in higher education). We also find that young people from non-white ethnic backgrounds are more likely to participate in higher education and are less likely in any of the other transition pathways than their white peers. Higher prior academic achievement is associated with a decreased likelihood of entering vocational training, an early work focus or experience of long-term NEET.

Moreover, we find two significant interactions between socioeconomic risk and our indicators of agency. First, the combination of high risk and high goal certainty is associated with a higher likelihood of entering employment after some further education (main effect for family resources: RRR=.539, p=.104; main effect for goal certainty: RRR=.827, p=.015; interaction effect: RRR=1.169, p=.011). The finding could indicate that high levels of success expectations are associated with staying on in education, but that socioeconomic or other reasons might then compel the young person to get a job before entering university. Second, high levels of socioeconomic risk in combination with high academic self-efficacy is associated with an increased likelihood of being NEET after some further education (main effect for family resources: RRR=.543, p=.160; main effect for self-efficacy: RRR=.734, p=.004; interaction effect: RRR=1.360, p=.028), possibly suggesting a 'dark' side of high self-efficacy beliefs, i.e. that unrealistic self-perceptions might be harmful to individuals, promoting over-confidence and inappropriate persistence.

3. *Who is most satisfied with one's life?*

We ran a stepwise OLS regression to assess the association between structural and agency factors and life satisfaction at age 19/20, as well as the role of transition experiences, treating the cluster membership as a nominal variable. Model 1 includes structural and agency indicators. We find significant associations between life satisfaction and cumulative socioeconomic risk, area deprivation, our indicator of goal certainty and school engagement. Model 2 adds the dummies for the transition clusters. Adding the transition patterns as explanatory variables appears to fully mediate the influence of family socioeconomic risk, yet there remains a significant association between area disadvantage and the two agency indicators. The significant effect of these variables is not removed after adding the control variables to the model. We find independent associations between being female and earlier life evaluation (feeling happy with one's life) for later levels of satisfaction. There were no significant interaction effects between socioeconomic risk and the four agency dimensions. There are no significant differences in levels of life satisfaction between those in higher education, vocational training, and employment after some further education, while those who experienced NEET report the lowest level of life satisfaction. Those with an early work orientation and who experienced unemployment after some further education reported lower levels of life satisfaction then those in higher education.

Table 4. Predicting life satisfaction (OLS regression)

	Model 1		Model 2		Model 3	
	Unstandardised Coefficient	Std. Err	Unstandardised Coefficient	Std. Err	Unstandardised Coefficient	Std. Err
Socioeconomic resources						
Family resources	**-.033***	.013	-.017	.013	-.015	.015
IMD	**-.003****	.001	**-.003***	.001	**-.003***	.001
Urban	-.046	.050	-.031	.050	-.022	.053
Agency						
Likely to apply to uni	.016	.018	.009	.018	.014	.019
Goal certainty	**.070*****	.022	**.060***	.022	**.054#**	.022
Self efficacy	.006	.016	.004	.016	.019	.018
School engagement	**.067*****	.018	**.053****	.018	**.045#**	.019
Transition patterns (ref=Higher Education)						
Vocational training			.089	.059	.115	.059
Employed after some education			-.046	.036	-.044	.038
Work focus employed at 16			**-.152*****	.044	**-.140****	.048
Unemployed after some education			**-.253*****	.067	**-.249*****	.069
NEET			**-.553*****	.132	**-.627*****	.136
Controls						
Female					**.094****	.030
Non-white					-.069	.022
Academic attainment at age 11					-.018	.022
Life satisfaction at wave 2					**.099*****	.018
R²	.025		.043		.056	

*** p <.oo1; ** p <.005; * p<.01; # p<.05 ; Data source: Longitudinal Study of Young People in England (n=9558)

Schoon, Lyons-Amos

A socio-ecological model of agency: The role of structure and agency in shaping education and employment transitions in England

Discussion

In this paper we present a socio-ecological model of human agency taking into account multiple dimensions of agency, as well as their inter-linkage with social and economic resources in proximal settings as well as the wider social context in shaping the transition from school to work. We show that individuals steer the course of their lives, and actively cope with given structural constraints. In particular, our findings suggest that agency can give rise to the creation of niches that enable the experience of competence and life satisfaction, especially among those who do not follow the academic track. For some however, the lack of socioeconomic and psycho-social resources is too overwhelming and they encounter long-term experience of NEET or are not able to transform their educational credentials into employment opportunities. The findings thus highlight that for a better understanding of variations in youth transitions it is important to consider structural constraints as well as the role of the agent.

The findings illustrate the heterogeneity of transition pathways after completion of compulsory schooling in a current cohort of young people in England. We identified six distinct transition patterns, suggesting that the assumption of polarised transitions is not sufficient to capture the diverse experiences of young people (Schoon, 2015; Schoon & Lyons-Amos, 2016). In addition to the 45% of young people who continued in full-time education after compulsory school leaving age, we find that 42% of the sample succeeded in making the transition into the labour market (6.5% who engaged in vocational training, 14.5% entered employment after some period in further education, 21.1% showed an early work orientation, entering full-time employment more or less immediately after completing compulsory schooling) . However, About 13% of the sample encountered precarious transitions (7.1% experienced NEET after some further education, and 5.6% encountered long periods of NEET). The findings show that not all young people expect to go to university, that many succeeded in making the transition into employment by age 20, although some are struggling. Strategies aiming to provide the training and skills needed for the 21st century thus need to provide alternatives to the current provision of post-compulsory education, which is very much focused on gaining 4 year academic qualifications Wolf, 2016), and provide viable pathways for those who do not expect, or cannot afford to go to university (Schoon & Lyons-Amos, 2016).

The study found significant associations between transition experiences and family socioeconomic resources as well as area characteristics, confirming the assumption of cumulative (dis)advantage and multiple deprivation in the transition to adulthood, i.e. less privileged young people are leaving education early or encounter more problems in establishing themselves in the labour market. Moreover, it matters where one lives. The opportunities and constraints in local labour markets are an independent risk factor shaping young people's lives. Yet, the findings also suggest that individuals are not passively exposed to the experience of disadvantage, and illustrate the heterogeneity of how young people respond to adverse socioeconomic conditions.

For example, the associations between socioeconomic resources and the four indicators of agency are very small (ranging from -.01 to -.08), confirming evidence reported in a meta-analytic study of associations between parental SES and self-esteem (Twenge & Campbell, 2002), and suggesting potential self-protective mechanisms among adolescents that can reduce the effect of socioeconomic circumstances on the expression of individual agency. In particular, the findings suggest that agency indicators predict transition outcomes independent of structural constraints, i.e. they have a unique effect after controlling for family background, area effects, prior academic attainment, gender and ethnicity. Given the constraining forces in their immediate and wider social context, young people can to some extent actively steer the course of their lives. The study highlights the importance of conceptualising the role of the agent for a better understanding of diversity in youth transitions, rather than solely focusing on structures and socioeconomic resources (see also Hitlin & Johnson, 2015). Moreover, there are interactive as well as domain-specific effects suggesting that young people select a specific pathway or niche that can offer them developmentally appropriate challenges through which they can experience competence and which enables them to feel satisfied about their lives.

Interestingly there is no significant difference in self-efficacy between those in higher education and those experiencing long-term NEET, potentially

pointing to a possible 'dark' side of high competence beliefs, which for some can imply that they overestimate their abilities. This assumption is also supported by the significant interaction effect suggesting that high socioeconomic risk and high academic self-concepts are associated with an increased likelihood of experiencing NEET after some further education. We also found a significant interaction between high risk and goal certainty and the likelihood of entering employment after some further education, which might point to a turning point, where young people experiencing socioeconomic adversity were initially on track to higher education but might have changed their path after finding a promising full-time job. We furthermore find that high levels of school engagement are associated with a reduced likelihood of leaving school directly after compulsory schooling, to enter employment or prolonged periods of NEET, pointing to the significant role of engagement in the school context and underlying preferences for academic work as an influence on later life choices (Schoon, 2008). The findings suggest that within the available opportunities young people choose a pathway that fits with their preferences and which they see as achievable (Eccles et al., 1993). Although we find some interactive effects, cumulative risk effects and the independent effect model appear to be most appropriate in describing the inter-linkages between structure and agency shaping youth transitions.

Regarding the subjective evaluation of one's life at age 19/20 we find that both external (structural) and internal (individual) factors independently shape how one feels about one's life as a whole. Moreover, transition experiences fully mediate the influence of family socioeconomic resources on later life satisfaction, suggesting cumulative processes, and the amplification of prior difficulties or advantages. Yet, transition experiences do not fully mediate the influence of area characteristics and individual agency factors. Agency factors, especially goal certainty and school engagement, are significantly associated with life satisfaction after controlling for structural factors, gender, ethnicity, transition experiences, prior academic attainment and life-satisfaction. Forethought and self-directedness are independent predictors of later life satisfaction, pointing to their role in shaping one's life course transitions as well as one's outlook on life.

There are multiple pathways to a successful transition, defying the notion of polarisation or assumption of a 'right way'. Most satisfied with their lives are those in higher education as well as those who established themselves in the labour market either through vocational training or some further education, while those who experienced NEET were least satisfied. Moreover, life satisfaction is influenced by area characteristics and transition experiences, suggesting that it matters where one lives and how one's life is shaping up. Previous studies have shown that neighbourhood effects are strongest during early childhood and late adolescence (Brooks-Gunn, Duncan & Aber, 1997). This study confirms the importance of area deprivation for young adults and points to local opportunities, and potentially also collective socialisation (Ainsworth, 2002; Ioannides & Loury, 2004), which influence the type of role models a young person is exposed to outside the home. The explained variance in our model is however low, and other factors and processes are likely to also play a role in shaping the subjective evaluation of one's life.

In interpreting the findings a number of limitations have to be considered. Like in all longitudinal studies we are faced with the problem of missingness in response. We checked the robustness of findings using multiple imputations (Mi-command in STATA), which confirmed the stability of the solution. We also had to make do with the information available in the data set. For example, some of our indicators of agency and the measures of life satisfaction are based on single items, which are less stable than multi-item scales. However, single-item assessments of education aspirations are widely used in large-scale surveys, suggesting satisfactory face validity (Sewell & Hauser, 1993; Schulenberg & Schoon, 2012), as are single item measures of life satisfaction (Lucas & Donnellan, 2012). The correlations between our two indicators of subjective evaluation of one's life assessed during early adolescence and young adulthood are low, suggesting that either the measures tap into different aspects, or that there is considerable change in how young people evaluate their lives at different stages of the life course. Future research is needed to clarify this relationship further. Moreover, our study observes only a short period in the transition to employment, and changes in transitions at a later time point are very likely. Yet, we focus on a crucial period in the lives of young

people when major decisions about which path to follow are made. However, our findings might be unique to the English context, reflecting its liberal regulation of school-to-work transitions and the impact of the 2008 recession on education and employment opportunities of young people. Future studies have to assess if the established patterns also apply in other countries, characterised by different transition regimes and economic circumstances. The role of gender and ethnicity, which showed highly significant associations with transition experiences as well as adult life satisfaction, also has to be studied in more detail in future studies.

Despite these limitations this study illustrates the fruitfulness of combining assumptions regarding school-to-work transitions from different disciplines and of examining interactions between person and environment (Schoon, 2015b). The socio-ecological model of agency enables us to gain a better insight into the processes underlying diverse youth transitions. We show that transition experiences are influenced by both socioeconomic and psycho-social resources and we have gained a better understanding of the interplay between structure and agency. Young people from less privileged background are less likely to participate in higher education. Yet, young people do not passively follow a pre-determined alternative of college or nothing, and given that constraints are not overpowering, can create alternative sustainable pathways. They are able to navigate their lives by balancing the resources available to them. The socio-ecological model of agency enables us to a. show the association between structural factors and indicators of agency; b. identify processes linking socioeconomic resources in the proximal and wider context to the timing and sequencing of transitions, taking into account agency factors; and c. illustrate how agency indicators can steer the direction of a pathway that corresponds to one's preferences, intentions, and self-concepts independently or in interaction with structural constraints. Whether this path will continue to fit one's preferences, intentions, and forethoughts remains to be seen, as both socioeconomic and psycho-social resources can change over time, introducing new opportunities and challenges.

Acknowledgements

References

Ainsworth, J. W. (2002). Why does it take a village? The mediation of neighborhood effects on educational achievement. *Social Forces, 81*(1), 117-152. https://doi.org/10.1353/sof.2002.0038

Arnett, J. J. (2000). Emerging Adulthood. A theory of development from the late teens to the late twenties. *American Psychologist, 55*(5), 469-480. https://doi.org/10.1037/0003-066X.55.5.469

Ashby, J. S., & Schoon, I. (2010). Career success: The role of teenage career aspirations, ambition value and gender in predicting adult social status and earnings. *Journal of Vocational Behavior, 77*(3), 350-360. https://doi.org/10.1016/j.jvb.2010.06.006

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26. https://doi.org/10.1146/annurev.psych.52.1.1

Bandura, A. (2006). Toward a Psychology of Human Agency. *Perspectives on Psychological Science, 1*(2), 164-180. https://doi.org/10.1111/j.1745-6916.2006.00011.x

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development, 72*(1), 187-206. https://doi.org/10.1111/1467-8624.00273

BIS. (2013). *Participation rates in higher education*. Department for Business Innovation and Skills (BIS). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/306138/13-p140-HEIPR_PUBLICATION_2011-12_2_.pdf

Blanden, J., Gregg, P., & Macmillan, L. (2011). *Intergenerational persistence in income and social class: The impact of within-group inequality*. Retrieved from Bonn: Bourdieu, P., & Passeron, J. (1977). *Reproduction in Education, Society and Culture* (Vol. 5). London: Sage Publications.

Breen, R., & Goldthorpe, J. H. (2001). Class, mobility and merit - The experience of two British birth cohorts. *European Sociological Review, 17*(2), 81-101. https://doi.org/10.1093/esr/17.2.81

Breen, R., Luijkx, R., Muller, W., & Pollak, R. (2009). Nonpersistent Inequality in Educational Attainment: Evidence from Eight European Countries. *American Journal of Sociology, 114*(5), 1475-1521. https://doi.org/10.1086/595951

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

Brooks-Gunn, J., Duncan, G. J., & Aber, L. (1997). *Neighborhood Poverty: Context and Consequences for Children*. New York: Russel Sage Foundation.

Buchmann, M. C. (1989). *The Script of Life in Modern Society: Entry Into Adulthood in a Changing World*. Chicago: Chicago University Press.

Buchmann, M. C., & Kriesi, I. (2011). Transition to Adulthood in Europe *Annual Review of Sociology, 37*, 481-503. https://doi.org/10.1146/annurev-soc-081309-150212

Bukodi, E., & Goldthorpe, J. H. (2013). Decomposing 'Social Origins': The Effects of Parents' Class, Status, and Education on the Educational Attainment of Their Children. *European Sociological Review, 29*(5), 1024-1039. https://doi.org/10.1093/esr/jcs079

Conger, R. D., Conger, K. J., & Martin, M. J. (2010). Socioeconomic Status, Family Processes, and Individual Development. *Journal of Marriage and the Family, 72*(3), 685-704. https://doi.org/10.1111/j.1741-3737.2010.00725.x

Dannefer, D. (2003). Cumulative advantage/disadvantage and the life course: Cross fertilizing age and social science theory. *The Journals of Gerontology, 58*, 327-337. https://doi.org/10.1093/geronb/58.6.S327

Diener, E., Inglehart, R., & Tay, L. (2013). Theory and Validity of Life Satisfaction Scales. *Social Indicators Research, 112*(3), 497-527. https://doi.org/10.1007/s11205-012-0076-y

DiPrete, T. A., & Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments *Annual Review of Sociology*, Vol. 32, pp. 271-297). Palo Alto: Annual Reviews. https://doi.org/10.1146/annurev.soc.32.061604.123127

Duckworth, K., & Schoon, I. (2012). Beating the odds: Exploring the aspects of social risk on young people's school-to-work transitions during recession in the UK. *National Institute Economic Review, 222*, 38-51. https://doi.org/10.1177/002795011222200104

Duncan, G. J., & Brooks-Gunn, J. (1997). *Consequences of growing up poor*. New York: Russell Sage Foundation Press.

Eccles, J. S. (2008). Agency and Structure in Human Development. *Research in Human Development, 5*(4), 231-243. https://doi.org/10.1080/15427600802493973

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Maciver, D. (1993). Development during adolescence - The impact of stage-environment fit on young adolescents experiences in schools and in families. *American Psychologist, 48*(2), 90-101. https://doi.org/10.1037//0003-066X.48.2.90

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109-132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Elder, G. H. (1998). The life course as developmental theory. *Child Development, 69*(1), 1-12. https://doi.org/10.2307/1132065

Elder, G. H., & Shanahan, M. J. (2007). The life course and human development *The Handbook of Child Psychology. 6th edition*. New York: Wiley.

Emirbayer, M., & Mische, A. (1998). What is agency? *American Journal of Sociology, 103*(4), 962-1023. https://doi.org/10.1086/231294

Erikson, R. (1984). Social class of men, women and families. *Sociology-the Journal of the British Sociological Association, 18*(4), 500-514. https://doi.org/10.1177/0038038584018004003

Erikson, R., & Jonsson, J. O. (Eds.). (1996). *Can education be equalized? The Swedish case in comparative perspective*. Boulder, Colo. ; Oxford: Westview Press.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of evidence. *Review of Educational Research, 74*, 59-109. https://doi.org/10.3102/00346543074001059

Fuchs, S. (2001). Beyond agency. *Sociological Theory, 19*(1), 24-40. https://doi.org/10.1111/0735-2751.00126

Furlong, A., & Cartmel, F. (1997). *Young People and Social Change*. Buckingham: Open University Press.

Galindo-Rueda, F., & Vignoles, A. (2005). The Declining Relative Importance of Ability in Predicting Educational Attainment *Journal of Human Resources, 40*(2), 335-353.

Giddens, A. (1991). *Modernity and self-identity: self and society in the late modern age*. Cambridge: Polity Press.

Gregg, P., & Wadsworth, J. (2001). Everything you ever wanted to know about measuring worklessness and polarization at the household level but were afraid to ask. *Oxford Bulletin of Economics and Statistics, 63*, 777-806. https://doi.org/10.1111/1468-0084.63.spe1.9

Guo, G., & Harris, K. M. (2000). The mechanisms mediating the effects of poverty on children's intellectual development. *Demography, 37*(4), 431-447.

Heckhausen, J. (1999). *Developmental regulation in adulthood: age-normative and sociostructural constraints as adaptive challenges*. Cambridge: Cambridge University Press.

Heckhausen, J., & Chang, E. S. (2009). Can ambition help overcome social Inequality in the transition to adulthood? Individual agency and societal opportunities in Germany and the United States. *Research in Human Development, 6*(4), 235-251. https://doi.org/10.1080/15427600903281244

Hitlin, S., & Elder, G. H., Jr. (2007). Time, self, and the curiously abstract concept of agency. *Sociological Theory, 25*(2), 170-191. https://doi.org/10.1111/j.1467-9558.2007.00303.x

Hitlin, S., & Johnson, M. K. (2015). Reconceptualizing Agency within the Life Course: The Power of Looking Ahead. *American Journal of Sociology, 120*(5), 1429-1472. https://doi.org/10.1086/681216

Ioannides, Y., & Loury, L. (2004). Job information networks, neighbourhood effects and inequality. *Journal of Economic Literature, 42*(4), 1056-1093. https://doi.org/10.1257/0022051043004595

Jencks, C., & Mayer, S. E. (1990). The social consequences of growing up in a poor neighbourhood. In L. E. J. Lynn & M. G. H. McGeary (Eds.), *Inner-city poverty in the United States* (pp. 111-187). Washington, DC: National Academy Press.

Jones, G. (2002). *The youth divide: diverging paths to adulthood*. York: Joseph Rowntree Foundation.

Kerckhoff, A. C. (2001). Education and social stratification processes in comparative perspective. *Sociology of Education*, 3-18. https://doi.org/10.2307/2673250

Kins, E., & Beyers, W. (2010). Failure to Launch, Failure to Achieve Criteria for Adulthood? *Journal of Adolescent Research, 25*(5), 743-777. https://doi.org/10.1177/0743558410371126

Lerner, R. M. (1996). Relative plasticity, integration, temporality, and diversity in human development: A developmental contextual perspective about theory, process, and method. *Developmental Psychology, 32*(4), 781-786. https://doi.org/10.1037/0012-1649.32.4.781

Loyal, S., & Barnes, B. (2001). 'Agency' as a red herring. *Philosophy of the Social Sciences, 31*(4), 507-524. https://doi.org/10.1177/004839310103100403

Lucas, R. E., & Donnellan, M. B. (2012). Estimating the Reliability of Single-Item Life Satisfaction Measures: Results from Four National Panel Studies. *Social Indicators Research, 105*(3), 323-331. https://doi.org/10.1007/s11205-011-9783-z

McLanahan, S. (2004). Diverging destinies: How children are faring under the second demographic transition. *Demography, 41*(4), 607-627. https://doi.org/10.1353/dem.2004.0033

McLeod, J. D., & Lively, K. J. (2003). Social structure and personality. In J. DeLamater (Ed.), *Handbook of Social Psychology* (pp. 77-102). New York: kluwer.

Moore, K. A., Vandivere, S., & Redd, Z. (2006). A sociodemographic risk index. *Social Indicators Research, 75*(1), 45-81. https://doi.org/10.1007/s11205-004-6398-7

Mortimer, J. T. (2003). *Working and growing up in America*. Cambridge, Mass. ; London: Harvard University Press.

Murry, V. M., Berkel, C., Gaylord-Harden, N. K., Copeland-Linder, N., & Nation, M. (2011). Neighborhood Poverty and Adolescent Development. *Journal of Research on Adolescence, 21*(1), 114-128. https://doi.org/10.1111/j.1532-7795.2010.00718.x

OECD. (2012). *Education at a Glance 2012. Highlights*.

Oishi, S. (2014). Socioecological Psychology. In S. T. Fiske (Ed.), *Annual Review of Psychology, Vol 65* (Vol. 65, pp. 581-609).

ONS. (2013). Urban and Rural Area Definitions for Policy Purposes in England and Wales. Retrieved from Office of National Statistics. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/239477/RUC11methodologypaperaug_28_Aug.pdf

Roberts, S. (2011). Beyond 'NEET' and 'tidy' pathways: considering the 'missing middle' of youth transition studies. *Journal of Youth Studies* 14, 21-39. https://doi.org/10.1080/13676261.2010.489604

Rutter, M. (1981). Stress, Coping and Development - Some Issues and Some Questions. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 22*(4), 323-356. https://doi.org/10.1111/j.1469-7610.1981.tb00560.x

Salmela-Aro, K. (2009). Personal goals and well-being during critical life transitions: The four C's-Channelling, choice, co-agency and compensation. *Advances in Life Course Research, 14*(1-2), 63-73. https://doi.org/10.1016/j.alcr.2009.03.003

Scarpetta, S., Sonnet, A., & Manfredi, T. (2010). *Rising youth unemployment during the crisis: How to prevent negative long-term consequences on a generation?* Retrieved from Paris:

Schoon, I. (2006). *Risk and Resilience: Adaptations in Changing Times* Cambridge: Cambridge University Press.

Schoon, I. (2007). Adaptations to changing times: Agency in context. *International Journal of Psychology, 42*(2), 94-101. https://doi.org/10.1080/00207590600991252

Schoon, I. (2008). A transgenerational model of status attainment: The potential mediating role of school motivation and education. *National Institute Economic Review, 205*, 72-82. https://doi.org/10.1177/0027950108096590

Schoon, I. (2014). Parental worklessness and the experience of NEET among their offspring. Evidence from the Longitudinal Study of Young People in England (LSYPE). *Longitudinal and Life Course Studies, 6*(6), 129-150. https://doi.org/10.14301/llcs.v5i2.279

Schoon, I. (2015a). Diverse pathways: Rethinking the transition to adulthood. In P. R. Amato, A. Booth, S. McHale, & J. Van Hook (Eds.), *Families in an era of increasing inequality. Diverging destinies* (pp. 115-136). London: Springer. https://doi.org/10.1007/978-3-319-08308-7_9

Schoon, I. (2015b). Let's work together: Towards interdisciplinary collaboration. *Research in Human Development 12,* 350-355. https://doi.org/10.1080/15427609.2015.1068050

Schoon, I. and Lyons-Amos, M. (2016). *Diverse pathways in becoming an adult: the role of structure, agency and context. Research in Social Stratification and Mobility*. ISSN 02765624. http://www.sciencedirect.com/science/article/pii/S0276562416300178

Schoon, I., & Silbereisen, K. R. (Eds.). (2009). *Transitions from School to Work: Globalisation, Individualisation, and Patterns of Diversity*. New York Cambridge University Press.

Schulenberg, J. E., Bryant, A. L., & O'Malley, P. M. (2004). Taking hold of some kind of life: How developmental tasks relate to trajectories of well-being during the transition to adulthood. *Development and Psychopathology, 16*(4), 1119-1140. https://doi.org/10.1017/S0954579404040167

Schulenberg, J. E., & Schoon, I. (2012). The transition to adulthood in the UK, the US, and Finland: Differential social role pathways, their predictors and correlates. *Longitudinal and Life Course Studies, 3*(2), 164-172.

Settersten, J., R. A. (2007). The new landscape of adult life: Road maps, signposts, and speed lines *Research in Human Development, 4*, 239-252. https://doi.org/10.1080/15427600701663098

Settersten, R. A., Furstenberg, F. F., & Rumbaut, R. G. (2005). *On the frontier of adulthood: Theory, research, and public policy*. Chicago, Ill. ; London: University of Chicago Press. https://doi.org/10.7208/chicago/9780226748924.001.0001

Sewell, W. H., & Hauser, R. M. (1993). *A review of the Wisconsin Longitudinal Study of Social and Psychologocial Factors in Aspirations and Achievements 1963-1993*. University of Wisconsin-Adison, Center for Demography and Ecology: CDE Working Paper No.92-01.

Shanahan, M. J. (2000). Pathways to adulthood in changing societies: Variability and mechanisms in life course perspective. *Annual Review of Sociology, 26*, 667-692. https://doi.org/10.1146/annurev.soc.26.1.667

Shavit, Y., Arum, R., & Gamoran, A. (2007). Stratification in higher education: A comparative study. Stanford: Stanford University Press.

Twenge, J. M., & Campbell, W. K. (2001). Age and birth cohort differences in self-esteem: A cross-temporal meta-analysis. *Personality and Social Psychology Review, 5*(4), 321-344. https://doi.org/10.1207/S15327957PSPR0504_3

Wolf, A. (2016). *Remaking tertiary education: can we create a system that is fair and fit for purpose?* Retrieved from: http://epi.org.uk/wp-content/uploads/2016/11/remaking-tertiary-education-web.pdf

Endnotes

[i] We focus on the opportunity structures associated with socio-economic disadvantage, not potential cultural explanations associated with patterns of behaviours, beliefs or values that are transmitted through socialisation.

[ii] A Key Stage is a stage of the state education system in England, Wales and Northern Ireland. Key Stage 2 reflects attainment at the later stage of primary education, often known as junior schools.

The consequences of contact with the criminal justice system for health in the transition to adulthood

Michael H. Esposito   University of Washington, Seattle, USA
esposm2@uw.edu
Hedwig Lee   University of Washington, Seattle, USA
Margret T. Hicken   University of Michigan, Ann Arbor, USA
Lauren C. Porter   University of Maryland, College Park, USA
Jerald R. Herting   University of Washington, Seattle, USA

Abstract

A rapidly growing literature has documented the adverse social, economic and, recently, health impacts of experiencing incarceration in the United States. Despite the insights that this work has provided in consistently documenting the deleterious effects of incarceration, little is known about the specific timing of criminal justice contact and early health consequences during the transition from adolescence to adulthood—a critical period in the life course, particularly for the development of poor health. Previous literature on the role of incarceration has also been hampered by the difficulties of parsing out the influence that incarceration exerts on health from the social and economic confounding forces that are linked to both criminal justice contact and health. This paper addresses these two gaps in the literature by examining the association between incarceration and health in the United States during the transition to adulthood, and by using an analytic approach that better isolates the association of incarceration with health from the multitude of confounders which could be alternatively driving this association. In this endeavor, we make use of variable-rich data from The National Longitudinal Study of Adolescent to Adult Health (n = 10,785) and a non-parametric Bayesian machine learning technique- Bayesian Additive Regression Trees. Our results suggest that the experience of incarceration at this stage of the life course increases the probability of depression, adversely affects the perception of general health status, but has no effect on the probability of developing hypertension in early adulthood. These findings signal that incarceration in emerging adulthood is an important stressor that can have immediate implications for mental and general health in early adulthood, and may help to explain long lasting implications incarceration has for health across the life course.

Keywords

Incarceration, population health, causal inference, machine learning

Introduction

The United States (U.S.) currently has the highest incarceration rate in the world (Western, 2006). Indeed, recent estimates show that nearly 732 of every 100,000 adults (1,352 of every 100,000 men, and 126 of every 100,000 women) are in the U.S. correctional population, whereas only 510 of 100,000 adults are incarcerated in Cuba (the country of at least 500,000 people with the second highest incarceration rate), and only 233 of 100,000 individuals are incarcerated in Chile (the OECD country with the second highest incarceration rate after the U.S.) (Annie E. Casey Foundation, 2013; Glaze & Parks, 2011; Walmsley, 2013). Perhaps more consequential is that the U.S. leads the world in the confinement of youth, with estimates showing that 336 of every 100,000 youth are held in confinement (Annie E. Casey Foundation, 2013; Glaze & Parks, 2011). For reference, in South Africa (the country with the second highest youth incarceration rate in the world), only 69 of every 100,000 youth are held in confinement (Annie E. Casey Foundation, 2011; Hazel, 2008). Because of the high rate of incarceration across the U.S. population, many youths have experienced some form of contact with the criminal justice system, either through their own incarceration or through the incarceration of a family member.

Given the staggeringly high prevalence of incarceration in the U.S., researchers have turned towards investigating the impact that incarceration has on the lives of those who experience it. This work has shown that having been incarcerated adversely affects social and economic well-being along multiple dimensions. Indeed, research has shown that former inmates earn lower wages and face greater difficulties in securing employment than equivalent individuals who have never been incarcerated (Holzer, Raphael & Stoll, 2004; Western, 2002). In addition, formerly incarcerated individuals are more likely to experience housing difficulties (Geller & Curtis, 2011; Western, Kling & Weiman, 2001), have reduced social standing (Schnittker & Bacak, 2013), are more likely to experience marital dissolution (Lopoo & Western, 2005), and struggle maintaining parenting roles (Turney & Wildeman, 2013).

A relatively new body of work has begun to document that incarceration adversely affects mental and physical health. Current and former inmates are more likely to contract infectious diseases, have higher risk of death, and are more likely to develop either chronic physical or mental conditions than members of the general population (Golzari, Hunt & Anoshiravani, 2006; Maruschak, 2010; Massoglia, 2008; Massoglia & Pridemore, 2015). Second-degree contact with incarceration, typically defined as having a parent or other family member incarcerated, has also been shown to be predictive of poor mental and physical health in both childhood and adulthood (Lee, Wildeman, Wang, Matsuko & Jackson, 2014; Roettger & Boardman, 2012; Wildeman, Schnittker & Turney, 2012). These adverse health effects are thought to be maintained by numerous mechanisms. For instance, recall that income, occupation and employment opportunities are all limited by experience of incarceration. As each of these factors are inputs to good health, their restriction by incarceration translates into worse health for former inmates (Mirowsky & Ross 2003; Schnittker & John, 2007). Beyond the adverse social and economic consequences of incarceration, the experience of having once been incarcerated is itself a chronic stressor that may directly affect mental and physical health through psychosocial and biological pathways and may also work indirectly to impact health through economic strains and deterioration in family functioning, particularly if introduced during critical periods in the life course (Massoglia, 2008; Pervanidou & Chrousos, 2012; Roettger & Boardman, 2012).

Despite the above insights, critical gaps still exist in our understanding of the link between incarceration and health. In this paper, we address two areas in which our current understanding of the relation between incarceration and health is limited. Specifically, we (1) estimate the relation between incarceration and health over a limited time span (during the transition to adulthood). Previous research that has examined the health consequences of incarceration has focused almost exclusively on health well into adulthood, overlooking the possibility that incarceration at young ages may have immediate effects on health. Then, (2) we also employ methods

designed to better isolate the association of incarceration on health from the large array of confounders which could be alternatively driving this association. In this endeavor, we make use of variable-rich data from The National Longitudinal Study of Adolescent to Adult Health (Add Health) and a non-parametric Bayesian machine learning technique (i.e. Bayesian Additive Regression Trees or BART). This combination of rich data and robust statistical learning approach allows us to effectively control for a high-dimensional set of features that surround the association between previous incarceration and health, and obtain a less biased estimate of previous incarceration's effect on health during this important period in the life course.

Background

Incarceration, health, and the transition from adolescence to adulthood

To date, most empirical studies linking incarceration to health have examined the effect of incarceration at any point during the life course on health at mid-life or later (Massoglia, 2008; Schnittker & John, 2007; Schnittker, Massoglia & Uggen, 2012). Limited attention has been paid towards whether and how experiences of incarceration, specifically during the transition to adulthood, might serve to impact health relatively early in adulthood. This gap is problematic as life course theory suggests that the transition to adulthood (e.g. approximately at ages 15-30 years) is a critically important stage for the development of social and economic wellbeing, as well as good mental and physical health, and that the disruption at this stage may set in motion a lifetime of poor health, and increased exposure to other health-adverse conditions and behaviours (Alwin & Wray, 2005; Ben-Shlomo & Kuh; 2002; Cable, 2014). Indeed, life course epidemiology has enumerated the ways in which adverse social exposures (such as incarceration) can have important impacts on health. Beyond the direct biological stress impact of incarceration, we are informed by the chains of risk formulation, which indicate that early experience of incarceration sets in motion additional risk behaviour, or clusters of behaviours related to short-term and long-term health

(Cable, 2014). Incarceration works through these mechanisms (see below) to promote negative health, beyond just the biological impact of the stress generated from having been incarcerated. While our task here is to simply identify the presence of an incarceration effect, and not to model these chains or pathways, it is important to highlight some of these mechanisms, and note that they inform our general approach.

The transition to adulthood is a critical period in the life course where individuals acquire the human capital skills and material resources needed to secure future resources, and develop productive social relationships (Settersten, Furstenberg & Rumbaut, 2005). Incarceration during this critical period removes individuals from opportunities to develop these skills and secure other forms of social and material resources, potentially setting up a chain of risks both in terms of limiting protective resources and in terms of engaging in additional health risk behaviours (Ben-Shlomo & Kuh, 2002). Instead of having the opportunity to obtain additional educational credentials, or develop social networks that might lead to marriage or job opportunities, for instance, incarcerated individuals are left sidelined both economically and socially (Pettit & Western, 2004). Even when individuals return to the community from prison they face substantial barriers to re-building human and social capital due to factors such as employment discrimination, social stigma and isolation (Rose & Clear, 1998; Western et al., 2001).

As many of the skills and resources that are fostered in the transition to adulthood can be used to protect health over the life course, the disruption of this critical period by incarceration is likely to result in worse health over the course of life. For instance, educational attainment is a strong determinate of multiple health outcomes in early, mid, and late adulthood. Adults with higher levels of educational attainment have a lower risk of mortality, are less likely to smoke, have better self-rated health, and have lower levels of depression and anxiety than their less educated counterparts (Cutler & Lleras-Muney, 2006; Mirowsky & Ross, 2003). Because incarceration during the transition to adulthood restricts one's ability to attain higher levels of education it may also serve to

restrict one's ability to attain good health. Similarly, because marriage buffers against numerous adverse health outcomes (such as premature mortality, or the development of functional limitations), the restricted marriage opportunities that result from incarceration during the transition to adulthood may leave individuals less well equipped to protect against challenges to their health (Robards, Evandrou, Falkingham & Vlachantoni, 2012).

The direct chronic stress (i.e. stress that results from repeated, and/or prolonged exposure to adverse events) brought about by being incarcerated at this life stage might also have both immediate, and long lasting health implications, as argued by a biological programing perspective (Ben-Shlomo & Kuh, 2002). Stress can stem from multiple sources related to incarceration, including severed relationships with family and/or family instability, shame and stigma both during and post-imprisonment, as well as financial strain stemming from fines and fees incurred in court and other monetary costs associated with imprisonment and reduced job opportunities post-imprisonment. Along with being immediately damaging to physical and or mental health, the elevated chronic stress generated from incarceration during the transition to adulthood can lead to increased exposure to other health degenerating factors, which themselves worsen health later in the life course. For example, a recent study found that young adults who experienced incarceration were more likely to engage in unhealthy, stress-inducing behaviours (e.g. smoking cigarettes or unhealthy binge eating) than their counterparts who had no prior experience with incarceration (Porter, 2014). This poor health behaviour can lead to immediate poor physical health, as well as poor health behaviours later into adulthood. Because poor health behaviours later into adulthood can beget poor adult health, the elevated chronic stress brought about by incarceration at this life stage could function to damage health far after incarceration has occurred (Haas, 2007; Hale, Bevilacqua & Viner, 2015). Given the potentially severe, negative health consequences that incarceration during the transition to adulthood could have on long-term health, more work is needed to increase our understanding of the relationship between incarceration at this transitional stage and health at this specific point in early adulthood.

Capturing or Estimating the Effect of Incarceration on Health

The likelihood of coming into contact with the criminal justice system is not stochastic; individuals from marginalised groups are more at risk of being incarcerated than their more advantaged counterparts (Western & Pettit, 2010). Similarly, the risk of poor health is not uniformly distributed across the U.S. population; many marginalised individuals are more at risk of experiencing chronic poor health and disease compared to their more advantaged peers (Phelan, Link, & Tehranifar, 2010; Phelan & Link, 2015; Williams, Mohammed, Leavell & Collins, 2010).

These two points suggest that the fundamental problem in obtaining unbiased estimates of incarceration's health effect lies in the possibility that the health disadvantage that former inmates face is not a consequence of incarceration, but rather, is a reflection of inequalities in their pre-incarceration conditions (Schnittker et al., 2012). Parsing the degree to which prior incarceration *influences* health from the multitude of factors that occur before incarceration and health, and are associated with both incarceration and health (i.e. *confound the association*) is especially important. If incarceration is saddling individuals with poor health, policy interventions that address experiences during and post-incarceration would be likely to improve the health of former inmates. If instead, the poor health of previously incarcerated individuals is wholly a reflection of features that come before incarceration occurred, interventions that address incarceration and its consequences would do little to enhance the health of those who have experienced incarceration. We expect that both phenomena are operating. Criminal justice contact during the transition to adulthood likely has a lasting impact on health, but some of this association may be explained by the fact that marginalised populations have both higher risks of incarceration, and higher risks of poor health (Phelan & Link, 2015). This means that the association between incarceration and health will remain sizable and significant, but not as strong once adjustments are made for confounding variables.

Identifying the health consequences of incarceration is difficult because incarceration does not lend itself to random assignment. To study the health consequences of prior incarceration, it is necessary to compare the health of individuals with and without a history of incarceration in observational data, where the processes of how they came to their level of health, and how they came to their incarceration history are largely unknown, ineffectively modelled, and yet highly intertwined (Schnittker & John, 2007). Given that the best course of action for improving former inmates' health is dependent on the degree to which incarceration *itself* influences health, more research is needed which uses observational data to explore the potentially causal relationship between incarceration and health. Our innovation is to use a statistical learning technique that selects and flexibly models a subset of important confounders from a large array of candidates, to compare the health of individuals who had experienced incarceration, with their health had they never experienced incarceration, and obtain less biased approximations of incarceration's health influence. The method is likely superior to standard regression approaches that include relevant controls, or instrumental regression approaches where there are, at best, weak instruments available to isolate incarceration's specific effects.

Data and methods

Analytical framework

To estimate the effect that incarceration during the transition to adulthood has on health we would like to observe the following quantity:

$$\frac{1}{\sum_{i} A_i} A_i \left(E(Y_i \mid A_i = 1, C_i) - E(Y_i \mid A_i = 0, C_i) \right) \quad (1)$$

In Equation 1, i index*es* individuals (1 to n), A is an indicator of whether an individual had previously been incarcerated (i.e. $A = 1$ if a youth had been incarcerated before their health was measured, and $A = 0$ if otherwise), Y gives an individual's health, and C is a vector of each individual's confounders (Ho, Imai, King & Stuart, 2007).

This quantity, or *the average treatment effect on the treated* (ATT), represents a comparison of scenarios. The left-hand side of the minus sign (i.e. $E[Y_i \mid A_i = 1, C_i]$) gives the average level of health among a group of individuals (with C pre-incarceration traits) who were observed to have experienced incarceration. The right-hand side of the equation (i.e. $E[Y_i \mid A_i = 0, C_i]$) gives an estimate of the average health of the *same* group of individuals (with the same C background traits) had they never experienced incarceration. Because individuals and their background characteristics are held constant in the scenarios being compared, the difference given by the ATT represents a counterfactual of *how different the health of individuals who had experienced incarceration would be had they actually never experienced incarceration.*

Note that a more complete C, and more accurately modeled Equation 1, leads to a less biased estimate of incarceration's influence on health; the more relevant confounding factors that are held constant across counterfactuals, and the more accurately modeled each counterfactual, the higher the probability is that the difference in $E[Y_i \mid A_i = 1, C_i]$ and $E[Y_i \mid A_i = 0, C_i]$ is attributable to a change in prior incarceration status. To minimize bias in our calculation of Equation 1, we use a rich data set (The National Longitudinal Study of Adolescent to Adult Health) to flesh out C, and a machine learning approach (Bayesian Additive Regression Trees) to accurately specify the models needed to estimate Equation 1's counterfactual states.

Data: The National Longitudinal Study of Adolescent to Adult Health

We used data from the National Longitudinal Study of Adolescent to Adult Health (Add Health), which is an ongoing nationally representative, school-based study of adolescents in grades 7 to 12 that began in 1994 (Harris, et al., 2009). In a sample of 132 U.S. schools all students were given an in-school questionnaire. A sub-sample of 20,745 was chosen for separate in-home child and parent interviews. Add Health currently includes four waves of data: Waves 1 (W1: 1994-95); 2 (W2: 1996); 3 (W3: 2001-02); 4 (W4: 2007-08). In W4,

respondents ranged in age from 24-32 years. For our purposes the utility of the Add Health data is two-fold. First, the Add Health data contain a relatively large sample of individuals across the life-stages that we are interested in (i.e. adolescence and early adulthood). Second, these data contain multiple detailed measures of social, economic, psychological, and health-related factors. The richness of Add Health's measures allows us to select a diverse array of features for our set of confounders (*C*); as a more fully defined *C* allows for a more accurate estimation of the health effects of incarceration, this latter feature makes Add Health invaluable to our efforts.

Measurement

Analytical samples and prior incarceration

Incarceration may occur at any point over the course of the study. To ensure that we only control for confounders that occur *prior* to incarceration (and thus avoid either controlling away incarceration's influence or inducing post-treatment bias[i]), we split the data into two analytical samples (defined below) based upon the timing of an individual's first instance of incarceration (Acharya, Blackwell & Sen, 2015; Gelman & Hill, 2007).

At W4, respondents were asked, *How many times have you been in jail, prison, a juvenile detention center, or other correctional facility?* If a respondent had been incarcerated, they were then asked, *How old were you when you (first) went to jail, prison, juvenile detention center or other correctional facility?* Using this information, in conjunction with each individual's age across the Add Health waves, allows us to sort individuals into two analytical samples: individuals who were first incarcerated after W1 and prior to W2 are sorted into Sample I, those who were first incarcerated after W3 and before W4 are included in Sample II, and those individuals who had never experienced incarceration are included in both samples as controls. While we do not have exact dates of incarceration, using age at first incarceration does allow us, with reasonable accuracy, to ensure that W1 covariates occurred prior to incarceration for Sample I members, and that W1 and W3 covariates occurred prior to incarceration for Sample II individuals. To illustrate this sorting process, an individual who

reported that they were incarcerated at age 19, and was 19 years old between W1 and W2, would be sorted into Sample I. In both samples incarceration is defined as a binary variable, where a 1 indicates that an individual had been incarcerated, and a 0 indicates that an individual had never been incarcerated.

Confounders (C)

To define confounders for each analytical sample, we draw information from the Add Health waves that predate an individual's first instance of incarceration. For Sample I, this means that we only control for W1 covariates, and for Sample II, this means that we control for both W1 and W3 covariates. Specific measures are selected as confounders based on the literature of the social and economic associates of incarceration and health. For example, we include a variable measuring the degree to which an individual feels that they are connected to others at their school as: (1) social control theory suggests that low social attachment increases the probability of deviance/incarceration: (2) many health researchers report social connectedness as an associate of good health (Hawe & Shiell, 2000; Hirschi, 1969). Likewise because level of educational attainment is positively associated with health, and negatively associated with risk of incarceration, we include a measure of highest level of educational attainment in our confounder set (Mirowsky & Ross, 2003; Pettit & Western, 2004).

In total we select 93 variables from W1 for our Sample I confounder set. For Sample II, our confounders consist of the 93 variables from W1, plus 36 additional covariates from W3.[iiiii] The confounders we selected fell into the following general categories: demographic characteristics (e.g. gender; race), prior health status and behaviours (e.g. body mass index (BMI) prior to incarceration; frequency of being in ill-health) engagement in risky behaviour (e.g. frequency of drug use; frequency of fights), social connectedness (e.g. degree of closeness to individuals at school; degree of closeness to individuals in neighborhood), disposition characteristics (e.g. self-perceptions of intelligence; belief in ability to solve problems ''rationally''), parental characteristics (e.g. level of education, parental incarceration history), and contextual residential characteristics (e.g. crime rate in

residential area; poverty rate in residential Census tract). A full list of the confounders used in our analysis can be found in the supplemental material.

Health Outcomes (Y)

For this paper we examine three health outcomes: (1) an indicator of cardiovascular health that is increasing in prevalence among young adults (i.e. hypertension or raised blood pressure) (Nguyen et al., 2011), (2) a measure of general health status (i.e. excellent/very good self-reported health), and (3) a measure of mental health status (i.e. depression). These measures of general health and mental health were chosen as they have been examined in the adult population in relationship to incarceration, and also capture general wellbeing (Diamond, Wang, Holzer, Thomas & Cruse, 2001; Massoglia, 2008; Schnittker & John, 2007; Schnittker et al., 2012). Hypertension was chosen so that we may examine whether incarceration results in the immediate emergence of physiological problems that typically take more time to emerge. All health measures are obtained from W4.

In Add Health systolic and diastolic blood pressure (SBP, DBP) were recorded three times for each individual. As an estimate of true SBP and DPB, Add Health averaged the second and third reading of blood pressure. From this, hypertension was classified according to guidelines from the American Heart Association: individuals who had an average SBP greater than or equal to 140 mmHg, or an average DBP of at least 90 mmHg were classified as having hypertension (Calhoun et al., 2009). Individuals who had been told by a doctor that they had high blood pressure or hypertension, or were taking hypertension medication were also classified as having hypertension.[iv]

Next, excellent/very good self-rated health was derived from a question asking respondents, "in general, how is your health?" Available responses were, *excellent, very good, good, fair,* or *poor.* Individuals who responded with either *excellent* or *very good* were grouped into one category with the reference group including those who reported *good, fair or poor* health. The decision to divide self-rated health along the aforementioned lines was based on two factors. First, exploratory analysis showed that larger disparities existed in the higher categories of self-rated health, rather than the more commonly examined "poor" or "fair" dichotomization of self-rated health. This is particularly so for younger populations who are generally healthy and have yet to develop chronic diseases. This categorization is consistent with other inquires that have examined self-rated health among children and young adults (e.g. Braveman & Gottlieb, 2014; Mandal, Edelstein, Ma & Minkovitz, 2013).

Finally, our depression measure is a nine-item version of the Center for Epidemiological Studies depression scale (CES-D) (Eaton, Muntaner, Smith, Tien & Ybarra, 2004). Respondents were asked, on a scale from 0 to 3 (with 0 being never or rarely, and 3 being most of the time or all of the time) if they felt the following in the past seven days: *you were bothered by things that usually don't bother you; you could not shake off the blues, even with help from your family and your friends; you felt that you were just as good as other people* (for which the scale was reversed); *you had trouble keeping your mind on what you were doing; you were depressed; you were too tired to do things; you enjoyed life (also reversed); you were sad; you felt that people disliked you* (Cronbach's α = 0.80) (Boardman & Alexander, 2011). The results from each of these questions were summed, and [following the advice of Boardman & Alexander, 2011)], we classified any individual who scored over 10 as having depression. Note that our definition of depression is based on depressive symptomatology/states, and thus is not a clinical measure of depression. However, the CES-D depression scale is good measure of risk of developing clinical depression or anxiety disorders (Perreira, Deeb-Sossa, Harris & Bollen, 2005). For parsimony when we use the term depression in results, we mean depressive symptoms, in particular a depressive symptoms score of over 10.

Analysis: Bayesian Additive Regression Trees

Estimating the ATT is conceptually straightforward. To calculate this quantity we simply: 1) regress the health outcome of interest (Y) on an indicator of incarceration (A) and the set of confounders chosen from Add Health (C), 2) use the resulting model to predict $E[Y_i \mid A_i = 1, C_i]$ and $E[Y_i \mid A_i = 0, C_i]$, and 3) record the difference between the two counterfactual states. In practice, however, this process is made

difficult by the dimensionality of (or the number of variables in) the confounder set C, as well as the specification of these variables in the model (Ho et al., 2007).

To expand, recall that defining *C* as a rich, robust vector of potential confounders is theoretically desirable; as Hill, Weiss and Zhai (2011) state, "the richer the pre-treatment information we can condition on, the more we may be willing to believe that [the ATT represents a real treatment effect]." While advantageous in this manner, working with a large set of confounders to define the ATT comes with a number of empirical challenges. For instance, modeling a large number of confounders can be accompanied by an array of computational problems. Traditionally employed models, such as linear or logistic regression, often perform poorly (e.g. they provide nonsensical/unstable parameter estimates), or even fail to converge when used to describe the relationship between an outcome and a high-dimensional set of covariates (Hill et al., 2011).[v]

Moreover, model specification (or the process of deciding how covariates should be included in the model) is a major concern when working with a large set of potential confounders. Properly specified models are desirable as they lead to more accurate (i.e. less likely to mistake random noise in the data for real structure, or vice-versa) representations of reality than do their poorly specified counterparts (Green & Kern, 2012; Hill, 2011), but the probability of incorrectly specifying a model scales with the number of covariates under consideration. That is, with every additional predictor included in the model comes an increase in the number of models that could possibly be fitted, and with an increase in the number of models that could possibly be fitted comes a higher risk of choosing a model that is poorly specified (Ho et al., 2007).

Adjudicating between specifications of high-dimensional data can be a daunting task. Relying on prior knowledge to narrow the range of plausible model specifications is useful, but is often insufficient on its own (i.e. prior literature is often informative enough to allow researchers to determine which factors ought to be considered as potential confounders, but not informative enough to describe

the appropriate functional relationships between said predictors) (Ho et al., 2007).[vi] Even relying on principled model selection methods in conjunction with prior knowledge to adjudicate between models can be infeasible in high-dimensions. Approaches that search among plausible specifications for the "best fitting model" (such as stepwise or subset selection approaches) begin to become impracticable in the face of a large number of potential confounders because of the sheer number of specifications that exist in such a scenario (Green & Kern, 2012; Ho et al., 2007; James, Witten, Hastie & Tibshirani, 2015).

To alleviate the practical concerns that come with examining a high-dimensional space, we use Bayesian Additive Regression Trees (BART) to generate the model needed to estimate the ATT (Chipman, George & McCulloch, 2010). Through an iterative process this nonparametric algorithm uses an ensemble of small regression-trees (i.e. regressions that model only a very small subset of confounding variables at one time) to select a subset of variables and their specifications, and produce an accurate yet generalizable model representation of the data (Hill, 2011). That BART is able to produce stable, coherent estimates in face of a (very) large number of covariates, and is adept at letting the data decide upon a model specification (that strikes a balance between accuracy and parsimony) means that we are able to largely avoid the computational and specification challenges that often accompany modeling a large set of covariates. Thus, our machine-learning approach increases our confidence that our models account for confounders in a way that makes the ATT representative of a change in health due to incarceration history.

Results

Because the number of variables used in our analysis is large, we do not present a full table of descriptive statistics in the body of the paper (that information can be found here: http://bit.ly/2ac14gi). Instead, to contextualise our analytical samples, we provide summary statistics on a handful of demographic covariates in table 1:

**Table 1: Selected demographic statistics for analytical samples**

	Incarcerated Individuals		Never Incarcerated (n = 9,399)
Statistic	Sample I (n = 800)	Sample II (n = 586)	
Mean age of first incarceration	18.6	25.2	-
W4 Mean Age	28.6	28.03	28.31
Percent White	52	49	57
Percent Black	24	29	20
Percent Male	74	64	59
Percent Born in US	96	97	93
Percent with mother with a college degree	16	17	27
Median Household Income (of Census Block Group)	26,370	26,635	28,853

.

From table 1 we see that the average age of first incarceration in Sample I is 18.6 years. The mean age of first incarceration for Sample II (25.2 years) is somewhat late into the transition to adulthood. At Wave 4, where our chosen health outcomes were recorded, the mean age for all groups was approximately 28 years. Males and people of colour made up a larger proportion of the sample of former inmates than they did the sample of never incarcerated individuals. Moreover, it appears as if formerly incarcerated individuals were from more marginalised socioeconomic groups than individuals who had never been incarcerated; for instance, the proportion of never-incarcerated individuals who had a mother with a college degree was approximately 10% higher than the proportion of former inmates who similarly educated mothers.

Next, we examine how the health outcomes chosen from W4 vary by incarceration history. Figure 1 plots the difference in the proportion of previously incarcerated individuals with each health outcome and the proportion of never incarcerated individuals with each health outcome:
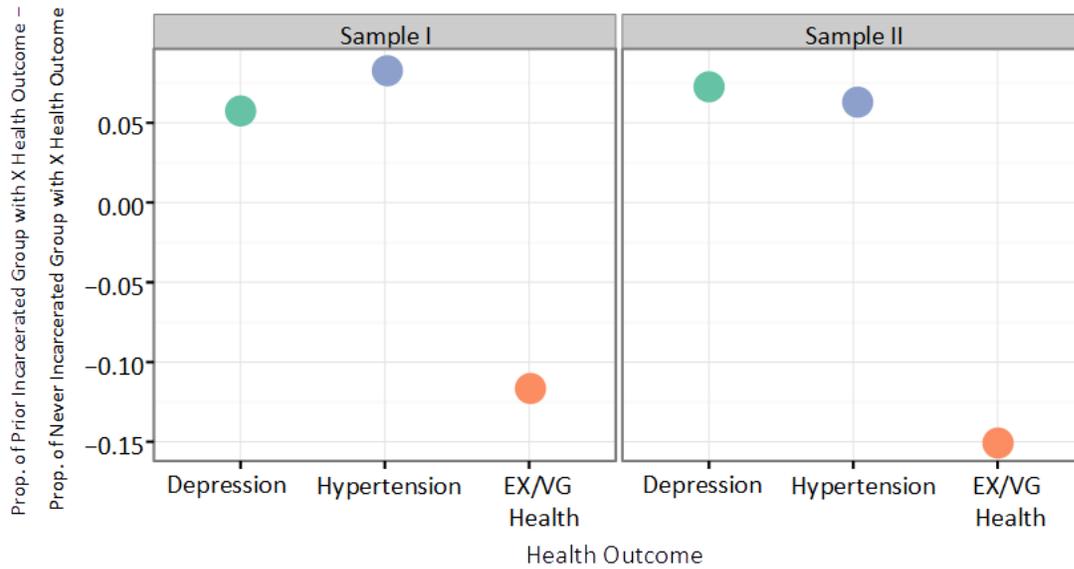
Figure 1: Difference in the proportion of previously incarcerated individuals with each health outcome and the proportion of never incarcerated individuals with each health outcome. Note that EX/VG Health indicates excellent/very good health.

As we can see from figure 1, individuals with a history of incarceration were in worse health than their counterparts who had never been incarcerated. In Sample I, individuals who had been incarcerated were 6% more likely to suffer from depression, 12 points less likely to report being in excellent or very good health, and 8 points more likely to have hypertension than their peers with no history of incarceration. The inequalities in Sample II follow a similar pattern; individuals who were incarcerated between W3 and W4 were 8% more likely to be depressed, 15% less likely to report excellent or very good health, and 7% more likely to have hypertension than their counterparts who had never been incarcerated.

To examine if the health inequalities between previously incarcerated and never incarcerated individuals was a product of incarceration rather than a product of features that occur prior to incarceration, we estimate Equation 1 using the BART methodology described above.[vii] Figure 2 plots the estimated ATTs (and their 95% confidence intervals) recovered from our models:
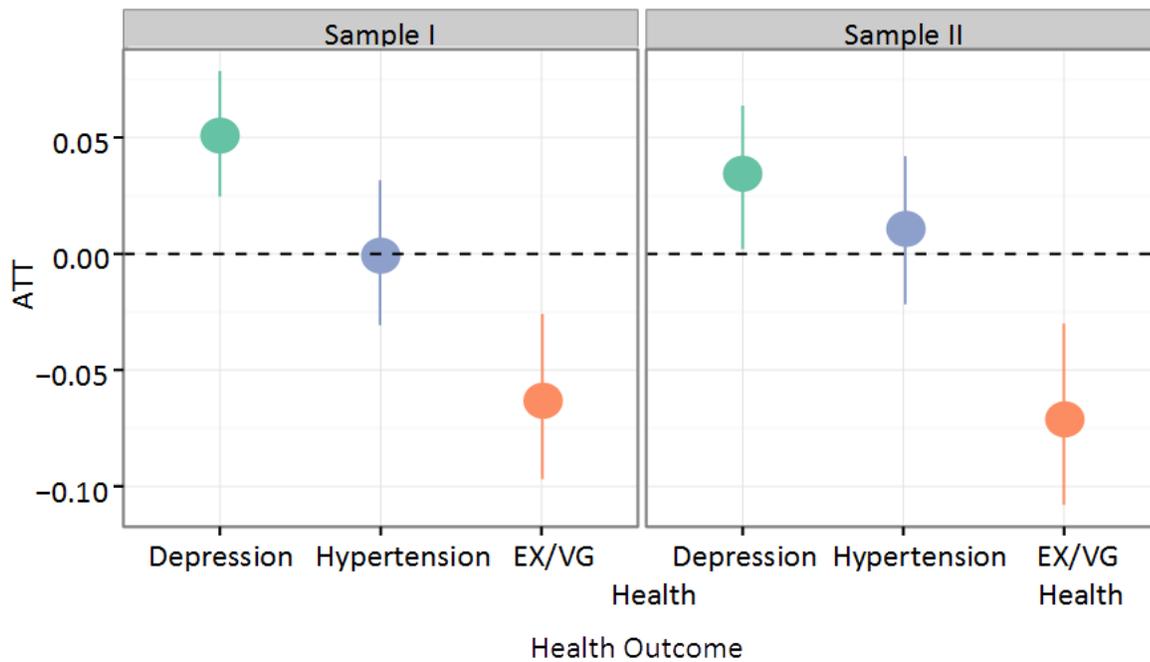
Figure 2: Estimated ATTs for each health outcome and each sample; 95% uncertainty intervals are marked.

Across analytical samples we see similar results. Net of the large set of confounders, the probability of being depressed would have been approximately 5% lower had individuals with a history of incarceration never been incarcerated. Note that this estimate is similar in magnitude to the difference in proportions given in figure 1. That this health disadvantage persists in full after making individuals equivalent on all factors except prior incarceration, suggests that the elevated risk of depression among incarcerated individuals is largely a consequence of their incarceration. Similarly, having a prior history of incarceration appears to come with a 10% decrease in the probability of reporting excellent/very good health. This estimate is nearly half of what the difference in excellent/very good self-rated health was before accounting for confounders. This suggests that, while incarceration negatively affects general health status, a large portion of the association between health status and incarceration is simply a reflection of differences in factors that occur prior to incarceration.

In contrast to the previous two findings, our results show no adverse effects of incarceration on hypertension. The relationship between prior incarceration and hypertension represented in figure 1 dissipates after accounting for confounders, suggesting that those who were incarcerated before W4 would have been just as likely to experience hypertension had they never been incarcerated.

Discussion

In this paper, we expand the current understanding of how incarceration and health relate by examining the association between incarceration and health in the transition to adulthood. Understanding how incarceration impacts health during this period in the life course has been largely unexplored. Our results provide some support for the idea that incarceration during emerging adulthood is harmful to health.

Taking account of a large number of features that occur prior to incarceration, experiences with incarceration during late adolescence and early adulthood appear to increase an individual's risk of depression, and negatively impact general health status. Given that health trajectories begin to solidify in young adulthood, these findings, which show incarceration effects early in adulthood, suggest that early confinement may have implications for health much later in the life course.

Our results, however, suggest that experiences with incarceration during this period have no discernible effect on risk of hypertension. The results for hypertension are surprising, given that earlier work suggests that incarceration significantly influences cardiovascular health (Binswanger, Krueger & Steiner, 2009; Massoglia, 2008). It may be that Add Health respondents are simply not old enough to have manifested high blood pressure after exposure to stressors. Recent estimates indicate that approximately 10% of individuals aged between 18 and 45 years suffer from hypertension, compared to 40% of those between ages 45-64 (though Add Health reports substantially higher prevalence of hypertension among young adults than other national datasets do [such as the National Health and Nutrition Examination Survey]) (Keenan & Rosendorf, 2011; Nguyen et al., 2011). In addition, young adults may be more resilient to stressors, and may be more likely to receive help from family members while confined, and after release. Since prior studies typically utilize samples of mid-age adults (such as the National Longitudinal Study of Youth 1979), this age difference may also account for the discrepancy in results. Our measure of incarceration includes both juvenile and adult confinement. Previous research indicates that these experiences differ (Loughran et al., 2010), and these differences may have different implications for health. Therefore, our unadjusted estimates of the association between confinement and health are likely to be somewhat conservative.

Our results suggest a number of directions for future work.

First, the relationship between incarceration during the transition to adulthood and health should continue to be explored by incorporating additional measures of mental and physical health as outcomes (including clinical measures of depression), and by empirically identifying the mechanisms that propagate this relationship. Our findings signal that incarceration is an important stressor that can have serious implications for life course outcomes even in the early stages of adulthood. Consequently, research should continue to increase our understanding of how incarceration in this emergent life stage matters for health.

Second, future researchers might use BART to examine other well-known consequences of incarceration (e.g. children's education outcomes, family dissolution, employment opportunities) (Lopoo & Western, 2005; Turney & Wildeman, 2013; Western, 2002). As with the association between incarceration and health, the associations among incarceration and many of these alternative outcomes are likely confounded by a large number of preexisting features. Therefore, using BART to isolate incarceration's influence on these outcomes could provide more exact insight into the impact that incarceration has on individuals' lives.

Third, future research should continue to make use of cutting-edge methods to examine incarceration's health effect. Though BART is useful in that it can coherently account for many confounding forces, it is limited in that it can only control for measured covariates. While we believe that our set of confounders is quite comprehensive, it is possible that we have not controlled for all of the features that jointly influence health and incarceration. As new methodologies that leverage observational data for causal inference become available, researchers should implement them, and compare their results with ours. A catalogue of results, produced via methodologies with varying strengths and weaknesses, would allow us to better quantify our confidence in estimates of incarceration's effect on health (e.g. if estimates of this quantity are similar across studies of varying approaches, we would be confident that our understanding of how incarceration and health relate is not a by-product of bias that comes from drawing causal inferences from observational data).

Fourth, investigating if (and how) the effect of incarceration on health varies across individuals is an important next step. Though data limitations kept our study from fully examining this point, incarceration is not a unidimensional or simple dichotomous variable. Indeed, the experience of incarceration can vary tremendously: while some incarcerated individuals are exposed to rather severe conditions (such as solitary confinement, over-crowded living spaces, or persistent physical abuse), others are confined in much less (overtly) harmful settings (e.g. low-security, rehabilitation-oriented facilities) (Kifer, Hemmens, & Stohr, 2003; Smith, 2006; Wolff, Blitz, Shi, Siegel & Bachman, 2007). This variation in incarceration experience may well lead to variation in incarceration's health effect (e.g. one might predict that individuals who have been exposed to long spells of solitary confinement while incarcerated might be more harmed than individuals who never experienced solitary confinement while detained). Additionally, a dose-response relationship (where the adverse health consequences of incarceration become more severe the longer an individual is incarcerated) may exist. So future research should focus on identifying the degree to which incarceration's health effect depends on the conditions under which an individual was incarcerated.

To this end, we believe that examining the association between juvenile incarceration and health in other countries would be productive. In the U.S. contact with the juvenile justice system does not necessarily serve to successfully rehabilitate or reintegrate individuals back into society. This may not be the case in other contexts, where incarceration may be a necessary, therapeutic family-level intervention. In these more rehabilitative contexts, incarceration's effect on health may prove to be less negative. As this is the case, the association between incarceration and health may vary across contexts, and should be explored in more depth (Andrews & Bonta, 2010; Jones & Newburn, 2005). Finally, a recent review of evidence by the National Academy of Sciences and the Institute of Medicine has determined that the U.S. population, on average, has lower life expectancies and higher rates of disease and injury compared to people in other high-income countries (Institute of Medicine, 2013). Moreover, this health disadvantage can be evidenced at all ages, and the high rates of illness and death along those younger than age 50 years is particularly troubling (Hummer & Lawrence, 2015). The reasons for these troubling health disadvantages are multifactorial. However, it is likely that the high rates of incarceration for both youth and adults in the U.S. may also be driving some of these population differences. Future research should examine the important role incarceration may play as a driver of population health differences between the U.S. and other high-income countries, particularly at young ages.

Acknowledgments

References

Alwin, D., & Wray, L. (2005). A life-span developmental perspective on social status and health. *Journal of Gerontology: Series B, 60*(2), S7-S14. https://doi.org/10.1093/geronb/60.Special_Issue_2.S7

Acharya, A., Blackwell, M., & Sen, M. (2015). Detecting direct effects and assessing Alternative mechanisms. Under review. Accessed 15 March 2014 from http://j.mp/1Ee6Eaq.

Annie E. Casey Foundation. (2013). *Reducing youth incarceration in the United States: KIDS COUNT data snapshot: youth incarceration in the United States.* Accessed 14th March   2015 from http://tinyurl.com/ofv2f4u.

Andrews, D., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy and Law, 16*(1), 39-55. https://doi.org/10.1037/a0018362

Ben-Shlomo, Y., and Kuh, D. (2002). A life course approach to chronic epidemiology: conceptual models, empirical challenges, and interdisciplinary perspectives. *International Journal of Epidemiology, 31,* 286-293. https://doi.org/10.1093/ije/31.2.285

Boardman, J., & Alexander, K. (2011). Stress trajectories, health behaviors, and the mental health of black and white young adults. *Social Science and Medicine, 72*(10): 1659-1666. https://doi.org/10.1016/j.socscimed.2011.03.024

Binswanger, I., Krueger, P., & Steiner, J. (2009). Prevalence of chronic medical conditions among jail and prison inmates in the United States compared with the general population. *Journal of Epidemiology and Community Health, 63*(11), 912-919. https://doi.org/10.1136/jech.2009.090662

Braveman, P., & Gottlieb, L. (2014). The social determinants of health: it's time to consider the causes of the causes. *Public Health Reports, 129(*Suppl., 2), 19–31.

Cable, N. (2014). Life course approach in social epidemiology: an overview, application and future implications. *Journal of Epidemiology, 24*(5), 347-352. https://doi.org/10.2188/jea.JE20140045

Calhoun, D., Jones, D., Textor, S., Goff, D., Murphy, T., Toto, R., White, A., Cushman, C., White, W., Sica, S., Ferdinand, K., Giles, T., Falkner, B., & Carey, R. (2009). Resistant hypertension: diagnosis, evaluation, and treatment: a scientific statement for high blood pressure research from the American Heart Association Professional Education Committee of the Council for High Blood Pressure Research. *Hypertension, 51,* 1403-1419. https://doi.org/10.1161/HYPERTENSIONAHA.108.189141

Chipman, H., George, E., & McCulloch, R. (2010). BART: Bayesian Additive Regression Trees. *Annals of Applied Statistics, 4,* 266-298. https://doi.org/10.1214/09-AOAS285

Cutler, D., & Lleras-Muney, A. (2006). Education and health: evaluating theories and evidence. NBER Working Paper 12352. National Bureau of Economic Research, Chicago, IL. https://doi.org/10.3386/w12352

Diamond, P., Wang, E., Holzer, C., Thomas, C., & Cruse, A. (2001). The prevalence of mental illness in prison. A*dministration and Policy in Mental Health and Mental Health Services Research, 29*(1), 21-40. https://doi.org/10.1023/A:1013164814732

Eaton, W., Muntaner, C., Smith, C., Tien, A., & Ybarra, M. (2004). Center for Epidemiologic Studies Depression Scale: Review and revision (CESD and CESD-R). In M. Maruish (Ed.), The use of psychological testing for treatment planning and outcomes assessment. (pp. 363-   377). Mahwah, NJ: Lawrence Erlbaum.

Glaze, L., & Parks, E., (2011). Correctional populations in the United States, 2011. U.S. Bureau of Justice Statistics. NCJ 239972.

Geller, A., & Curtis, M. (2011). A sort of homecoming: incarceration and the housing security of urban men. *Social Science Research, 40*(4), 1196-1213. https://doi.org/10.1016/j.ssresearch.2011.03.008

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* New York: Cambridge University Press.

Golzari, M., Hunt, J., & Anoshiravani, A. (2006). The health status of youth in juvenile detention facilities. *Journal of Adolescent Health, 38,* 776-782. https://doi.org/10.1016/j.jadohealth.2005.06.008

Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2014). Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics, 24(*1),

Green, D., & Kern, H. (2012). Modeling heterogeneous treatment effects in survey experiments with Bayesian Additive Regression Trees. *Public Opinion Quarterly, 76*(3), 491-511.

Haas, S. (2007). The long-term effects of poor childhood health: an assessment and application of retrospective reports. *Demography, 44*(1), 113-135. https://doi.org/10.1353/dem.2007.0003

Hale, D., Bevilacqua, L., & Viner, R. (2015). Adolescent health and adult education and employment: a systematic review. *Pediatrics, 126*(1), 128-140. https://doi.org/10.1542/peds.2014-2105

Harris, K.M., Halpern, C., Whitsel, E., Hussey, J., Tabor, J., Entzel. P., & Udry, J. (2009). The National Longitudinal Study of Adolescent Health: research design. Accessed 14th March 2015 from http://www.cpc.unc.edu/projects/addhealth/design

Hawe, P., & Shiell., P. (2000). Social capital and health promotion: a review. *Social Science and Medicine. 51*(6), 871-85. https://doi.org/10.1016/S0277-9536(00)00067-8

Hazel, N. (2008). *Cross-national comparisons of youth justice.* London: Youth Justice Board. Accessed 14th March 2015 from http://tinyurl.com/o348r3k.

Hill, J. (2011). Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics, 20*(1), 217-240. https://doi.org/10.1198/jcgs.2010.08162

Hill, J., Weiss, C., & Zhai, F. (2011). Challenges with propensity score strategies in a high-dimensional setting and a potential alternative. *Multivariate Behavioral Research*, *46*(3), 477-513. https://doi.org/10.1080/00273171.2011.570161

Hirschi, T. (1969). *Causes of delinquency*. University of California Press.

Ho, D., Imai, K., King, G., & Stuart, E. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15,* 199-236. https://doi.org/10.1093/pan/mpl013

Holzer, H., Raphael, S., & Stoll, M. (2004). How willing are employers to hire ex-offenders? *Focus, 23*(2), 40-43.

Hummer, R., & Lawrence, E. (2015). *Racial/ethnic differences in early life mortality in the United States.* Poster presented at the annual meeting of the Southern Demographic Association, San Antonio, Texas.

Institute of Medicine. (2013). *U.S. health in international perspective: shorter lives, poorer health.* Accessed November 2015 from http://tinyurl.com/oe8ahzu.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2015). *An introduction to statistical learning: with applications in R.* New York: Springer.

Jones, T., & Newburn T. (2005). Comparative criminal justice policy-making in the United States and the United Kingdom. *The British Journal of Criminology, 45*(1), 58-80. https://doi.org/10.1093/bjc/azh067

Kapelner, A. & Bleich, A. (2016). bartMachine: Machine Learning with Bayesian Additive Regression Trees. *Journal of Statistical Software, 70*(4), 1-40. https://doi.org/10.18637/jss.v070.i04

Keenan, N., & Rosendorf, K. (2011). Prevalence of hypertension and controlled hypertension - United States, 2005-2008. *MMWR Surveillance Summaries, 60*(S): 94-97.

Kifer, M., Hemmens, C., & Stohr, M. K. (2003). The goals of corrections: perspectives from the line. *Criminal Justice Review, 28*(1), 47-49. https://doi.org/10.1177/073401680302800104

Lee, H., Wildeman, C., Wang, E., Matusko, N., & Jackson, J. (2014). A heavy burden: the cardiovascular health consequences of having a family member incarcerated. *American Journal of Public Health*, *104*(3), 421-427. https://doi.org/10.2105/AJPH.2013.301504

Loughran, T. A., Mulvey, E. P., Schubert, C. A., Chassin, L. A., Steinberg, L., Piquero, A. R., Cota-Robles, S., Fagan J., Cauffman E., & Losoya, S. H. (2010). Differential effects of adult court transfer on juvenile offender recidivism. *Law and Human Behavior*, *34*(6), 476-488. https://doi.org/10.1007/s10979-009-9210-z

Lopoo, L., & Western, B. (2005). Incarceration and the formation and stability of marital unions. *Journal of Marriage and Family, 67*(3), 721-734. https://doi.org/10.1111/j.1741-3737.2005.00165.x

Mandal, M., Edelstein, B. L., Ma, S., & Minkovitz, C. S. (2013). Changes in children's oral health status and receipt of preventive dental visits, United States, 2003–2011/2012. *Preventing Chronic Disease, 10*, E204. https://doi.org/10.5888/pcd10.130187

Maruschak, L. (2010). *HIV in Prisons, 2001-2010*. NCJ 238877. Washington D.C.: U.S. Bureau of Justice Statistics.

Massoglia, M. (2008). Incarceration as exposure: the prison, infectious disease, and other stress-related illnesses. *Journal of Health and Social Behavior, 49*(1), 56-71. https://doi.org/10.1177/002214650804900105

Massoglia, M., & Pridemore, W. (2015). Incarceration and health. *Annual Review of Sociology, 41,* 291-310. https://doi.org/10.1146/annurev-soc-073014-112326

Mirowsky, J., & Ross, C. (2003). *Education, social status, and health: social institutions and social change.* New York: Aldine de Gruyter.

Nguyen, Q., Tabor, J., Entzel, P., Lau, Y., Suchindran, C., Hussey, J., Halpern, C., Harris, K., & Whitsel, E. (2011). Discordance in national estimates of hypertension among young adults. *Epidemiology, 22*(4), 532-541. https://doi.org/10.1097/EDE.0b013e31821c79d2

Pettit, B., & Western, B. (2004). Mass imprisonment and the life course: race and class inequality in U.S. incarceration. *American Sociological Review, 69*(2), 151-169. https://doi.org/10.1177/000312240406900201

Perreira, K. M., Deeb-Sossa, N., Harris, K. M., & Bollen, K. A. (2005). What are we measuring? An evaluation of the CES-D across race/ethnicity and immigrant generation. *Social Forces 83*(4), 1567-1601. https://doi.org/10.1353/sof.2005.0077

Pervanidou, P., & Chrousos, G. (2012). Metabolic consequences of stress during childhood and adolescence. *Metabolism, 61*(5), 611-619. https://doi.org/10.1016/j.metabol.2011.10.005

Phelan, J., Link, B., & Tehranifar, P. (2010). Social conditions as fundamental causes of health   inequalities: theory, evidence, and policy implications. *Journal of Health and Social Behavior, 51*(1): S28-S40. https://doi.org/10.1177/0022146510383498

Phelan, J., & Link, B. (2015). Is racism a fundamental cause of inequalities in health? *Annual Review of Sociology, 41*, 311-330. https://doi.org/10.1146/annurev-soc-073014-112305

Porter, L. (2014). Incarceration and post-release health behavior. *Journal of Health and Social Behavior, 55*(2), 234-249. https://doi.org/10.1177/0022146514531438

Robards, J., Evandrou, M., Falkingham, J., & Vlachantoni, A. (2012). Marital status, health and morality. *Maturitas, 73*(4): 295-299. https://doi.org/10.1016/j.maturitas.2012.08.007

Roettger, M., & Boardman, J. (2012). Parental incarceration and gender-based risks for increased body mass index: evidence from the National Longitudinal Study of Adolescent Health in the United States. *American Journal of Epidemiology*, *175*(7), 636–644. https://doi.org/10.1093/aje/kwr409

Rose, D. R., & Clear, T. R. (1998). Incarceration, social capital and crime: implications for social disorganization theory. *Criminology*, *36*(3), 441-480. https://doi.org/10.1111/j.1745-9125.1998.tb01255.x

Settersten Jr, R., Furstenberg, F., & Rumbaut, R. (2005). *On the frontier of adulthood: theory, research and public policy.* Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226748924.001.0001

Schnittker, J., & Bacak, V. (2013). A mark of disgrace or a badge of honor? Subjective status among former inmates. *Social Problems, 60*(2), 234-254. https://doi.org/10.1525/sp.2013.60.2.234

Schnittker, J., & John, A. (2007). Enduring stigma: the long-term effects of incarceration on health. *Journal of Health and Social Behavior*, *48*(2), 115-130. https://doi.org/10.1177/002214650704800202

Schnittker, J., Massoglia, M., & Uggen, C. (2012). Out and down: incarceration and psychiatric disorders. *Journal of Health and Social Behavior, 53*(4), 448-464. https://doi.org/10.1177/0022146512453928

Smith, P. S. (2006). The effects of solitary confinement on prison inmates: a brief history and review of the literature. *Crime and Justice, 34*(1), 441-528. https://doi.org/10.1086/500626

Turney, K., & Wildeman, C. (2013). Redefining relationships: explaining the countervailing consequences of parental incarceration for parenting. *American Sociological Review, 78*(6), 949-979. https://doi.org/10.1177/0003122413505589

Walmsley, R. (2013). *World prison population list*, 10th edition. International Centre for Prison Studies. Accessed September 2015 from http://tinyurl.com/qc9ktdc.

Western, B. (2002)*.* The impact of incarceration on wage mobility and inequality. *American Sociological Review, 67*, 526-546. https://doi.org/10.2307/3088944

Western, B. (2006)*. Punishment and inequality in America.* New York: Russell Sage.

Western, B., & Pettit, B. (2010). Incarceration and social inequality. *Daedalus, 139*(3), 8-19. https://doi.org/10.1162/DAED_a_00019

Western, B., Kling, J. R., & Weiman, D. F. (2001). The labor market consequences of incarceration. *Crime & Delinquency*, *47*(3), 410-427. https://doi.org/10.1177/0011128701047003007

Williams, D., Mohammed, S., Leavell, J., & Collins, C. (2010). Race, socioeconomic status and health: complexities, ongoing challenges and research opportunities. *Annals of New York Academy of Sciences, 1186*, 69-101. https://doi.org/10.1111/j.1749-6632.2009.05339.x

Wildeman, C., Schnittker, J., & Turney, K. (2012). Despair by association? The mental health of mothers with children by recently incarcerated fathers. *American Sociological Review 77*(2), 216-243.

Wolff, N., Blitz, C., Shi, J., Siegel, J., & Bachman, R. (2007). Physical violence inside prison: rates of victimization. *Criminal Justice and Behavior, 34,* 588-599. https://doi.org/10.1177/0093854806296830

Endnotes

i Post-treatment bias is, in part, a form of selection bias that occurs when conditioning on variables that are affected by the treatment of interest (i.e. incarceration, in our case), and affect the outcome of interest (i.e. health) (Acharya et al., 2015). This form of bias can be extremely conservative or anti-conservative, and can leave estimates with no substantive, causal interpretation. So avoiding this form of bias is paramount to our efforts. For a more detailed overview of post-treatment bias see Acharya et al. (2015).

ii Note that these additional 36 variables allow for Sample II's estimates to be at least as accurate as Sample I's estimates. For example, self-rated health at W1 is controlled for in Sample I and Sample II, where self-rated health at W3 is controlled for only in Sample II (given that Sample I individuals had already experienced incarceration by the time their W3 health was recorded). If W3 self-rated health were not controlled for in Sample II, but W1 self-rated health was still controlled for in Sample I, our estimates for Sample II would not account for an individual's general health status immediately before they were incarcerated, but our Sample I estimates would. This may lead to more biased results for Sample II.

iii Because the size of the confounder set is so large, a rationale explaining each variable's inclusion is beyond the scope of this paper.

iv While hypertension is not often analyzed as an outcome in young adult populations, a remarkable number of individuals in Add Health are coded as having hypertension (approximately 20% of individuals at W4) (Nguyen et al., 2011). Because the proportion of individuals with hypertension is quite sizable in these data, and because (as will be discussed later) the prevalence of hypertension is elevated among the observed groups of individuals with a history of incarceration, we believe that it is useful to analyze hypertension as an outcome in this manuscript.

v These concerns manifest themselves in our analysis: logistic regression models of our health outcomes (in which all covariates were specified in a simple, additive fashion) either failed to converge, or failed to identify some parameter estimates when the full set of confounders were included as covariates.

vi For instance, while prior research strongly suggests that income, educational attainment, and race are confounders of incarceration's health effect (e.g. Mirowsky & Ross, 2003; Western & Pettit, 2010), it provides little guidance on how these factors ought to enter into a model predicting health as a function of incarceration. Questions such as, "does incarceration's health effect vary by pre-incarceration income" (i.e. does the model need an interaction between incarceration and income?), or "does the degree to which incarceration's negative health effect vary by education vary by race" (i.e. does the model need a three-way interaction between incarceration, race, and education) are not covered in detail in existing work.

vii Because BART is a nonparametric, "black-box" machine-learning approach, we are not provided parameter estimates for our models (Goldstein, Kapelner, Bleich, & Pitkin, 2014; Green & Kern, 2012). To establish estimates of what variables matter and to gain a (rough) sense of the structure of each BART model, we calculate each variable's *inclusion proportion.* This quantity measures the number of times a specific variable was used in BART's tree models, divided by the total number of variables used in all of the BART's tree models (Chipman et al., 2010; Kaplener & Bleich,2016) The more often a variable is used in predicting the response, the higher its inclusion proportion will be.

   Plots of the top 20 variables by inclusion proportion (for each model) are included in the supplemental materials. Note that, for the Sample I and Sample II self-rated health and depression models, social-psychological features (e.g. assessment of own abilities, frequency of "feeling down", or assessment of relationship with their peers) have relatively high variable inclusion proportions. Also note that factors measuring engagement in deviant behaviour (e.g. counts of times that an individual has stolen something, has 'graffitied', and (in Sample II) was a victim of a violent crime) have relatively large inclusion proportions in our models of hypertension.

What young English people do once they reach school-leaving age: A cross-cohort comparison for the last 30 years

Jake Anders　　　Department of Learning and Leadership, UCL Institute of Education,
jake@jakeanders.uk　University College London, UK
Richard Dorsett　LLAKES Centre for Research on Learning and Life Chances, UCL Institute of Education,
　　　　　　　　University College London; National Institute of Economic and Social Research, UK

Abstract

This paper examines how young people's early transitions into the labour market have changed between cohorts born in 1958, 1970, 1980, and 1990. We use sequence analysis to characterise transition patterns and identify three distinct pathways in all cohorts. An 'Entering the Labour Market' group has declined significantly in size (from 91% in the earliest cohort, to 37% in the most recent), an 'Accumulating Human Capital' group has grown in its place (from 4% to 51%), but also a 'Potentially Difficult Transition' group has grown alongside this, reaching 12% in the most recent cohort. These trends appear to reflect behavioural rather than compositional changes. Females and those who are from a non-white ethnic background have gone from being more likely to be in the 'Potentially Difficult Transition' group, to being less likely. Coming from a low socioeconomic status background has remained a strong predictor of having a transition of this type across all four cohorts. These early transitions are important, not least since we show they are highly predictive of longer-term outcomes.

Keywords

School to work transitions, cross-cohort comparison, sequence analysis

Introduction

In recent years, there has been growing concern about the number of young people failing to make a successful transition from education into employment. Increasingly, this appears to be a structural, rather than cyclical, problem. We see evidence of this from that fact that although youth unemployment in the UK was falling in the late 1990s and early 2000s, it started rising again as early as 2004, long before the general downturn in the economy (OECD, 2008). This is an important issue, not least because making a successful transition from education into the labour market is important for young people's long-term economic success; periods of unemployment during these early years may have long-term scarring effects on later employment and earnings prospects (Arulampalam, 2001; Gregg, 2001; Gregg & Tominey, 2005).

In this paper, we examine how early transitions have changed over the last thirty years. We focus in particular on the group of young people whose early experiences suggest that they are making a potentially difficult transition in the sense that they neither continue in education nor do they find stable employment. We assess the changing size of this group and examine the extent to which it is possible to predict, on the basis of characteristics at the time of reaching school-leaving age, which individuals may experience a potentially difficult transition from school to work.

Our approach is to use sequence analysis (Abbott, 1995) to quantify the similarity between individuals' transitions over a period of 29 months from the September following their 16th birthday. Previous research has shown that young people's transitions into work may be highly differentiated (Fergusson, Pye, Esland, McLaughlin, & Muncie, 2000). Sequence analysis provides a means of comparing the full detail of individuals' labour market trajectories. This permits a fuller comparison than the more usual methods of studying labour market states at single point in time (Andrews & Bradley, 1997) or specific changes in states (Berrington, 2001). In taking this approach, we build on previous research that uses sequence analysis to study young people's transitions from education into the labour market (Anyadike-Danes & McVicar, 2005, 2010; Dorsett & Lucchino, 2014; Halpin & Chan, 1998; Martin, Schoon, & Ross, 2008; Quintini & Manfredi, 2009).

The major contribution of this paper is that it uses detailed survey data based on four birth cohorts, each roughly a decade apart. Previous research (Schoon, McCulloch, Joshi, Wiggins, & Bynner, 2001) has examined how transitions have changed between individuals born in 1958 and individuals born in 1970. We extend this to include also individuals born in 1980 and individuals born in 1990. This provides a major update to the existing empirical literature. By focusing on these more recent cohorts, we are able to consider individuals for whom the school to work transitions are relatively recent (at the time of writing). More specifically, the transitions of the 1980 and 1990 cohorts took place in the 1996-1999 and 2006-2009 periods, respectively, while the transitions of the 1970 and 1958 cohorts took place in the 1974-77 and 1986-89 periods, respectively. Ours is the first study to conduct cross-cohort analysis using sequence analysis over such an extended period and, by using more recent data, the results are more closely related to the present-day labour market.

This paper proceeds as follows. In the next section we describe the datasets used in this analysis. In the third section, we describe our methodological approach. Our analysis results in the identification of a typology of transition pathways, discussed in the fourth section, along with an account of how pathways have changed over time. We also examine the extent to which it is possible to use characteristics at age 16 to predict which type of transition young people will experience, focusing especially on those who make a potentially difficult transition (in the sense of not being characterised by either education participation or stable employment). We consider the extent to which these relationships have changed over the four cohorts analysed in this paper. In the fifth section, we extend the horizon over which individuals' transitions are considered, examining the extent to which the early transitions that we have considered are predictive of longer-term transitions, up to approximately age 24. The sixth section concludes.

Data

Our analysis uses information on month-by-month transitions for young people in England, starting in the September following their 16th birthdays and continuing for 29 months. This is

dictated by the nature of the available data, which are drawn from four birth cohort surveys:

- The National Child Development Study (NCDS) is a longitudinal survey of all individuals born in one week in 1958. Background variables (used later to predict transitions) were taken from interviews with the participant and their parents at age 16 (NCDS Sweep 3, 1974) and activity histories assembled using recall interviews at age 23 (NCDS Sweep 4, 1981). The analysis sample has around 6,000 individuals.

- The British Cohort Study (BCS) is a longitudinal survey of individuals born in 1970. Background variables were taken from interviews with the participant and their parents at age 16 (BCS Sweep 4, 1986). Activity histories were assembled primarily using recall interviews at age 26 (BCS Sweep 5, 1996). The analysis sample contains around 8,600 individuals.

- Cohort 8 of the Youth Cohort Study (YCS) is a longitudinal survey of individuals born in 1980. Background variables were taken from interviews with the participant at age 16 (YCS Cohort 8, Sweep 1, 1996), who also provided information about their parents. Activity histories were constructed using annual interviews between ages 17 and 19. The analysis sample has around 8,700 individuals.

- The Longitudinal Study of Young People in England (LSYPE) is a longitudinal survey of individuals born in 1989-90. Background variables were taken from interviews with the participant and their parents up to and including age 16 and activity histories were constructed using annual interviews between ages 17 and 19 (LSYPE Waves 1-5, 2005-2010). The analysis sample has around 9,350 individuals.

Our methodological approach requires complete, month-by-month activity histories without any gaps. With the YCS and LSYPE, activity histories were provided with the dataset. However, with the NCDS and BCS, these needed to be constructed using the recall questions about young people's activities, along with their start and end dates. Constructing these histories required some data cleaning. This involved reconciling overlapping activity spells and, in the case of the NCDS, imputing education as the status where this was not recorded in the data but could be safely assumed.[i] Furthermore, we avoided

dropping individuals missing a small number of months' activities by filling in gaps where activity status was unknown. Where there was a gap of a single month, this was imputed to have the same status as the subsequent month. Where there was a gap of two months and the same activity was recorded before and after the gap, the missing two months were imputed to also have that same status.

While these steps reduced the number of observations that were dropped, there was still some loss of sample. In the NCDS, there are partial activity histories for 9,697 individuals but the analysis (i.e. full activity history) sample is only 8,356. For the BCS, there was a partial history for 9,760 individuals but the analysis sample is 9,518. In the YCS, the respective figures are 9,265 and 8,682; and in the LSYPE they are 9,371 and 9,347.[ii] As is inevitable in longitudinal surveys, our sample is also affected by attrition. We deal with this using an inverse probability weighting strategy, applying our own weighting scheme for the NCDS and BCS, and provided weights for the YCS and LSYPE.

The monthly histories distinguish between four activity status types: employment; education; unemployment; and other inactive. The exception is the LSYPE for which unemployment and 'other inactive' are combined into a single NEET ('not in education, employment or training') status.

Other than sample loss as discussed above, the other concern with the data is that the activity histories rely on the recall of survey respondents. Paull (2002) finds evidence of recall bias in similar longitudinal data, noting that this is more likely among younger respondents and those with the most transient employment histories. As indicated above, the NCDS and BCS rely on quite long recall periods, while the problem is much reduced with the YCS and LSYPE since recall is only over the period of a single year. As such, we should bear in mind the potential increase in recorded short spells in YCS/LSYPE compared to NCDS/BCS that may be driven not by a change in behaviour, but by a change in data collection. Less worrying in a comparative sense is Paull's suggestion of bias among younger respondents; because all cohorts consider the same age group, any such bias should affect all datasets equally.

Table 1 summarises the analysis sample, showing the size of each of the cohorts, along with mean levels of those characteristics later used to predict transitions pathways. There is a good balance of the

genders in the BCS, YCS and LSYPE, but males are somewhat over-represented in the case of the NCDS, suggesting that perhaps more of the female participants have been excluded from the analysis due to missing labour market histories over the period. The proportion of the sample from a minority ethnic group also changes between the cohorts, but this seems more likely to be tracking the changing ethnic composition of the population of England over this period. Similarly, we see the increased levels of parental education across the cohorts, with a rising proportion of parents having completed a degree.

Table 1 also provides an indicator of the proportion of individuals who experience a NEET spell at some point during our period of analysis. We see that 15% of the NCDS sample experienced being NEET for at least one month. This figure drops to 10% among the BCS and then rises to 25% among the YCS and 26% among the LSYPE. Overall, therefore, we see a long-term increase in the proportion of individuals who will be NEET at some point. However, given the caution above about shorter recall periods in the YCS and LSYPE potentially increasing the reporting of short spells, we might be concerned that this is, at least in part, driven by differences in data rather than capturing a real change. In the next section, we describe how we use sequence analysis to compare individuals' transition patterns. As a preliminary comment, we note that, among other advantages, this alleviates the problem of differential reporting of short spells since a change in just one month does not greatly affect the similarity between two individuals' sequences.

Table 1. Descriptive statistics of each cohort

	NCDS	BCS	YCS	LSYPE
N	8,356	9,518	8,682	9,347
Male	0.57	0.49	0.51	0.48
Non-White	0.01	0.02	0.09	0.14
Single parent family	0.07	0.04	0.15	0.25
Parent has A Levels (no degree)	0.11	0.05	0.07	0.22
Parent has a degree	0.01	0.05	0.07	0.17
Home owner occupied	0.31	0.31	0.80	0.74
Home socially rented	0.42	0.05	0.15	0.19
Living in workless household	0.06	0.03	0.08	0.13
Ever NEET?	0.15	0.10	0.25	0.26

Notes: NCDS results weighted using author's own attrition weighting scheme. No weights applied to BCS analysis, as number excluded due to attrition was too small to model. YCS and LSYPE analysis weighted using dataset-provided attrition weights.

Methods

Our analytical approach involves three steps. First, we use sequence analysis to quantify dissimilarity between individuals' experiences. Second, we use these measures to identify similar-looking clusters. Third, we look at predictors of cluster membership. These steps are conducted separately for each cohort to allow more flexibility in the estimation of dissimilarity matrices and clusters; this follows the precedent set by Schoon, McCulloch, Joshi, Wiggins and Bynner (2001) and Kneale, Lupton, Obolenskaya and Wiggins (2010) in this type of work. Other issues specific to applying this method across multiple cohorts are discussed below.

Comparing individuals' transition experiences

Sequence analysis, also known as optimal matching, [iii] provides a means of quantifying the difference between activity histories (Durbin, Eddy, Krogh, & Mitchison, 1998). Most commonly, there are two broad approaches followed. [iv] The first of these is to calculate the minimum number of insertions and deletions ('indels') needed to transform one sequence into another. With monthly status data of the type used here, this means adding in, or removing, months of doing a particular activity into the sequence that the individual actually experienced. An example is shown in Figure 1 Panel A. Indels can 'warp' time (Lesnard, 2006; Martin & Wiggins, 2011); that is, the transformed histories may no longer adhere to calendar time. While this may not be an issue in some applications, it is unattractive in this case, since young people's transitions are often influenced by specific fixtures in calendar time (such as the start/end of the academic year).

Figure 1. Example of substitutions carried out to transform Sequence A into Sequence B

Panel A - Insertions and Deletions Only

Sequence A	X	X	X	Y	Z

Sequence B	~~W~~	X	X	X	Y

↑
Z
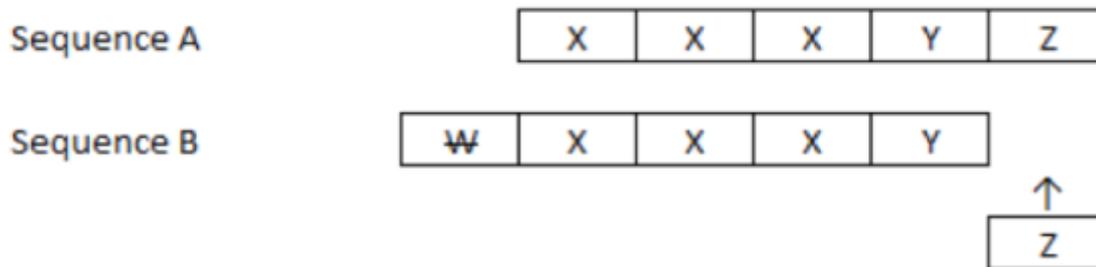
Panel B - Substitutions Only

Sequence A	X	X	X	Y	Z

Sequence B	W	X	X	X	Y

↓↑ ↓↑ ↓↑
X Y Z
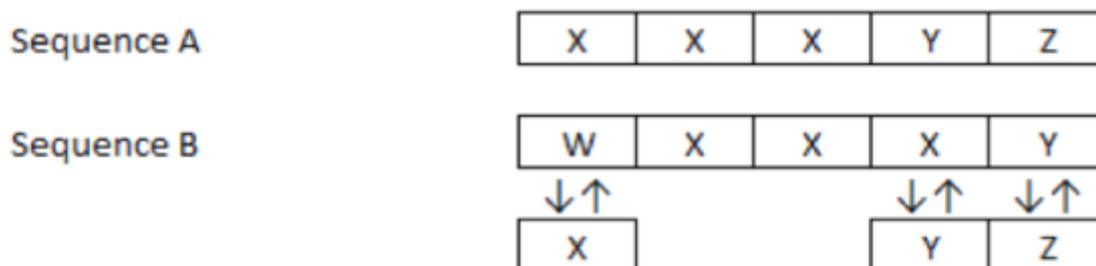
The second approach, followed here, is based on the number of substitutions from one state to another needed to transform one sequence into another. An example of this approach is demonstrated in Figure 1 Panel B. This approach maintains timelines within a sequence, but might in some circumstances exaggerate differences between sequences that are actually quite similar but slightly offset. Consider the example in Figure 1, which needs just one deletion and one insertion, compared with three substitutions.

To arrive at a measure of dissimilarity using this method, we must specify the 'cost' associated with each substitution; that is, how much of a change each type of substitution represents. Simplest would be to count the number of substitutions as the dissimilarity measure. However, some substitutions might involve a qualitatively more extreme change in status than others. Ideally, we would like the cost of substituting between two very similar states to be lower than that of substituting between two very different states. There are various ways of achieving this. In some applications, a matrix of substitution costs can be specified, based on prior knowledge of relevant differences between the states (Anyadike-Danes and McVicar, 2005, for example). However, arbitrary choice of substitution costs is one of the criticisms most often levelled at applications of sequence analysis (Wu, 2000). We adopt a more data-driven approach, using the inverse of the probability of transition between the two states being substituted. The less likely a transition between two states, the greater is the cost associated with a substitution of these two states. We allow these probabilities (and therefore costs) to vary over time, based on a moving average of the probability of transition in the months around the point in time at which the substitution is made. This method is referred to as calculating the Dynamic Hamming Distance (DHD) between two sequences (Lesnard, 2006). We implement this approach using

the TraMineR package for R (Gabadinho, Ritschard, Müller, & Studer, 2011a).

The advantage of sequence analysis is that it provides a means of measuring the differences between individuals' histories in a way that captures their full detail. The analysis in this paper focuses on one domain - the school-to-work transition - but in principle the approach could be broadened to compare individuals' experiences across multiple domains. Pollock (2007), for instance, performs an analysis of employment, housing, partnership and child-rearing experiences. The appeal of such an approach is that it addresses the likelihood that such processes are inter-related and therefore allows a richer understanding of individuals' circumstances. The focus in this paper is necessarily more narrow. While broadening to multiple domains is possible, there are two features of our study that mean doing so may be less appropriate. First, the study considers people aged 16-18, for whom housing, partnering and child-rearing play less of a role than among the broader population. Second, and related, the time period observed for each individual is relatively short (29 months) and, while there may be several transitions for the labour market domain considered, the lower incidence of changes to housing, partnering and child-rearing status would ideally be based on a longer observation period.

## Identifying groups of individuals with similar trajectories

With the dissimilarity measures, we then carried out cluster analysis in order to group together sets of sequences that were similar. Technically, we used the non-hierarchical k-medoids/Partitioning Around Medoids (PAM) method of cluster analysis (Kaufman & Rousseeuw, 1987). The non-hierarchical approach does not impose the same constraints on cluster formation as hierarchical approaches, while k-medoids rather than k-means is more robust to outliers. PAM begins by randomly choosing the requested number of 'medoids', which are actual individuals within the dataset. All other individuals are then assigned to the cluster of the medoid to which they are most similar. There is then an iterative process of swapping current individuals selected as medoids with other potential candidates, with swaps being made where this reduces within cluster variance, until no further swaps that reduce variance are available.

We started the analysis with two clusters and then repeated, adding one more cluster each time until there were twenty. In selecting which of these to use as our preferred cluster solution, we were guided by the average silhouette distance (Rousseeuw, 1987) as a primary diagnostic. This is reported for each of the requested solutions and for each of the four cohorts is shown in Figure 2. However, we also used some qualitative assessment of the sequences found within each cluster. Nevertheless, in all cases the average silhouette distance of the solution used is above 0.7, which is the 'rule of thumb' for the resulting cluster solution indicating that a strong structure has been found, suggested by Kaufman and Rousseeuw (1990, p. 88). Ultimately, we settle on seven cluster solutions in all of the datasets analysed.

Figure 2. Average silhouette distance of the cluster solutions
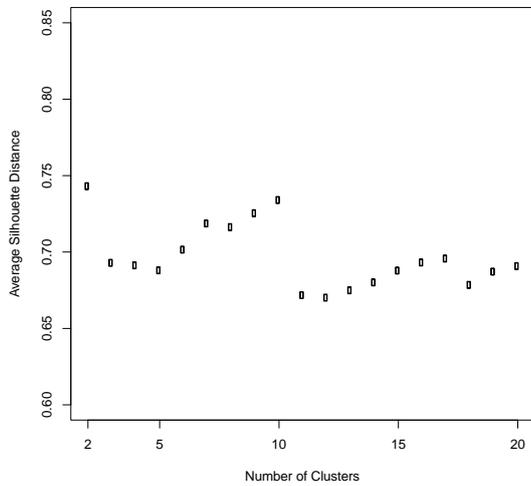
NCDS: BCS:
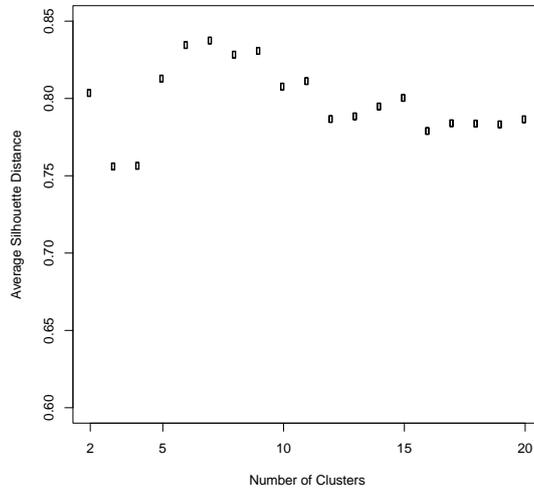
YCS: LSYPE:

Notes: Graphs report the average silhouette distance for each cluster solution from two to twenty clusters in each cohort. Graphs share common axes to allow comparison of the average silhouette distances in different d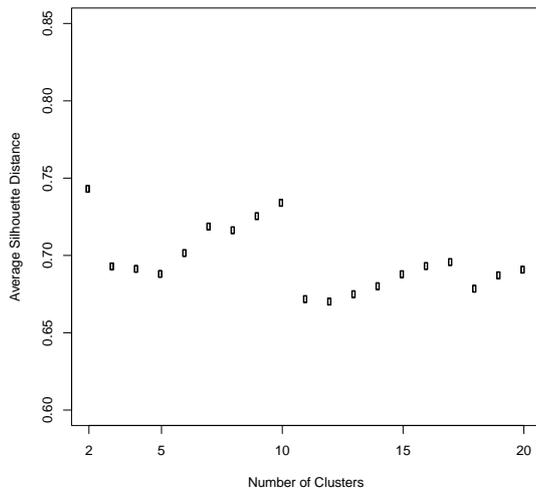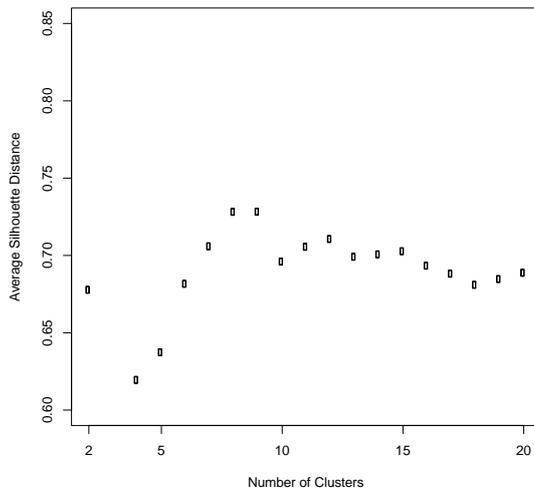atasets. A rule of thumb suggested by Kaufman and Rousseeuw (1990, p. 88) for "reasonable structure" is greater than 0.5 and for "strong structure" is greater than 0.7.

As noted above, we conduct this analysis separately for each cohort. This approach allows more flexibility for different types of cluster to emerge in the different cohorts, although if there are changes are in the prevalence of different types of transitions (rather than changes in the actual types of transitions) it is unlikely that the results would differ that much from a single cluster analysis across all cohorts. Furthermore, if there are such changes across cohorts, pooled analysis might result in not extracting a cluster that is important in one cohort but not across all four. While separate analysis could, in principle, make comparison a more difficult task, the next step of our approach helps to overcome this.

Grouping the clusters into substantive groups

In order to explore different kinds of transition, we choose to group together the seven clusters into three groupings on substantive grounds. We do this, rather than modelling the seven clusters separately or using the statistical approach of reducing the number of clusters sought, for three reasons:

1. Some of the clusters are very small and it would not be viable to conduct modelling using these.
2. The cluster analysis diagnostics indicated that the seven cluster solutions typically represented the strongest structure, while maintaining consistency across cohorts. Using a solution with a smaller number of clusters would not represent the strongest structure in the data.
3. Cluster solutions with a smaller number of clusters do not preserve separate clusters for transitions including significant periods of unemployment or economic inactivity. Instead, solutions with fewer clusters separate transitions by differences in the year in which individuals move from education to employment. This is unsurprising, since the number of individuals involved is much larger than the number of individuals with significant periods of unemployment or inactivity. However, there are clear substantive reasons for thinking it important to separate out individuals with substantial experience of unemployment or inactivity, especially as these groups do emerge in the cluster solutions that have the strongest structure, as indicated by average silhouette difference.

We appreciate that our process of grouping includes subjective decisions. However, we believe that the outcome represents a sound grouping based on both the structure of the data, indicated by the best-fitting cluster solutions, and the substantive importance of preserving a group in which individuals experience significant periods of unemployment and/or inactivity.

Predicting who will experience a pattern of transitions that may be a cause of concern

Being able to predict what kind of transition to the labour market individuals are likely to have, before this process has begun, is of clear potential use to policy makers. It potentially makes it easier to target support on those likely to be at risk of experiencing a pattern of transitions that might be a cause for concern. This work is in a similar spirit to that of Caspi, Entner Wright, Moffitt, and Silva (1998), who use childhood characteristics to model the probability of experiencing unemployment during the transition into the labour market. We use multinomial logistic regression models in order to assess how accurately age 16 characteristics that are common to the four datasets can predict outcomes.

Since the aim is to examine change between cohorts, we only make use of variables that can be derived to be comparable across cohorts. It is important to note that we make no claim that the associations found are causal (especially as there are relatively few available control variables to include in the regression models). The predictors we include in these models are gender, ethnicity (a dichotomous variable of white or non-white), highest parental education (specifically having achieved A-Levels or higher), housing tenure (specifically social renting or owner occupation), whether living with just one parent, and whether an individual's household is workless.

We estimate four separate models on the four datasets. Comparing these models provides evidence on how the roles of different predictive factors change, or remain the same, over time. In addition, we estimate a combined model on the pooled sample of all four datasets, which allows us to formally test whether the differences between these models are statistically significant from one another.

Comparing models from these different cohorts raises a kind of period-cohort identification problem. In other words, how do we interpret our

findings? Is it that the times have changed or the population has changed? We explore this by assessing the relative importance of cohort and characteristic influences by predicting group membership for members of the LSYPE cohort using the relationships estimated for members of the NCDS cohort.

Results

Cluster solutions

We categorise the seven clusters identified in each cohort's transitions into three broader groups which we label as follows:

- 'Entering the Labour Market' includes individuals who make a relatively early entry into the labour market, leaving education and finding a seemingly stable job before or within the period of analysis;
- 'Accumulating Human Capital' includes individuals who remain in education throughout the period of analysis and are, hence, likely to have received higher education ahead of their labour market entry;
- 'Potentially Difficult Transition' includes individuals whose experience includes extended periods of unemployment or economic inactivity.

As discussed above, these are different from a directly estimated three cluster solution. Nevertheless, there is always a single cluster in the directly estimated three cluster solutions very highly correlated with the 'Accumulating Human Capital' grouping. However, the 'Entering the Labour Market' grouping tends to be split into two groups, reflecting different ages of transitions from education to employment during the period. Meanwhile, individuals we group as making a 'Potentially Difficult Transition' end up spread across two or three of the clusters in a way that is not consistent across cohorts.

While clusters that fit into these three groupings exist in each of the four cohorts, the relative sizes of these groupings have changed dramatically over time. In order to show this, we present 'index plots' of young people's labour market states over the

periods considered. An index plot is a month-by-month representation of the sequence, where each horizontal line represents one young person's transition, with changes in colour showing changes in labour market state. Showing an index plot for a whole cohort is rather impenetrable, but showing plots for the clusters identified above gives a useful overview of the transitions experienced by individuals in the cluster.

Shown first, in Figure 3, is the 'Entering the Labour Market' group. These are clusters in which visual inspection reveals individuals who are either in employment throughout the period considered or who enter employment straight after education. The exception is the second LSYPE cluster which is a little more ambiguous (note that it is rather small).[v] This group has diminished significantly between the cohorts, from over 90% in the earliest to under 40% in the most recent. In addition, for those who do still follow this route, a visual inspection of the individual transitions that make up these clusters over the four cohorts suggests that earlier entry into the labour market may have become a less stable path with increasing evidence of short spells of unemployment.[vi]

Table 2 reveals the extent to which the composition of the 'Entering the Labour Market' group has changed over time. In the NCDS, the characteristics of individuals in this group essentially mirror those of the population as a whole, as one might expect for a group that makes up over 90% of the sample. However, by the BCS cohort, some differences have started to be evident. Most notably, young people whose parents' hold a degree make up only 3% of the group, compared to 5% of those in the population as a whole. The under-representation in this group of individuals with highly-educated parents persists in later cohorts too. Likewise, while there is little difference between the 'Entering the Labour Market' group and the rest of the cohort in the proportion of those who are non-white in the NCDS, by later cohorts a large gap has opened with young people with a non-white ethnic background under-represented in this group.

Figure 3. Plots of young people's individual transitions in four cohorts between the September following their 16th birthday and 29 months later: clusters placed in the 'Entering the Labour Market' group
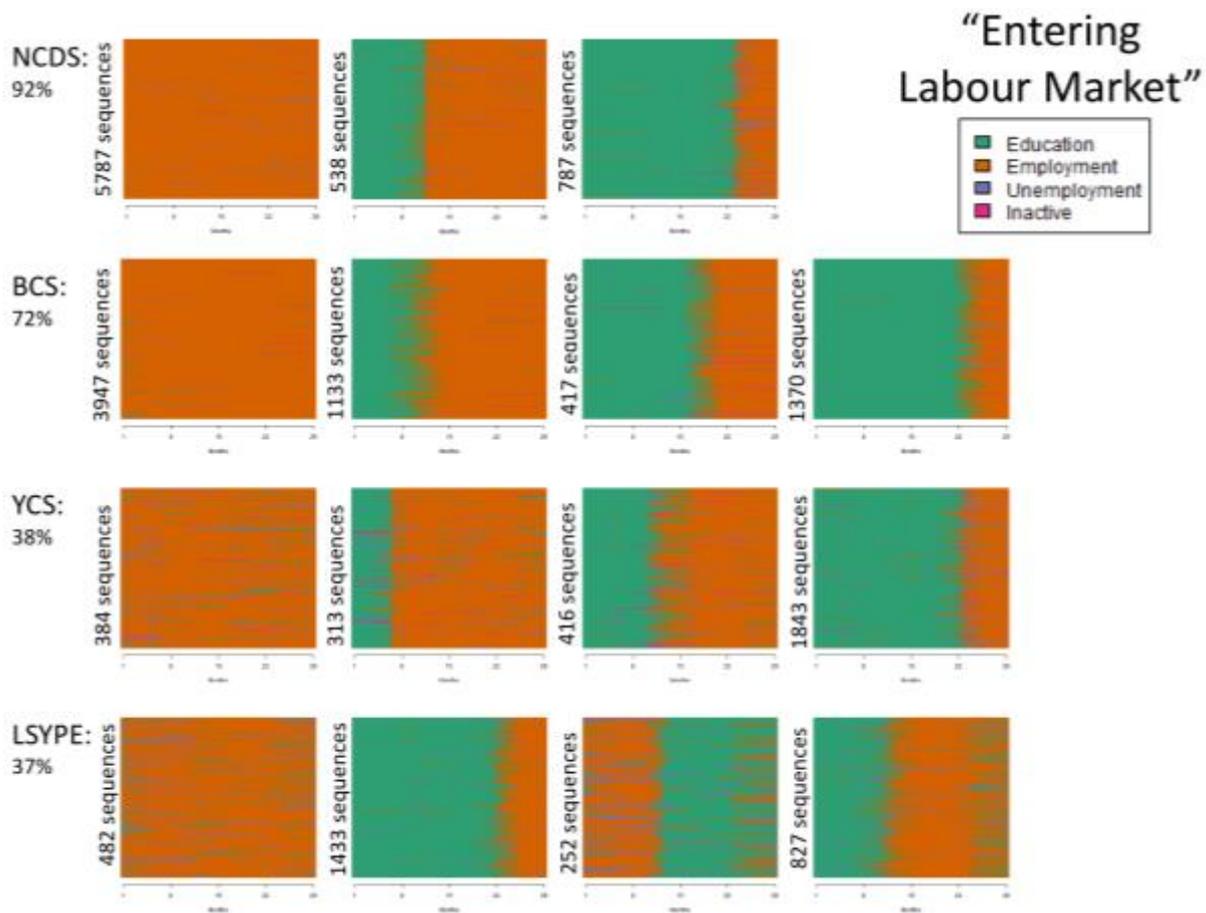
Notes: Total number of sequences from each dataset analysed as follows: NCDS: 8,356; BCS: 9,518; YCS: 8,682; LSYPE: 9,347. Horizontal axes track months from 1 to 29.

## Table 2. Descriptive statistics for identified groups within each cohort

NCDS	Entering the Labour Market	Accumulating Human Capital	Potentially Difficult Transition	Overall
N	7,110	852	394	8,356
Proportion	0.91	0.04	0.05	1.00
Male	0.59	0.57	0.23	0.57
Non-White	0.01	0.01	0.02	0.01
Single parent family	0.07	0.04	0.11	0.07
Parent has A Levels (no degree)	0.10	0.28	0.06	0.11
Parent has a degree	0.01	0.08	0.00	0.01
Home owner occupied	0.31	0.51	0.13	0.31
Home socially rented	0.43	0.14	0.53	0.42
Living in workless household	0.06	0.03	0.10	0.06

BCS	Entering the Labour Market	Accumulating Human Capital	Potentially Difficult Transition	Overall
N	6867	2282	369	9518
Proportion	0.72	0.24	0.04	1.00
Male	0.50	0.48	0.34	0.49
Non-White	0.02	0.05	0.04	0.02
Single parent family	0.03	0.04	0.05	0.04
Parent has A Levels (no degree)	0.04	0.10	0.01	0.05
Parent has a degree	0.03	0.14	0.01	0.05
Home owner occupied	0.28	0.44	0.11	0.31
Home socially rented	0.05	0.03	0.10	0.05
Living in workless household	0.03	0.03	0.07	0.03

YCS	Entering the Labour Market	Accumulating Human Capital	Potentially Difficult Transition	Overall
N	2,956	5,408	318	8,682
Proportion	0.40	0.55	0.05	1.00
Male	0.49	0.51	0.55	0.51
Non-White	0.03	0.13	0.07	0.09
Single parent family	0.16	0.14	0.17	0.15
Parent has A Levels (no degree)	0.04	0.09	0.07	0.07
Parent has a degree	0.03	0.09	0.08	0.07
Home owner occupied	0.75	0.85	0.64	0.80
Home socially rented	0.20	0.10	0.28	0.15
Living in workless household	0.08	0.08	0.15	0.08

LSYPE	Entering the Labour Market	Accumulating Human Capital	Potentially Difficult Transition	Overall
N	2,994	5,399	954	9,347
Proportion	0.37	0.51	0.12	1.00
Male	0.48	0.48	0.52	0.48
Non-White	0.07	0.19	0.13	0.14
Single parent family	0.27	0.22	0.36	0.25
Parent has A Levels (no degree)	0.20	0.25	0.16	0.22
Parent has a degree	0.12	0.22	0.12	0.17
Home owner occupied	0.74	0.79	0.52	0.74
Home socially rented	0.19	0.15	0.36	0.19
Living in workless household	0.10	0.13	0.26	0.13

Notes: NCDS results weighted using author's own attrition weighting scheme. No weights applied to BCS analysis, as number excluded due to attrition was too small to model. YCS and LSYPE analysis weighted using dataset-provided attrition weight

By contrast to the overall decline in the 'Entering the Labour Market' group, the size of the 'Accumulating Human Capital' group, shown in Figure 4, has grown significantly across the cohorts, from 4% in the earliest to around 50% in the most recent. These are clusters in which individuals remain in education throughout the period of analysis, and the growth reflects increases in both further and higher education across the cohorts

analysed. Perhaps unsurprisingly, given the large increase in the size of the group, there have been differences in the average characteristics of individuals in this group. For example, the proportion of young people whose ethnicity is not white has increased from 1% in the NCDS (similar to the population as a whole) to 19% in the LSYPE (compared to 14% in the population as a whole).

Figure 4. Plots of young people's individual transitions in four cohorts between the September following their 16ᵗʰ birthday and 29 months later: clusters placed in the "Accumulating Human Capital" group
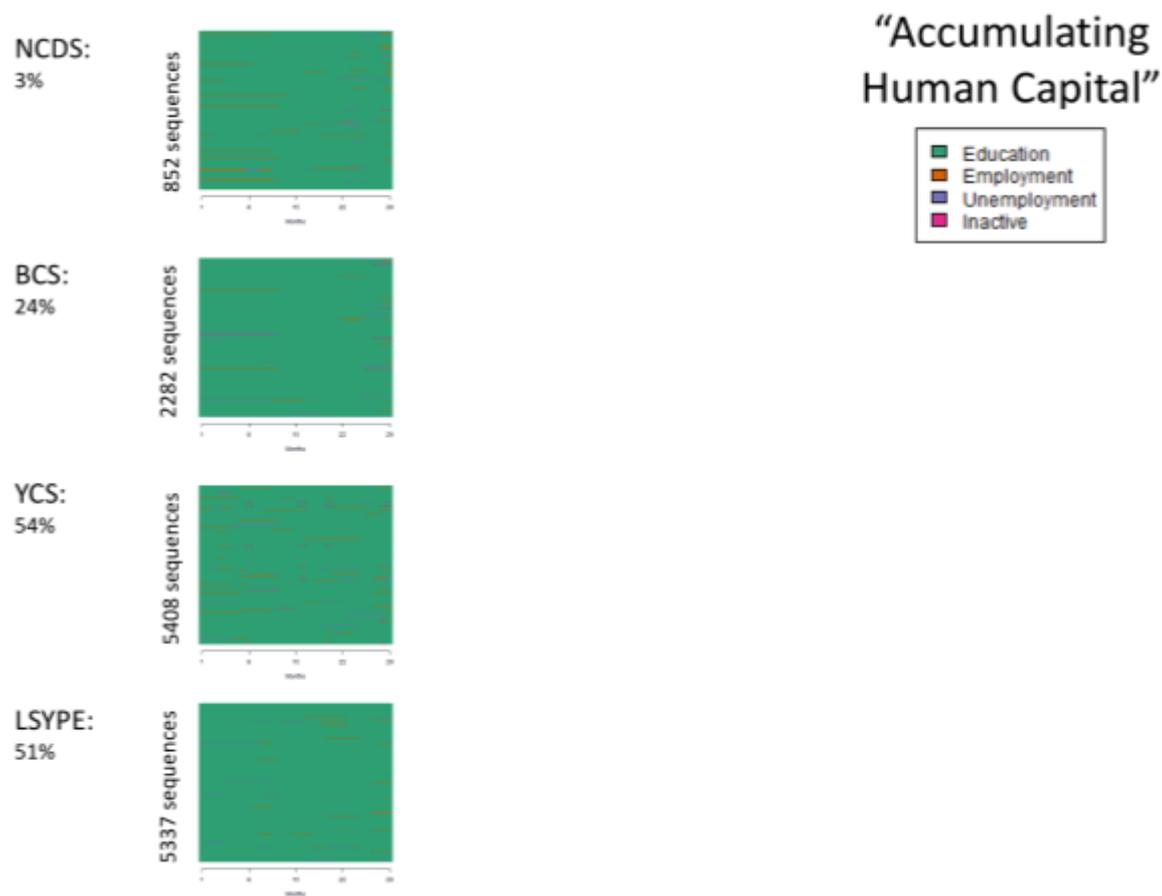
Notes: Total number of sequences from each dataset analysed as follows: NCDS: 8,356; BCS: 9,518; YCS: 8,682; LSYPE: 9,347. Horizontal axes track months from 1 to 29.

Lastly, the size of the 'Potentially Difficult Transition' group, shown in Figure 5, has also grown, although less dramatically than the 'Accumulating Human Capital' group, from 5% in the earliest cohort to 12% in the most recent. Unlike the other groups, whose respective fall and rise are relatively evenly spread through time, the growth of this group was concentrated between the 1980-born cohort and the 1990-born cohort. This group contains clusters in which individuals spend extended periods in inactivity or unemployment, seemingly not managing to settle into a job or education throughout this period of their lives. We should note that, for some, particularly where we see inactivity rather than unemployment, a transition of this type might be an active decision, for example individuals who become homemakers.

As such, we should not necessarily regard all individuals in this group as a cause for concern.

Relatedly, we also see a change in the behaviour of those who go straight from education into extended inactivity (predominantly young women, especially in earlier cohorts). In earlier cohorts, individuals who experience this kind of transition move into inactivity at around age 16. However, by the later cohorts, otherwise similar looking transitions show individuals moving into inactivity at around age 18, suggesting that such individuals are more likely to receive two additional years of education in later cohorts than they were in earlier ones. There may well be benefits for these individuals from the additional human capital they gain from these two years.

Figure 5. Plots of young people's individual transitions in four cohorts between the September following their 16th birthday and 29 months later: clusters placed in the 'Potentially Difficult Transition' group
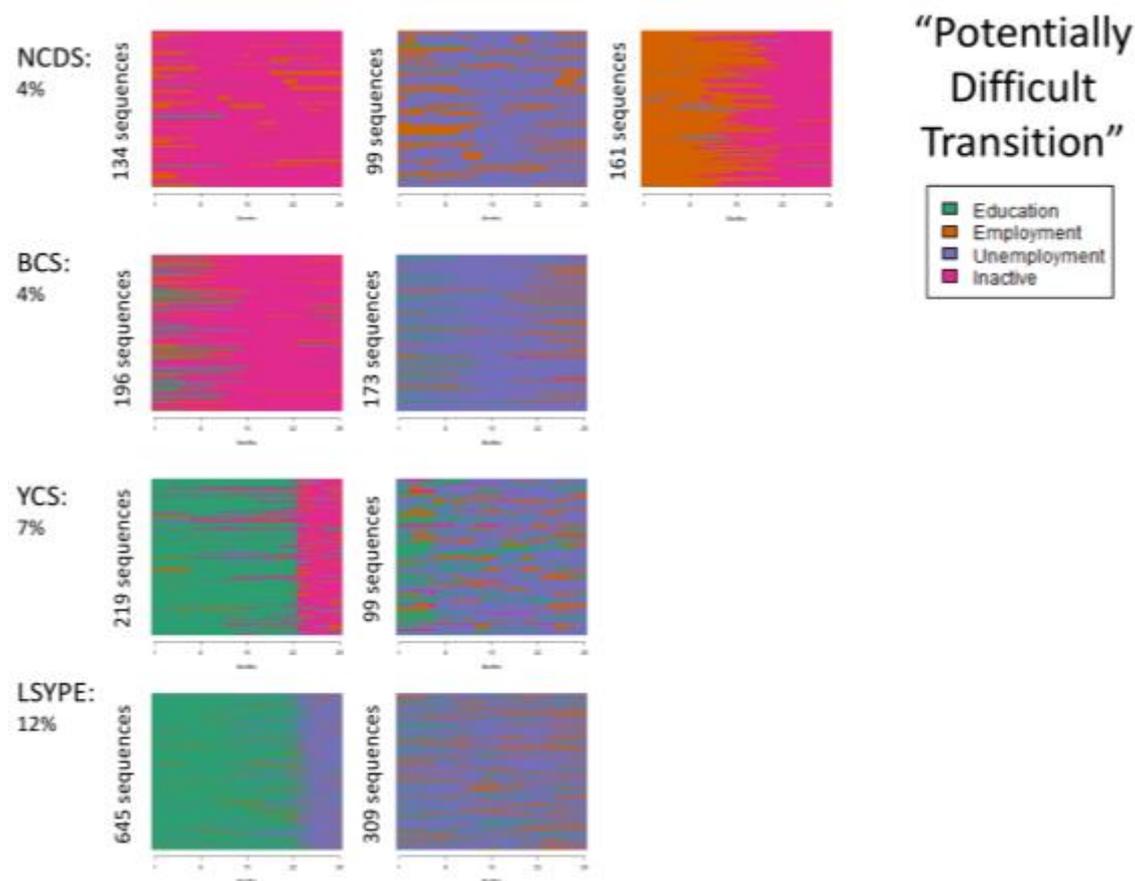

Notes: Total number of sequences from each dataset analysed as follows: NCDS: 8,356; BCS: 9,518; YCS: 8,682; LSYPE: 9,347. Horizontal axes track months from 1 to 29. In LSYPE analysis, purple represents both Unemployment and Inactivity.

A summary graphic of the size of each grouping is reported in Figure 6. It is encouraging to note that these findings accord with those of previous analyses of this period. Most directly, while our 'Potentially Difficult Transition' group makes up 5% of the YCS cohort (born in 1980) and 12% of the LSYPE cohort (born in 1990), the size of this group in the analysis by Dorsett and Lucchino (2014) falls somewhere in between (10%), for a sample born between 1975 and 1988. Similarly, the growth in the size of the Accumulating Human Capital group tracks the well-documented trend towards increased levels of post-compulsory education.

Figure 6. Percentage of each cohort allocated to each grouping
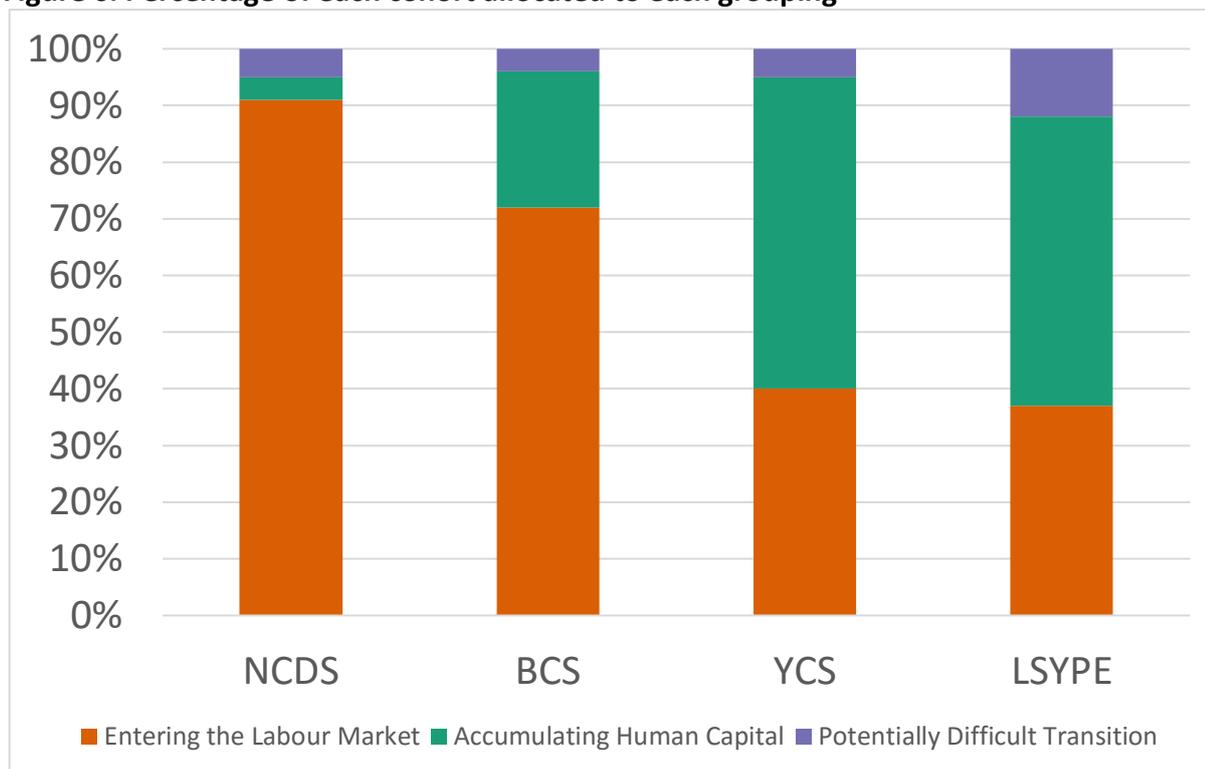
Notes: NCDS results weighted using author's own attrition weighting scheme. No weights applied to BCS analysis, as number excluded due to attrition was too small to model. YCS and LSYPE analysis weighted using dataset-provided attrition weights.

Predicting types of transitions

Table 3 reports the estimation results from multinomial logistic regression models of membership of a cluster in each of our three groupings.[vii] Our discussion focuses on the changing associations with the probability of experiencing a potentially difficult transition, given the particular policy interest of being able to predict transitions of this type.

The changing influence of gender and ethnicity are the most striking results, with both moving from being a significant predictor in one direction in the earliest cohort to being a significant predictor in the opposite direction by the most recent. First, in the case of ethnicity, individuals born in 1958 who are of non-white ethnicity are nearly five percentage points more likely to make a potentially difficult transition than their white peers. By contrast, in the 1990 cohort, individuals of non-white ethnicity are instead 2.5 percentage points less likely to experience a potentially difficult transition, compared to their white counterparts. Similarly, males born in 1958 are seven percentage points less likely to experience a potentially difficult transition than females from a similar background, while for the cohort born in 1990 the probability is two percentage points higher for males than females. To give some context for these results, we note that the potentially difficult transition grouping accounted for 5% of the 1958 cohort and 12% of the 1990 cohort.

Table 3. Estimated average marginal effects on the probability of an individual's membership of a cluster in each grouping (relative to the other two) from cohort-specific multinomial logistic regression models

	Entering the Labour Market				Accumulating Human Capital				Potentially Difficult Transition			
	NCDS	BCS	YCS	LSYPE	NCDS	BCS	YCS	LSYPE	NCDS	BCS	YCS	LSYPE
Non-White	-0.049**	-0.191***	-0.351***	-0.211***	0.002	0.171***	0.356***	0.237***	0.047**	0.019*	-0.005	-0.025**
	(-2.309)	(-6.927)	(-11.423)	(-12.609)	(0.197)	(6.706)	(12.017)	(15.145)	(2.500)	(1.727)	(-0.403)	(-2.277)
Male	0.071***	0.023**	-0.019	-0.001	-0.001	0.004	0.011	-0.018	-0.070***	-0.027***	0.008	0.019**
	(8.609)	(2.505)	(-1.534)	(-0.115)	(-0.298)	(0.455)	(0.860)	(-1.556)	(-9.036)	(-6.229)	(1.302)	(2.328)
Workless household	-0.023	-0.018	-0.012	-0.127***	0.006	-0.002	-0.010	0.071***	0.017	0.020**	0.022**	0.056***
	(-1.312)	(-0.730)	(-0.455)	(-6.204)	(0.603)	(-0.075)	(-0.374)	(3.488)	(1.176)	(2.404)	(2.090)	(4.855)
Lone parent	-0.003	-0.037	-0.022	0.056***	-0.008	0.017	0.029	-0.079***	0.011	0.020*	-0.007	0.023**
	(-0.153)	(-1.537)	(-1.116)	(3.950)	(-0.901)	(0.748)	(1.491)	(-5.433)	(0.745)	(1.880)	(-0.798)	(2.294)
Socially rented	0.014	0.081**	0.085**	-0.002	-0.012	-0.109***	-0.110***	-0.017	-0.002	0.029*	0.024	0.019
	(1.017)	(2.202)	(2.170)	(-0.087)	(-1.445)	(-3.159)	(-2.804)	(-0.579)	(-0.158)	(1.817)	(1.478)	(1.163)
Owner occupier	0.018	-0.022	-0.073**	-0.043*	0.027***	0.048*	0.086**	0.112***	-0.045***	-0.025*	-0.013	-0.069***
	(1.200)	(-0.750)	(-2.097)	(-1.686)	(3.689)	(1.765)	(2.517)	(4.260)	(-3.391)	(-1.649)	(-0.911)	(-4.410)
Parental A-Levels	-0.039***	-0.150***	-0.072*	-0.030**	0.048***	0.200***	0.081**	0.050***	-0.009	-0.049**	-0.009	-0.020*
	(-2.926)	(-6.294)	(-1.875)	(-2.099)	(10.748)	(12.162)	(2.149)	(3.455)	(-0.743)	(-2.304)	(-0.612)	(-1.801)
Parental degree	-0.053	-0.255***	-0.152***	-0.133***	0.096***	0.281***	0.117***	0.134***	-0.043	-0.026	0.035**	-0.000
	(-1.625)	(-12.044)	(-4.054)	(-8.250)	(15.987)	(17.954)	(3.188)	(8.430)	(-1.304)	(-1.580)	(2.490)	(-0.029)
N	8356	9518	8682	9144	8356	9518	8682	9144	8356	9518	8682	9144

Notes: Models also include regional dummy variables and missing variable dummies for the variables above. NCDS results weighted using author's own attrition weighting scheme. No weights applied to BCS analysis, as number excluded due to attrition was too small to model. YCS and LSYPE analysis weighted using dataset-provided attrition weights. T statistics reported in parentheses. Stars indicate statistical significance: * p=0.10; ** p=0.05; *** p=0.01.

The individual coefficients on each of our proxies for socioeconomic status (SES) are not straightforward to interpret in isolation, nor do they form any particularly obvious patterns. This partially reflects the changing importance of factors such as housing tenure as indicators for SES. Instead, to illuminate the combined role of SES, Tables 4.1, 4.2 and 4.3 present the predicted probability of an individual making each type of transition (Entering the Labour Market in 4.1, Accumulating Human Capital in 4.2 and Potentially Difficult Transition in 4.3) by gender, ethnicity and two combinations of the other model characteristics chosen to be an example of a 'high SES' individual and a 'low SES' individual. A 'high SES' individual is from a two-parent household, where at least one parent works, at least one parent holds a degree, and their house is owner-occupied. Conversely, a 'low SES' individual is from a lone parent, workless household, where the parent's highest qualification is below A-Level and their home is socially rented. Taken as a whole, these combinations remain indicative of advantage and disadvantage across all four cohorts.[viii]

Table 4.3 shows that the increase in the proportion of young people in clusters categorised as 'Potentially Difficult Transition' differs across ethnic/gender combinations. White females have a 6.9% probability of making a potentially difficult transition' in the NCDS, compared with a 1.5% probability for white males and 17.5% for non-white females. By the time of the LSYPE cohort, white females have a 10.2% probability of making a potentially difficult transition, slightly lower than the 12% probability for white males and higher than the probability for non-white females, which has fallen markedly to 7.1%.

The most obvious message from the predicted probabilities is that, throughout this period, young people from more advantaged backgrounds have been less likely than those from less advantaged backgrounds to make what we classify as a potentially difficult transition. There is evidence of this gap widening over time; from 4.9 percentage points in the NCDS to 17.4 percentage points in the LSYPE.

Table 4.1. Predicted probability of membership of a cluster in the 'Entering the Labour Market' grouping, by SES, gender and ethnicity in four cohorts

White male	Low SES	High SES	Overall	N
NCDS	95.8	67.9	96.5	4,338
BCS	81.6	31.2	73.4	4,564
YCS	53.6	23.6	40.7	3,429
LSYPE	30.6	34	40	3,161
White female	Low SES	High SES	Overall	N
NCDS	86.3	65.1	91.1	3,844
BCS	76.7	30.8	72.2	4,706
YCS	56.3	25.1	42.7	4,555
LSYPE	31.6	34	40	3,201
Non-White male	Low SES	High SES	Overall	N
NCDS	90.8	65.3	93.6	103
BCS	62.7	13.5	49.2	113
YCS	21.2	6	12.5	283
LSYPE	15.9	15.8	19.8	1,291
Non-White female	Low SES	High SES	Overall	N
NCDS	69.8	61.4	80.4	71
BCS	62.7	13.4	48	122
YCS	22.7	6.5	13.3	415
LSYPE	16.1	15.6	19.5	1,491
Overall	Low SES	High SES	Overall	N
NCDS	93.2	66.8	95	8,356
BCS	79.1	30.4	72.4	9,505
YCS	51.9	22	38.5	8,682
LSYPE	28.9	31.1	36.9	9,144

Notes: Predicted probabilities from underlying regression models reported in Table 3. Models also include regional dummy variables and missing variable dummies for the variables above. NCDS results weighted using author's own attrition weighting scheme. No weights applied to BCS analysis as number excluded due to attrition was too small to model. YCS and LSYPE analysis weighted using dataset-provided attrition weights. 'High SES' individual is from a two parent household, where at least one parent works, at least one parent holds a degree, and their house is owner occupied. 'Low SES' individual is from a lone parent, workless household, where the parent's highest qualification is below A-Level and their home is socially rented. 'Overall' are predictions based on the complete sample, not a weighted average of the 'Low SES' and 'High SES' predictions.

Table 4.2. Predicted probability of membership of a cluster in the 'Accumulating Human Capital' grouping, by SES, gender and ethnicity in four cohorts

White male	Low SES	High SES	Overall	N
NCDS	1.3	31.8	2.0	4,338
BCS	12.4	68.4	25.3	4,564
YCS	34.8	70.6	55.1	3,429
LSYPE	39.4	54.8	48	3,161
White female	Low SES	High SES	Overall	N
NCDS	1.3	33.6	2.1	3,844
BCS	11.8	68.3	25.1	4,706
YCS	34.1	70	53.8	4,555
LSYPE	42.2	56.6	49.8	3,201
Non-White male	Low SES	High SES	Overall	N
NCDS	1.4	33.9	2.1	103
BCS	27.7	86.1	49	113
YCS	68.9	90.8	84.7	283
LSYPE	61.7	76.7	71.6	1,291
Non-White female	Low SES	High SES	Overall	N
NCDS	1.1	35	2.0	71
BCS	27.7	85.8	48.3	122
YCS	68.9	90.8	84.3	415
LSYPE	64.8	78.2	73.3	1,491
Overall	Low SES	High SES	Overall	N
NCDS	1.3	32.7	2.0	8,356
BCS	12.4	68.9	25.7	9,505
YCS	37.4	72.9	57.7	8,682
LSYPE	43.9	59.1	52.4	9,144

Notes: Predicted probabilities from underlying regression models reported in Table 3. Models also include regional dummy variables and missing variable dummies for the variables above. NCDS results weighted using author's own attrition weighting scheme. No weights applied to BCS analysis as number excluded due to attrition was too small to model. YCS and LSYPE analysis weighted using dataset-provided attrition weights. 'High SES' individual is from a two parent household, where at least one parent works, at least one parent holds a degree, and their house is owner occupied. 'Low SES' individual is from a lone parent, workless household, where the parent's highest qualification is below A-Level and their home is socially rented. 'Overall' are predictions based on the complete sample, not a weighted average of the 'Low SES' and 'High SES' predictions.

Table 4.3. Predicted probability of membership of a cluster in the 'Potentially Difficult Transition' grouping, by SES, gender and ethnicity in four cohorts

White male	Low SES	High SES	Overall	N
NCDS	2.9	0.3	1.5	4,338
BCS	5.9	0.4	1.3	4,564
YCS	11.6	5.8	4.2	3,429
LSYPE	30	11.2	12	3,161
White female	Low SES	High SES	Overall	N
NCDS	12.4	1.3	6.9	3,844
BCS	11.5	0.9	2.7	4,706
YCS	9.7	4.9	3.5	4,555
LSYPE	26.1	9.4	10.2	3,201
Non-White male	Low SES	High SES	Overall	N
NCDS	7.8	0.8	4.2	103
BCS	9.7	0.4	1.8	113
YCS	9.9	3.2	2.8	283
LSYPE	22.4	7.5	8.6	1,291
Non-White female	Low SES	High SES	Overall	N
NCDS	29	3.6	17.5	71
BCS	9.7	0.8	3.7	122
YCS	8.5	2.7	2.4	415
LSYPE	19.1	6.2	7.1	1,491
Overall	Low SES	High SES	Overall	N
NCDS	5.5	0.6	2.9	8,356
BCS	8.5	0.6	1.9	9,505
YCS	10.7	5.1	3.8	8,682
LSYPE	27.2	9.8	10.7	9,144

Notes: Predicted probabilities from underlying regression models reported in Table 3. Models also include regional dummy variables and missing variable dummies for the variables above. NCDS results weighted using author's own attrition weighting scheme. No weights applied to BCS analysis as number excluded due to attrition was too small to model. YCS and LSYPE analysis weighted using dataset-provided attrition weights. 'High SES' individual is from a two parent household, where at least one parent works, at least one parent holds a degree, and their house is owner occupied. 'Low SES' individual is from a lone parent, workless household, where the parent's highest qualification is below A-Level and their home is socially rented. 'Overall' are predictions based on the complete sample, not a weighted average of the 'Low SES' and 'High SES' predictions.

An interesting question is whether changes in the size of the 'Potentially Difficult Transition' group are due to cross-cohort differences in composition or to cross-cohort changes in the influence of background characteristics. To explore this, we used the coefficients estimated using the NCDS to predict how group membership among individuals in the LSYPE would look had the influence of background characteristics not changed since the time of this first cohort. These predicted probabilities are reported in Table 5.1 for Entering the Labour Market, 5.2 for Accumulating Human Capital and 5.3 for Potentially Difficult Transition, in a similar way to those reported in Tables 4.1, 4.2 and 4.3. In each combination of ethnicity and gender we can compare young people's probabilities of being in each transition grouping in the NCDS, the LSYPE, and the LSYPE if the probabilities are affected by characteristics in the same way as they were in the NCDS cohort.[ix] For each ethnicity/gender combination, a comparison of the NCDS row with the 'NCDS associations/LSYPE cohort' row shows how changing composition over time affects the predicted probabilities. Similarly, a comparison of the LSYPE row with the 'NCDS associations/LSYPE cohort' row shows the changing influence of background characteristics, assuming composition is fixed.

Looking at the comparison of the NCDS probabilities with those of NCDS association on the LSYPE cohort we find that, across the full sample, the results suggest that the change in composition would be expected to, if anything, reduce the probability of making a potentially difficult transition from 2.9% to 2.2%. The biggest difference due to changing composition is among non-white females. In particular, those in the 'Low SES' group see their probability of making a potentially difficult transition fall from roughly 29% to 7.7% (among the 'High SES' group, there is no predicted change).

This implies that it is the change in the influence of background characteristics that is primarily responsible for the growth in this group. The second comparison (of the LSYPE row with the 'NCDS associations/LSYPE cohort' row) provides more detail; applying the NCDS associations to the LSYPE cohort predicts 2.2% will make a potentially difficult transition, whereas in fact 10.7% do. As such, it is the change in the relationship between characteristics and cluster memberships, rather than changes in the composition of the cohorts, that explains the growth in the proportion classified as a making a potentially difficult transition. The increased probability of making a potentially difficult transition is seen across all ethnicity/gender combinations for both high and low SES groups. However, it is among the low SES groups that the most dramatic differences are seen.

Table 5.1. Predicted probability of membership of a cluster in the 'Entering the Labour Market' grouping, by SES, gender and ethnicity for cohort born in 1989/90 and for same cohort assuming same influence of characteristics as that seen for cohort born in 1958

White male	Low SES	High SES	Overall	N
NCDS	95.8	67.9	96.5	4338
NCDS associations/LSYPE cohort	96.0	69.2	94	3161
LSYPE	30.6	34.0	40.0	3161
White female	Low SES	High SES	Overall	
NCDS	86.3	65.1	91.1	3844
NCDS associations/LSYPE cohort	86.6	66.4	90.6	3201
LSYPE	31.6	34.0	40.0.	3201
Non-White male	Low SES	High SES	Overall	
NCDS	90.8	65.3	93.6	103
NCDS associations/LSYPE cohort	91.0	66.6	92.0	1291
LSYPE	15.9	15.8	19.8	1291
Non-White female	Low SES	High SES	Overall	
NCDS	69.8	61.4	80.4	71
NCDS associations/LSYPE cohort	91.0	62.7	83.7	1491
LSYPE	16.1	15.6	19.5	1491
Overall	Low SES	High SES	Overall	
NCDS	93.2	66.8	95.0	8356
NCDS associations/LSYPE cohort	91.8	67.5	92.5	9144
LSYPE	28.9	31.1	36.9	9144

Notes: Predicted probabilities from underlying regression models reported in Table 3. Models also include regional dummy variables and missing variable dummies for the variables above. Missing value dummies are set to zero. NCDS results weighted using author's own attrition weighting scheme. LSYPE analysis weighted using dataset-provided attrition weights. 'High SES' individual is from a two parent household, where at least one parent works, at least one parent holds a degree, and their house is owner occupied. 'Low SES' individual is from a lone parent, workless household, where the parent's highest qualification is below A-Level and their home is socially rented. 'Overall' are predictions based on the complete sample, not a weighted average of the 'Low SES' and 'High SES' predictions.

Table 5.2. Predicted probability of membership of a cluster in the 'Accumulating Human Capital' grouping, by SES, gender and ethnicity for cohort born in 1989/90 and for same cohort assuming same influence of characteristics as that seen for cohort born in 1958

White male	Low SES	High SES	Overall	N
NCDS	1.3	31.8	2.0	4338
NCDS associations/LSYPE cohort	1.2	30.6	5.1	3161
LSYPE	39.4	54.8	48	3161
White female	Low SES	High SES	Overall	
NCDS	1.3	33.6	2.1	3844
NCDS associations/LSYPE cohort	1.2	32.3	5.4	3201
LSYPE	42.2	56.6	49.8	3201
Non-White male	Low SES	High SES	Overall	
NCDS	1.4	33.9	2.1	103
NCDS associations/LSYPE cohort	1.3	32.6	5.5	1291
LSYPE	61.7	76.7	71.6	1291
Non-White female	Low SES	High SES	Overall	
NCDS	1.1	35	2.0	71
NCDS associations/LSYPE cohort	1.3	33.7	5.5	1491
LSYPE	64.8	78.2	73.3	1491
Overall	Low SES	High SES	Overall	
NCDS	1.3	32.7	2.0	8356
NCDS associations/LSYPE cohort	1.2	31.8	5.3	9144
LSYPE	43.9	59.1	52.4	9144

Notes: Predicted probabilities from underlying regression models reported in Table 3. Models also include regional dummy variables and missing variable dummies for the variables above. Missing value dummies are set to zero. NCDS results weighted using author's own attrition weighting scheme. LSYPE analysis weighted using dataset-provided attrition weights. 'High SES' individual is from a two parent household, where at least one parent works, at least one parent holds a degree, and their house is owner occupied. 'Low SES' individual is from a lone parent, workless household, where the parent's highest qualification is below A-Level and their home is socially rented. 'Overall' are predictions based on the complete sample, not a weighted average of the 'Low SES' and 'High SES' predictions.

Table 5.3. Predicted probability of membership of a cluster in the "Potentially Difficult Transition", by SES, gender and ethnicity for cohort born in 1989/90 and for same cohort assuming same influence of characteristics as that seen for cohort born in 1958

White male	Low SES	High SES	Overall	N
NCDS	2.9	0.3	1.5	4338
NCDS associations/LSYPE cohort	2.8	0.3	0.9	3161
LSYPE	30	11.2	12	3161
White female	Low SES	High SES	Overall	
NCDS	12.4	1.3	6.9	3844
NCDS associations/LSYPE cohort	12.2	1.3	4.0	3201
LSYPE	26.1	9.4	10.2	3201
Non-White male	Low SES	High SES	Overall	
NCDS	7.8	0.8	4.2	103
NCDS associations/LSYPE cohort	7.7	0.8	2.5	1291
LSYPE	22.4	7.5	8.6	1291
Non-White female	Low SES	High SES	Overall	
NCDS	29	3.6	17.5	71
NCDS associations/LSYPE cohort	7.7	3.6	10.8	1491
LSYPE	19.1	6.2	7.1	1491
Overall	Low SES	High SES	Overall	
NCDS	5.5	0.6	2.9	8356
NCDS associations/LSYPE cohort	6.9	0.7	2.2	9144
LSYPE	27.2	9.8	10.7	9144

Notes: Predicted probabilities from underlying regression models reported in Table 3. Models also include regional dummy variables and missing variable dummies for the variables above. Missing value dummies are set to zero. NCDS results weighted using author's own attrition weighting scheme. LSYPE analysis weighted using dataset-provided attrition weights. 'High SES' individual is from a two parent household, where at least one parent works, at least one parent holds a degree, and their house is owner occupied. 'Low SES' individual is from a lone parent, workless household, where the parent's highest qualification is below A-Level and their home is socially rented. 'Overall' are predictions based on the complete sample, not a weighted average of the 'Low SES' and 'High SES' predictions.

Extending sequence analysis to age 24

A possible reservation about the results discussed so far is that they may not warrant particular attention since what is more important is how the school to work transitions play out in the longer run. Such a view may be justified if these early patterns do not persist. However, if they are predictive of transitions over a longer period, their importance is greatly increased.

In order to explore this, we also carried out an analysis of sequences beginning 30 months after turning 16 (i.e. following the end of the period we have been considering so far) up to approximately age 24, and compared the resulting groupings to those for the first 29 months post-16. This is only possible for the two datasets where the data are available: the NCDS and the BCS. We carry out

sequence and cluster analysis on the same basis as was done for the earlier time period analyses, except that it starts 30 months after the September following their 16th birthday and continues for 69 months.[x] This time we use 14-cluster (rather than 7-cluster) solutions, reflecting the greater heterogeneity possible within longer sequences. Again, our choice of a 14-cluster solution is primarily on the basis of average silhouette distances.

We once again aggregate these clusters into our three broad groupings: Entering the Labour Market, Accumulating Human Capital and Potentially Difficult Transition. One particular challenge with conducting extended sequence analysis on the NCDS is the quality of the monthly activity data available particularly once we extend

to age 24. The NCDS appears to have a rather systematic problem with gaps between different spells, which results in the loss of a substantial number of individuals from our analysis, reducing the sample size from 8,372 to 6,122. This loss seems concentrated among individuals in the 'Entering the Labour Market' group, and we suspect that this is responsible for inflating the size of the 'Accumulating Human Capital' grouping compared to that estimated in the shorter analysis. Consequently, there is a concern about the ability of the NCDS to support the longer-run analysis. The BCS analysis does not suffer from the same problem; extending to age 24 reduces the sample size only marginally (from 9,518 to 9,419). In view of this, we feel more confident about the BCS results.

In order to learn more about the relationship between the two sets of categorisations, we cross-tabulate the groupings into which individuals are placed in the shorter- (29 month) and longer-term (98 month) analyses. Considering first the NCDS, we see that a majority of individuals in the short-term groupings remain in the same grouping on the basis of the extended sequence analysis. There is also, for example, some movement from 'Entering the Labour Market' into 'Potentially Difficult Transition'. Similarly, some individuals initially classified as 'Potentially Difficult Transition' have seen a recovery by this later period. Overall, though, there is a strong correlation between the two sets of groupings. We should also note that the 'Potentially Difficult Transition' category grows primarily from individuals that were previously characterised as being 'Entering the Labour Market' and very few from the 'Accumulating Human Capital' grouping. In the BCS, the picture is much the same, except for the much-reduced size of the missing category, as discussed in the introduction to this section.

Table 6. NCDS: Cross-tabulation of groupings on basis of 16-18 sequence analysis and of groupings on basis of 18-25 sequence analysis

	18-24 Groupings				
16-18 Groupings	**ELM**	**AHC**	**PDT**	**Missing**	**Total (freq.)**
ELM	61.9	1.1	12.4	24.6	7,110
AHC	13.7	41.7	2.1	42.5	852
PDT	8.6	0.5	55.6	35.3	394
Missing	75.0	0.0	25.0	0.0	16
Total	54.6	5.2	13.4	26.9	8,372

Notes: ELM = Entering the Labour Market; AHC = Accumulating Human Capital; PDT = Potentially Difficult Transition. Reporting row proportions, except for the final (total) column, which reports frequencies.

Table 7. BCS: Cross-tabulation of groupings on basis of 16-18 sequence analysis and of groupings on basis of 18-25 sequence analysis

	18-24 Groupings				
16-18 Groupings	**ELM**	**AHC**	**PDT**	**Missing**	**Total (freq.)**
ELM	81.3	5.7	12.2	0.9	6,867
AHC	6.8	87.2	4.4	1.6	2,282
PDT	23.0	6.8	69.4	0.8	369
Total	61.1	25.3	12.6	1.0	9,518

Notes: ELM = Entering the Labour Market; AHC = Accumulating Human Capital; PDT = Potentially Difficult Transition. Reporting row proportions, except for the final (total) column, which reports frequencies.

What do we learn from this? Those who are making a potentially difficult transition in the earlier analysis are likely still to be making a potentially difficult transition on the basis of the

longer run analysis: in the NCDS 85.9% of those deemed to be 'Potentially Difficult Transition' on the early basis (and for whom we can derive a longer run grouping) are placed in this group over the longer term; in the BCS the comparable figure is 70.0%. In addition, as one might expect, the longer analysis also picks up an additional number of cases that we deem to be potentially difficult transitions, on the basis of their trajectories post-29 months. However, we next explore whether this changes the risks of various observable characteristics associated with making a potentially difficult transition.

Reassuringly, in Table 8 we find a fairly similar pattern in the estimated average marginal effects of making each type of transition in this analysis as we did in the 29-month analysis, although there are unexpected or surprisingly insignificant results associated with a few characteristics for the NCDS. Nevertheless, we conclude that this suggests that while the sequence and cluster analyses themselves do not necessarily pick up all the individuals who are making a potentially difficult transition in this shorter timeframe, our shorter-run analysis nevertheless identifies the observable groups that are likely to be at greater risk.

Table 8. Estimated average marginal effects on the probability of an individual's membership of a cluster in each grouping (relative to the other two) from cohort-specific multinomial logistic regression models – Sequences to age 25

	Entering the Labour Market		Accumulating Human Capital		Potentially Difficult Transition	
	NCDS	BCS	NCDS	BCS	NCDS	BCS
Non-White	-0.013	-0.196***	-0.025	0.167***	0.037	0.028
	(-0.338)	(-5.367)	(-0.891)	(5.766)	(1.245)	(1.176)
Male	0.221***	0.128***	-0.060***	0.023***	-0.161***	-0.151***
	(18.692)	(12.545)	(-7.229)	(2.674)	(-14.557)	(-19.004)
Workless household	-0.037	-0.063**	0.007	-0.005	0.031	0.068***
	(-1.343)	(-2.263)	(0.309)	(-0.197)	(1.590)	(4.492)
Lone parent	-0.042*	-0.055**	0.035**	0.034	0.007	0.021
	(-1.764)	(-1.998)	(1.970)	(1.457)	(0.438)	(1.066)
Socially rented	-0.009	0.088**	0.015	-0.136***	-0.006	0.048**
	(-0.393)	(2.159)	(0.838)	(-3.655)	(-0.379)	(1.971)
Owner occupier	0.039*	0.013	-0.004	0.055*	-0.035**	-0.069**
	(1.707)	(0.402)	(-0.221)	(1.944)	(-2.125)	(-3.097)
Parental A-Levels	0.027	-0.166***	0.010	0.211***	-0.037***	-0.045**
	(1.558)	(-6.704)	(0.886)	(12.653)	(-2.591)	(-2.147)
Parental degree	-0.006	-0.250***	0.085***	0.348***	-0.078**	-0.098***
	(-0.157)	(-8.369)	(4.301)	(20.348)	(-2.206)	(-3.403)
N	6122	8574	6122	8574	6122	8574

Notes: Models also include regional dummy variables and missing variable dummies for the variables above. NCDS results weighted using author's own attrition weighting scheme. No weights applied to BCS analysis, as number excluded due to attrition was too small to model. T statistics reported in parentheses. Stars indicate statistical significance: * p=0.10; ** p=0.05; *** p=0.01.

Conclusions

In this paper we have used sequence analysis to analyse young people's transitions into the labour market and how these have changed over the past thirty years or so. The advantage of sequence analysis is that it allows us to consider young people's transition patterns as a whole, rather than concentrating on specific individual transitions and their timings. Our results shed new light on how the nature of young people's very early transitions have evolved, and on how the factors influencing them have changed over time.

It is well-established that young people's employment is more sensitive than older people's to the underlying strength of the economy. There are several reasons that combine to give such pro-cyclicality. For instance, firms may be more willing to lose younger workers in a recession than older workers with valuable experience. Equally, coming out of a downturn, firms may feel more prepared to recruit younger workers than older (more expensive) workers. Kahn (2010) shows that the time at which young people enter the labour market affects their subsequent outcomes. Scarring is one channel through which this operates; Gregg (2001) and Gregg & Tominey (2005) provide compelling evidence that youth unemployment can adversely affect both employment and earnings prospects as an adult.

The results in this paper are influenced by cyclicality to the extent that economic conditions prevailing at the time of reaching school-leaving age vary across cohort. However, since these cohorts span three decades, the results also pick up structural changes. Such changes reflect more than just economic influences and instead capture changes over time in, for instance, social preferences, demographics, technology and institutions. Unsurprisingly, we find a substantial shift away from early labour market entry towards gaining significant amounts of additional education or training before entering a job. However, in addition we have documented a rise in the proportion of successive cohorts that experience potentially difficult transitions, with prolonged or numerous spells not in education, employment or training. This group has grown in size from 5% of the sample born in 1958 to 12% of the sample born in 1990, with pretty much all of this growth concentrated between the 1980- and 1990-born cohorts.

Focussing on the 'Potentially Difficult Transition' group, there are two particularly striking results. First, females have gone from being more likely than males to be members of this group in early cohorts to being less likely by the more recent cohorts. One reason for this is likely to be a decline in the proportion of young women who choose to move quickly from education into an extended period of inactivity associated with homemaking or starting a family. Alongside this we find that individuals who move from education into long-term inactivity have become more likely to remain in education for two additional years (leaving education at age 18, rather than age 16) before entering inactivity.

Second, we find that young people from a non-white ethnic background go from being more likely than whites to experience a potentially difficult transition to being less likely. Across this period the non-white population of England has grown significantly in size, has diversified and has become more established. We suspect that all three of these facts have contributed to the relative improvement in the probability that individuals of non-white ethnicity experience transitions likely to be precursors of future economic prosperity.

In addition, we find that socioeconomic status, as captured through a combination of indicators, remains a powerful predictor of young people's chances of experiencing a potentially difficult transition. This is unsurprising, but underlines that it is among those from disadvantaged backgrounds where there has been the greatest increase in difficult transitions.

Lastly, we assess the extent to which the patterns seen in these early years predict longer-term outcomes. The fact that we find a high degree of correlation suggests that those likely to face ongoing difficulties in the labour market are often identifiable at a very early stage. This points to the importance of early transitions.

Acknowledgements

References

Abbott, A. (1995). Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology, 21*, 93-113. https://doi.org/10.1146/annurev.so.21.080195.000521

Andrews, M., & Bradley, S. (1997). Modelling the Transition from School and the Demand for Training in the United Kingdom. *Economica, 64*(255), 387-413. https://doi.org/10.1111/1468-0335.00087

Anyadike-Danes, M., & McVicar, D. (2005). You'll never walk alone: Childhood influences and male career path clusters. *Labour Economics, 12*, 511-530. https://doi.org/10.1016/j.labeco.2005.05.008

Anyadike-Danes, M., & McVicar, D. (2010). My Brilliant Career: Characterizing the Early Labor Market Trajectories of British Women From Generation X. *Sociological Methods & Research, 38*(3), 482-512. https://doi.org/10.1177/0049124110362968

Arulampalam, W. (2001). Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages. *The Economic Journal, 111*(475), F585-F606. https://doi.org/10.1111/1468-0297.00664

Bell, D.N. and Blanchflower, D.G., 2011. Young people and the Great Recession. *Oxford Review of Economic Policy*, 27(2), pp.241-267. https://doi.org/10.1093/oxrep/grr011

Berrington, A. (2001). Transition to Adulthood in Britain *Transitions to Adulthood in Europe* (Vol. 10, pp. 67-102): Springer Netherlands. https://doi.org/10.1007/978-94-015-9717-3_4

Caspi, A., Entner Wright, B. R., Moffitt, T. E., & Silva, P. A. (1998). Early Failure in the Labor Market: Childhood and Adolescent Predictors of Unemployment in the Transition to Adulthood. *American Sociological Review, 63*(3), 424-451. https://doi.org/10.2307/2657557

Dorsett, R., & Lucchino, P. (2014). Explaining patterns in the school-to-work transition: An analysis using optimal matching. Advances in Life Course Research, 22(1-14). https://doi.org/10.1016/j.alcr.2014.07.002

Durbin, R., Eddy, S. R., Krogh, A., & Mitchison, G. (1998). *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/cbo9780511790492

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*, 225-250. https://doi.org/10.1007/s10680-007-9133-7

Fergusson, R., Pye, D., Esland, G., McLaughlin, E., & Muncie, J. (2000). Normalized dislocation and new subjectivities in post-16 markets for education and work. *Critical Social Policy, 20*(3), 283-305. https://doi.org/10.1177/026101830002000302

Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011a). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software, 40*(4), 1-37. https://doi.org/10.18637/jss.v040.i04

Gabadinho, A., Ritschard, G., Studer, M., & Müller, N. S. (2011b). Mining sequence data in R with the TraMineR package: A user's guide: University of Geneva, Switzerland. Retrieved from http://mephisto.unige.ch/pub/TraMineR/doc/TraMineR-Users-Guide.pdf

Gregg, P. (2001). The Impact of Youth Unemployment on Adult Unemployment in the NCDS. *The Economic Journal, 111*(475), F626-F653. https://doi.org/10.1111/1468-0297.00666

Gregg, P., & Tominey, E. (2005). The wage scar from male youth unemployment. *Labour Economics, 12*, 487-509. https://doi.org/10.1016/j.labeco.2005.05.004

Halpin, B., & Chan, T. W. (1998). Class Careers as Sequences: An Optimal Matching Analysis of Work-Life Histories. *European Sociological Review, 14*(2), 111-130. https://doi.org/10.1093/oxfordjournals.esr.a018230

Kahn, L. (2010) The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy, *Labour Economics*, 17(2), pp. 303-316. https://doi.org/10.1016/j.labeco.2009.09.002

Kaufman, L., & Rousseeuw, P. J. (1987). Clustering by means of Medoids *Statistical Data Analysis Based on the L1–Norm and Related Methods* (pp. 405-416): North-Holland.

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Chichester: Wiley Inter-Science. https://doi.org/10.1002/9780470316801

Kneale, D., Lupton, R., Obolenskaya, P., & Wiggins, R. D. (2010). A cross-cohort description of young people's housing experience in Britain over 30 years: An application of Sequence Analysis. DoQSS Working Paper 10-17, Institute of Education, University of London. Retrieved from http://repec.ioe.ac.uk/REPEc/pdf/qsswp1017.pdf

Lesnard, L. (2006). Optimal Matching and Social Science *Série des Documents de Travail du CREST*. Paris, France: Institut National de la Statistique et des Etudes Economiques. Retrieved from https://halshs.archives-ouvertes.fr/halshs-00008122/

Martin, P., Schoon, I., & Ross, A. (2008). Beyond Transitions: Applying Optimal Matching Analysis to Life Course Research. *International Journal of Social Research Methodology, 11*(3), 1-21. https://doi.org/10.1080/13645570701622025

Martin, P., & Wiggins, R. D. (2011). Optimal Matching Analysis. In M. Williams & W. P. Vogt (Eds.), *The SAGE Handbook of Innovation in Social Research Methods* (pp. 385-408). London: SAGE Publications Ltd. https://doi.org/10.4135/9781446268261.n22

OECD. (2008). *Jobs for Youth - United Kingdom*. Paris, France: Organisation for Economic Cooperation and Development. http://www.oecd.org/unitedkingdom/jobsforyouthdesemploispourlesjeunesunitedkingdom.htm

Paull, G. (2002). Biases in the Reporting of Labour Market Dynamics *IFS Working Paper Series*. London, UK: Institute for Fiscal Studies. Retrieved from http://www.ifs.org.uk/publications/2012

Pollock, R. (2007) Holistic trajectories: a study of combined employment, Housing and family careers by using multiple-sequence analysis. *Journal of the Royal Statistical Society*, Series A, 170(1), 167-183. https://doi.org/10.1111/j.1467-985X.2006.00450.x

Quintini, G., & Manfredi, T. (2009). Going Separate Ways? School-to-Work Transitions in the United States and Europe *OECD Social, Employment and Migration Working Paper*. Paris, France: Organisation for Economic Cooperation and Development. https://doi.org/10.1787/221717700447

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*, 53-65. https://doi.org/10.1016/0377-0427(87)90125-7

Ryan, P., 2001. The school-to-work transition: a cross-national perspective. *Journal of economic literature*, 39(1), pp.34-92. https://doi.org/10.1257/jel.39.1.34

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

Wu, L. L. (2000). Some Comments on Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect. *Sociological Methods and Research, 29*(1), 41-64. https://doi.org/10.1177/0049124100029001003

Endnotes

[i] For example, we made use of a separate variable reporting school leaving date and also used characteristics such as young people's highest educational qualification reported by age 23 to impute earlier education status.

[ii] We carry out a sensitivity analysis to assess the extent to which sample reduction is likely to influence our results by using the alternative approach of treating missing as a state in itself (Gabadinho, Ritschard, Studer, & Müller, 2011b, pp. 55-61). This makes little difference to our findings.

[iii] Optimal matching in this sense should not be confused with the identically named technique within the propensity score matching literature. Partly in order to avoid this ambiguity we use the term sequence analysis throughout.

[iv] Combinatorial approaches, such as those outlined by Elzinga (for example Elizinga & Liefbroer, 2007), are another alternative.

[v] This cluster contains individuals who have entered the labour market immediately at the end of compulsory schooling but return to education at a later point. This makes it slightly ambiguous if they should be classified as 'Entering the Labour Market' (since they do this but then leave again) or 'Accumulating Human Capital' (since they return to do this but are not in education throughout the period). In any case, it makes up only approximately 2.5% of the sample; little changes if it is reclassified as AHC or dropped entirely.

[vi] It is also possible, though, that some of this effect is explained by under-reporting of short spells in the NCDS/BCS, as discussed earlier.

[vii] We also fitted a single multinomial logistic regression model on the pooled sample from all cohorts, including a cohort regressor and all predictors interacted with these. These replicated the results obtained from the separate models, but allowed for inference testing of the differences between the influences of characteristics in each cohort. These significance tests are not reported in this paper but are available on request.

[viii] The distinction between high SES and low SES does not conform to any standard definition. The two groups were chosen in order to satisfy two criteria: first, that the characteristics used in the definition had a strong association with disadvantage such that it was plausible to view the high SES group as unambiguously "better off" (at least on average) than the low SES group and, second, that resulting groups were of a sufficient size to be of practical use.

[ix] The NCDS and LSYPE rows in Table 5 contain identical results to corresponding rows in Table 4 but are included for convenience of comparison.

[x] In addition, we carried out the same analysis over the whole time period (i.e. both the initial 29 months and the following 69 months) and achieved similar results to those reported later in this section.

Getting better all the time? Selective attrition and compositional changes in longitudinal and life course studies

Susanne Kelfve   Aging Research Center, Karolinska Institutet/Stockholm University & Linköping
Susanne.kelfve@ki.se   University, Sweden
Stefan Fors   Aging Research Center, Karolinska Institutet/Stockholm University, Sweden
Carin Lennartsson   Aging Research Center, Karolinska Institutet/Stockholm University, Sweden

Abstract

Longitudinal surveys are valuable tools for investigating health and social outcomes across the life course. In such studies, selective mortality leads to changes in the social composition of the sample, but little is known about how selective survey participation affects the sample composition, in addition to the selective mortality. In the present paper, we followed a Swedish cohort sample over six waves 1968–2011. For each wave we recalculated the distribution of baseline characteristics in the sample among i) the sample still alive and ii) the sample still alive and with complete follow-up. The results show that the majority of the compositional changes in the cohort were modest and driven mainly by mortality. However, for some characteristics, class in particular, the selection was considerable and in addition, was substantially compounded by survey non-participation. We suggest that sample selections should be taken into account when interpreting the results of longitudinal studies, in particular when researching social inequalities.

Keywords

Introduction

Longitudinal surveys are valuable tools for studying life course processes. Following individuals over many years provides among other things the opportunity to investigate life-course exposures that impact health and living conditions in later life – a prerequisite for understanding the complex process of aging (Kuh, Cooper, Hardy, Richards & Ben-Shlomo, 2014). However, the usefulness of a longitudinal survey is partly dependent on how well the participants represent the target population and on the completeness of follow-up among the participants. The results might be fallacious if individuals lost to follow-up differ from those who participate throughout the entire survey (Deeg, 2002; Lynn, 2011; Willson, Shuey, & Elder, 2007). While there are now methods developed to handle missing data efficiently, many researchers still resort to complete case analyses and listwise deletion is still the default way of handling missing data in statistical software packages (Bartlett, Carpenter, Tilling & Vansteelandt, 2014).

In this paper we investigate how a longitudinal sample, with 43 years follow-up time, changes over time because of selective mortality and selective survey non-participation.

Selective mortality

By default, life course studies include selection processes. As the people in the sample age, mortality reduces the number of individuals in the sample. However, mortality is selective: not all groups in society have the same chance to survive to high ages. For instance, women tend to live longer than men, a pattern well documented in most part of the world (Barford, Dorling, Smith & Shaw, 2006; Thorslund, Wastesson, Agahi, Lagergren & Parker, 2013). In Sweden, for example, women are expected to live approximately four years longer than men (Vaupel, Zhang, & van Raalte, 2011). Individuals from higher socioeconomic strata live longer than individuals from lower socioeconomic strata. Regardless of the socioeconomic indicator used, people in higher social strata have robust advantages: those with higher education, income, or status or from a higher social class, tend to live longer than those with lower education, income, or status or from a lower social class (Torssander & Erikson, 2010). Civil status has also been associated with mortality differences; married individuals tend to live longer than those who are unmarried (Fors, Lennartsson, & Lundberg, 2011; Lennartsson & Lundberg, 2007; Umberson,

1992). The literature devoted to explaining the mechanisms behind social stratification in mortality is extensive, and hypotheses include poorer living conditions, health behaviours, work conditions, and genes (Ferraro & Shippee, 2009; Mackenbach, 2012; Phelan, Link, & Tehranifar, 2010; Stringhini et al., 2011).

Diversity in mortality risk leads to higher mortality in more vulnerable groups. Thus, in any given population or sample, the distribution of factors such as sex, civil status, health status, and education gradually changes over time, eventually giving rise to a privileged group of individuals that have survived to high ages (Zajacova & Burgard, 2013). If the selective mortality is substantial, it may lead to 'cohort inversion', a phenomenon whereby a cohort becomes healthier over time as the most disadvantaged individuals (i.e. those with the poorest health) die. Cohort inversion has substantial implications for health inequality research, as it might lead to the impression that inequality decreases over time, whereas it has in fact increased at the individual level (Ferraro, Shippee, & Schafer, 2009). For example, if mortality is higher among individuals with low education, and we assume this mortality is associated with poor health, then over time, differences in poor health between the groups with high and low education will decrease. However, this does not necessarily mean that health differences diminish over time; the individuals with a lower level of education may still have the same degree of poorer health than those with a higher level of education. The only change is that those with the poorest health have died.

Selective survey participation

Longitudinal surveys also suffer from attrition for reasons other than mortality. Survey non-participation can occur for different reasons in one or several survey waves. Survey non-participation can be due to 1) inability to track respondents or to establish contact with them 2) migration, or 3) non-response. Non-response can, in turn, be driven by a) refusal to participate, or b) inability to participate (e.g. due to poor health or cognitive impairment).

A great deal of research has been devoted to analysing factors related to survey non-response. Non-response has consistently been associated with socio-demographic factors such as low education, low socioeconomic status, unemployment, and not being married (Galea & Tracy, 2007). Other factors, such as age, sex, and health, have shown varying

associations with response rate, partly depending on which age groups are included in the study. In contrast to non-response among people in younger age groups (Fejer et al., 2006; Korkeila et al., 2001), non-response among people 70 years and older is usually related to high age and female sex (Hardie, Bakke, & Mørkve, 2003; Kjøller & Thoning, 2005; Klein et al., 2011). The association between non-response and poor health and mortality is also specific to older age groups (de Souto Barreto, 2012; Kelfve, Thorslund, & Lennartsson, 2013). In longitudinal studies of older people, drop-outs tend to be older, more often cognitively impaired, and to have poorer health than those who remain in the studies (Chatfield, Brayne, & Matthews, 2005). The association between drop-outs and socioeconomic factors has been less clear (Banks, Muriel, & Smith, 2011).

Selection processes in longitudinal studies

To sum up, it has been shown that selective mortality changes the composition in a longitudinal sample over time and that survey non-response is unevenly distributed in the population. But less is known about how these two processes work together over time.

There is an association between non-response and mortality. Both baseline non-response and non-response in a later wave are associated with increased mortality risk (Ferrie et al., 2009). The question that arises is what implication this has for the composition of a longitudinal sample. Does survey non-participation independently affect sample composition, or are compositional changes in the sample driven mostly by mortality?

In a paper from 2013, Zajacova and Burgards showed how selective mortality changed the composition of baseline characteristics in longitudinal samples, gradually making the samples 'healthier, wealthier and wiser'. Using only baseline data, they isolated the effect of the selective mortality that occurs in a cohort over time. In the current study, we added selective survey participation to this analysis to ascertain whether the compositional changes in baseline characteristics that occur in a longitudinal sample because of selective mortality are affected by selective survey participation.

More specifically, we examined changes in the distribution of baseline characteristics over time in an ageing panel sample due to attrition caused by i) mortality and ii) mortality in combination with survey non-participation.

Method

Data

The analyses were based on a sample of 1,132 individual, born 1924 to 1934 and followed up from 1968 to 2011, first in the Swedish level-of-living survey (LNU) and later in The Swedish panel study of living conditions of the oldest old (SWEOLD). The LNU study is a nationally representative longitudinal survey of the Swedish adult population that was initiated in 1968 and is still running. The upper age limit for participation in the LNU study is 75 years. The SWEOLD study was initiated 1992 and includes all individuals previously included in LNU who have passed the age limit of 75 years.

The cohort sample used in the present study has been followed through six waves (LNU1968, LNU1974, LNU1981, LNU1991, LNU2000/ SWEOLD2002, and SWEOLD2011), from the age of 34 to 44 in 1968 to the age of 77 to 87 in 2011 (table 1). More than 90% participated in the first LNU wave. The response rates in the follow-up waves ranged between 73.1% (LNU2000) and 84.7% (SWEOLD2011). In the fourth follow-up, individuals too old to participate in LNU2000 (those aged over 75) were included in SWEOLD2002 instead. As a result, the fourth wave (T4) contains individuals included either in LNU2000 or in SWEOLD2002, providing five possible follow-up waves for all sampled individuals. Individuals born 1925 (n=67) were included in both LNU 2000 and SWEOLD 2002. In cases were individuals were non-response in one wave but interviewed in the other (n=9), their response/non-response in SWEOLD 2002 was used as the outcome for T4.

Table 1. Sample characteristics at baseline

Birth year	1924-1934
Age 1968, mean (SD)	39.0 (3.2)
Age span 1968	34-44
Age span 2011	77-87
Women %	49.9
Education beyond compulsory school (>7 years) %	40.0
Married/cohabitating %	84.8
Social class	
Manual workers %	45.0
Non-manual workers %	40.2
Other classes %	14.8
Sample response rates	
LNU 1968	90.5
LNU 1974	84.2
LNU 1981	78.8
LNU 1991	74.3
LNU 2000[a]	73.1
SWEOLD 2002	80.6
SWEOLD 2011	84.7
Total sample size	1132

LNU, The Swedish level-of-living Survey; SWEOLD, The Swedish panel study of living conditions of the oldest old [a]. Because of the age limit in LNU, part of the sample was not included in LNU 2000. Instead they were included in SWEOLD 2002. There is also an overlap of 67 individuals that were included in both year.

Both LNU and SWEOLD primarily use face-to-face interviews to gather data. The questionnaires used cover a broad range of topics, such as living conditions, family situation, health, health behaviours, and financial resources. One primary goal is to maintain a representative study population, and much effort has been made to achieve high response rates. In SWEOLD, indirect interviews are used when the older person is too frail or cognitively impaired to participate in an interview. Close relatives or health care personnel who know the respondent well are used as informants. For a more thorough description of the LNU and SWEOLD studies, see Fritzell and Lundberg (2007) and Lennartsson et al. (2014).

Baseline characteristics

Information on sex, age, education, and social class were collected from registers in 1967. Level of education was dichotomised into compulsory school (correspond to ≤ 7 years for this cohort in Sweden)

or above. Social class was measured as the proportion of manual workers, non-manual workers, and other classes (farmers, self-employed people, housewives, and people whose occupations were unclassified) in accordance with the Swedish socioeconomic classification (SEI) (Andersson, Erikson, & Wärneryd, 1981). The SEI is similar to the commonly used Erikson–Goldthorpe–Portocarero (EGP) class scheme (Bihagen, Nermo, & Erikson, 2010; Erikson & Goldthorpe, 1992). Self-reported civil status, analysed as the proportion of the sample who were married or cohabitating, was available for those interviewed at baseline. Register information on civil status from 1968 was used to ascertain the civil status of the baseline non-response group.

Analyses

To demonstrate how the sample changed over time as the result of mortality and non-participation, we divided the sample into three 3 groups: dead, drop outs (accumulated), and interviewed (with

complete follow-up), at each follow-up, as described in the flowchart in Figure 1. For each survey wave, the interviewed groups consisted of all individuals in the sample who responded to each survey up to and including that wave. Accordingly, the drop-outs consisted of all sample members still alive but who, for whatever reason, did not participate in at least one wave, up to and including that wave, i.e. the accumulated non-participation. The final group consisted of those in the sample who had died up to and including that wave. In 2011, after 43 years of follow-up, 47% of the original 1968 sample had died, and an additional 24% were alive but had been non-participant in at least one survey wave. Less than 29% of the original sample from 1968 were both alive 2011 and had no missing data points; i.e., had participated in all survey waves.

Figure 1. Flowchart of the sample 1968-2011, where each row corresponds to the distribution interviewed, drop-outs* and dead at the specific wave.
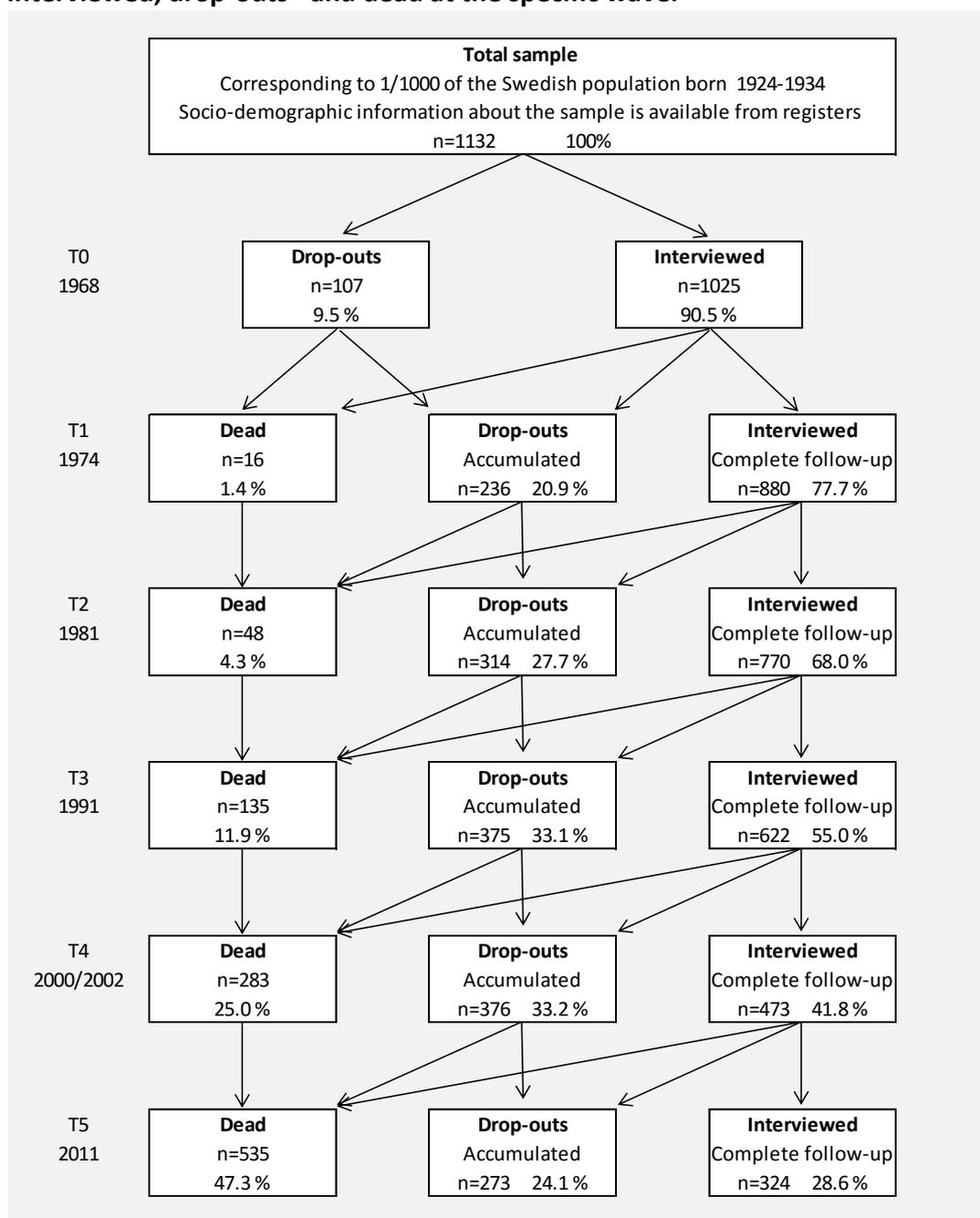
T0, baseline; T1, first wave; T2, second wave; T3, third wave; T4, fourth wave; T5, fifth wave.
*Those who dropped out at any wave were removed from the sample for all waves after drop-out.

The distribution of the baseline characteristics in the sample was then calculated for each wave. Initially, we describe how the distributions of baseline characteristics in the sample changed only as the result of mortality; that is, only those who died at each follow up were removed from the sample. In the next step we also removed the accumulated drop-outs from the sample, describing how the sample would look like if only those with complete follow-up were included in the analyses. None of the characteristics we measured were allowed to vary over time; all characteristics were measured at baseline only. All analyses were stratified by sex.

Then we tested the mortality risk for the accumulated drop-out group compared to the interview group with complete follow-up using Cox proportional hazard regression. Mortality was followed from 1st of January each survey year until 31st of December 2014.

Finally, we compared baseline characteristics and mortality risk between responders with complete follow-up, responders with incomplete follow-up, and non-responders, among those still alive at T5.

Results

The solid lines in figures 2 and 3 show how the distribution of baseline characteristics in the sample changed over time as mortality successively reduced the number of individuals in the sample. If mortality had been random rather than selective, the lines would be horizontal. Any slope in the lines indicates that mortality was associated with the characteristic measured. All estimates presented in table 2 and 3 are available in a supplementary table with 95% confidence intervals.

The dotted lines in figures 2 and 3 show the sample after excluding the drop-outs in addition to excluding those that had died. Hence, the dotted line shows the distribution of the baseline characteristics in a sample restricted to individuals with complete data. Any gap between the solid and the dotted lines is an indication that selective survey participation

has an additional impact (beyond mortality) on the composition of the longitudinal sample.

The distributions of demographic characteristics are shown in figure 2. Over time, the proportion of women in the sample increased from 50% to almost 60%. However, the solid line is almost horizontal up to the fourth wave (T4), indicating that the higher mortality among men does not affect the sample in any substantial way before the cohort reaches a higher age (age 57-67). Removing the drop-outs from the sample did not change the pattern; the dotted line is almost identical with the solid line. Hence, selective survey participation did not impact the effect of the selective mortality on the sex distribution in the sample.

Because of higher mortality among the oldest individuals in the sample, the mean baseline age in the sample increased. Higher mortality among the older individuals increased the mean baseline age by a bit less than half a year among women over the 43-year period. When drop-outs were also excluded, the difference increased to more than 0.6 years. Among men, mortality increased the mean baseline age with more than 0.8 years over the follow-up period. In contrast to women, the exclusion of drop-outs did not compound the effect of selective mortality among men.

Being married or cohabitating at baseline was associated with a lower mortality risk among both women and men. Over time, selective mortality gradually resulted in a sample that contained slightly more individuals who were married at baseline. Among men, however, this selection was compounded by selective survey participation. When the men who had dropped out were also excluded the proportion of men who, at baseline, were married increased by almost 5% units at T5 over the proportion attributable to mortality selection alone. This means that men who were unmarried at baseline had both a higher risk of dying and a higher risk of dropping out than men who were married at baseline.

Figure 2. Distribution of demographic characteristics in the 1924-1934 cohort in the LNU/SWEOLD sample 1968-2011 after mortality and drop-out*, presented with 95% confidence intervals
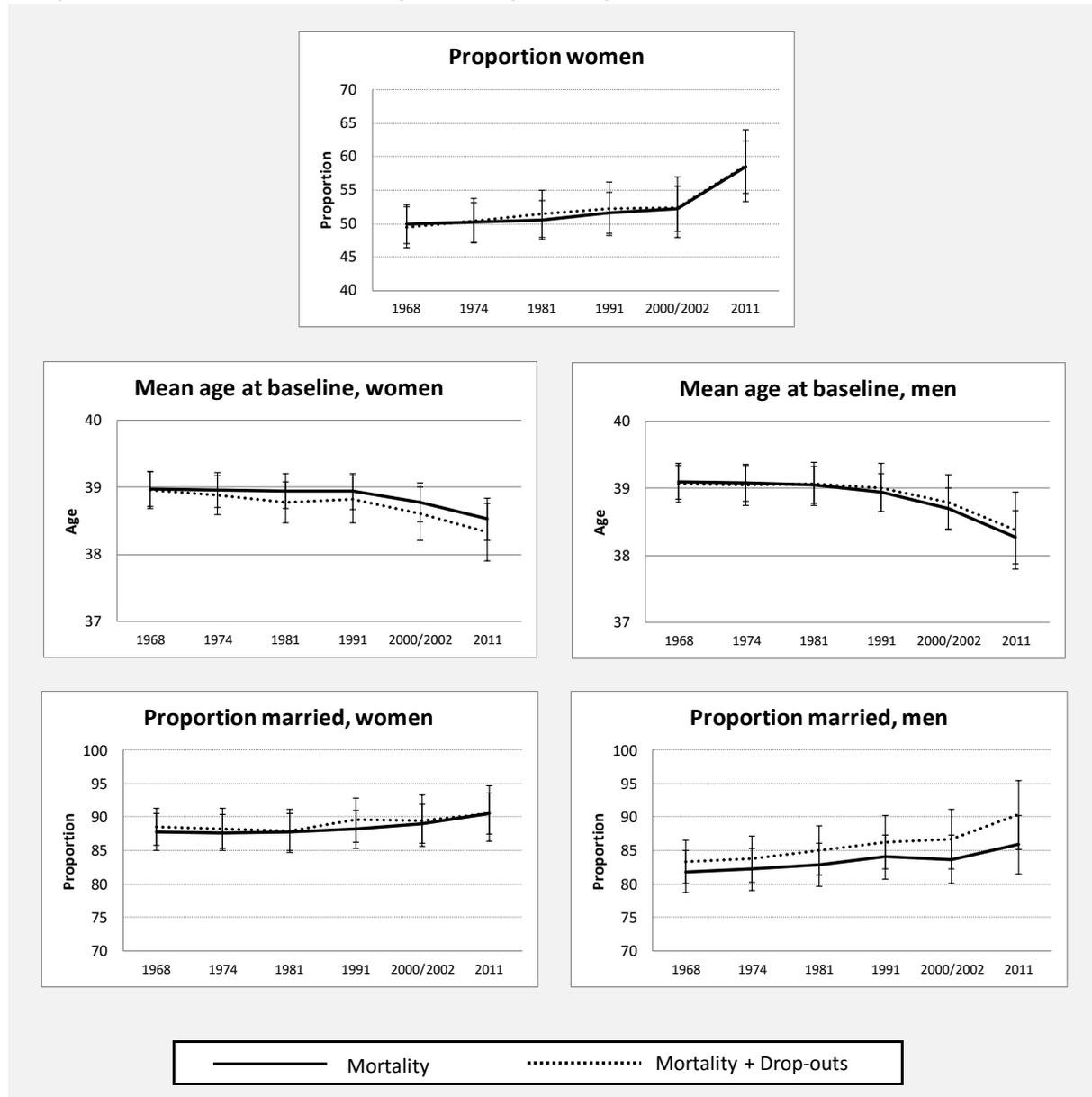
LNU, The Swedish level-of-living Survey; SWEOLD, The Swedish panel study of living conditions of the oldest old

*Those who dropped out at any wave were removed from the sample for all waves after drop-out.

Figure 3. Distribution of socioeconomic characteristics among the 1924-1934 cohort in the LNU/SWEOLD sample 1968-2011 after mortality and drop-out*, presented with 95 % confidence intervals
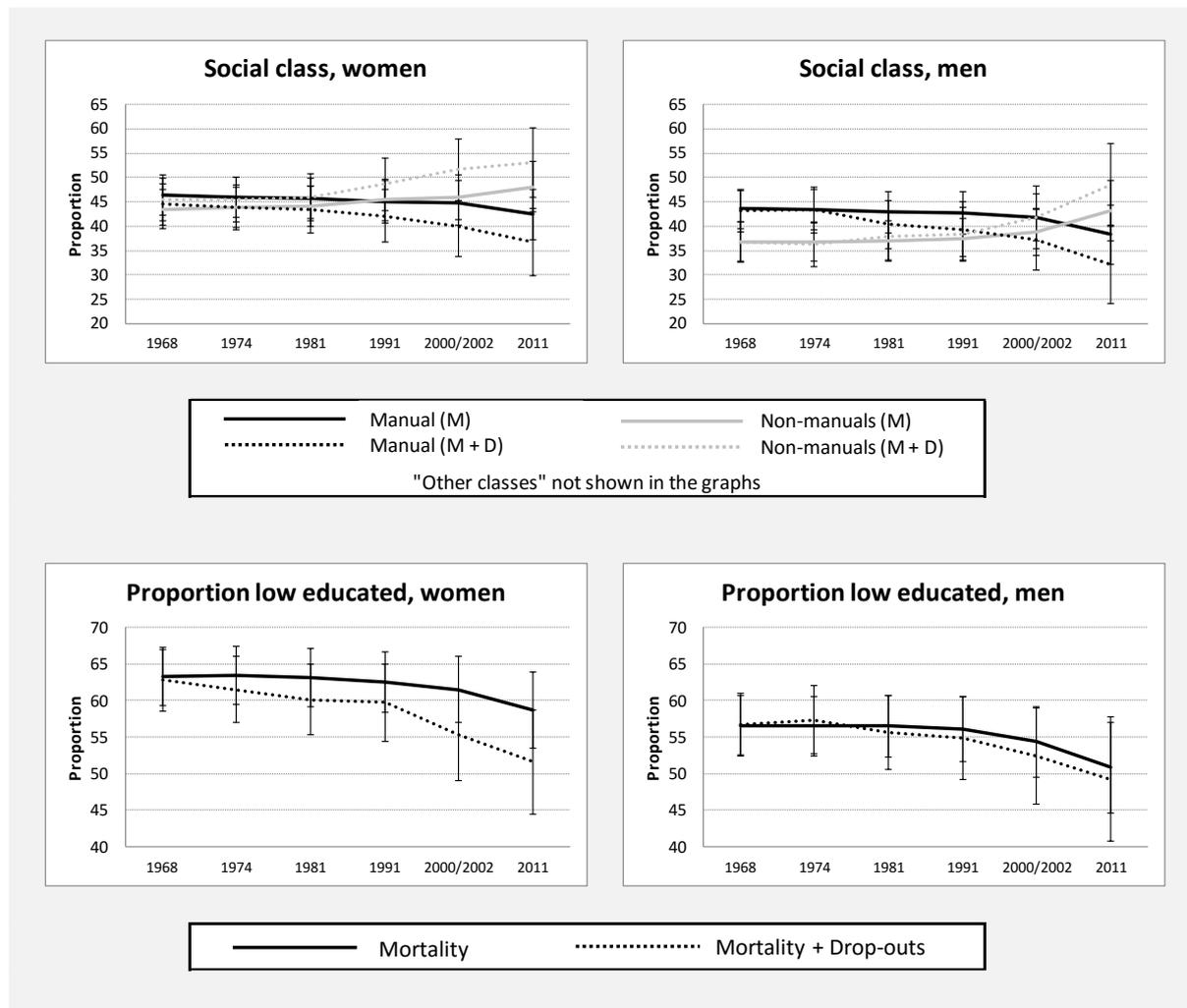
LNU, The Swedish level-of-living Survey; SWEOLD, The Swedish panel study of living conditions of the oldest old; M, mortality; D, drop-out

*Those who dropped out at any wave were removed from the sample for all waves after drop-out.

The distributions of socioeconomic characteristics are shown in figure 3. Changes in class structure were observed for both women and men. Over time, selective mortality decreased the proportion of manual workers from 46 to 42% among women and from 44 to 38% among men, and the proportion of non-manual workers in the sample increased. However, among both women and men, the changes in class structure were substantially compounded by selective survey non-participation. From the original sample in 1968 to the restricted sample in 2011, the proportion of manual workers decreased with more than 11% units for men and more than 9% units for women. This means that the 2011 sample consisted of a significantly larger proportion non-manual than

manual workers although the original sample had consisted of more manual than non-manual workers. The proportion of individuals from other classes (farmers, self-employed people, housewives, and people whose occupations were unclassified) showed only marginal fluctuation after mortality and survey non-participation were taken into account (not shown). For all survey waves, the proportion of other classes was approximately 20% for men and approximately 10% for women.

Selective mortality also changed the distribution of educational level among both women and men in the sample over time. Because of higher mortality among those with a compulsory school education than those with more than a compulsory school

education, the level of education in the sample increased as the proportion of those with more than a compulsory school education increased over time. In women but not men, selective survey participation compounded the changes in education. When the drop-outs were removed from the sample, the proportion women with more than a compulsory education increased with additional 7% units.

Table 2 show that the accumulated drop-out group consequently have a higher mortality risk than the interview group with complete follow-up, albeit not statistically significant at all times. Adjustment for sex and age slightly lowered the estimates for T0, T1, and T2, whilst the estimates for T3 up to T5 did not change after the adjustment.

Finally, table 3 show that at T5, when the people in the sample were between 77 and 87 years old, responders with complete follow-up were in general more educated, more likely to have had a non-manual occupation, and a lower mortality risk than responders with incomplete follow-up and the non-response group. In addition, non-responders were less likely to be women and to be married at baseline compared to responders with complete follow-up.

Table 2. Mortality risk for the accumulated drop-out group compared with the interview group with complete follow-up. For each survey wave, mortality is followed until 2014

Survey wave	N	Number of deaths	Crude HR	95 % CI	Adjusted[a] HR	95 % CI
T0 1968	1132	649	1.34	*1.05;1.70*	1.29	*1.01;1.64*
T1 1974	1116	633	1.24	*1.03;1.49*	1.19	*0.99;1.43*
T2 1981	1084	601	1.11	*0.93;1.32*	1.05	*0.88;1.25*
T3 1991	997	514	1.14	*0.96;1.37*	1.14	*0.95;1.36*
T4 2000/2002	849	366	1.27	*1.03;1.56*	1.28	*1.04;1.56*
T5 2011	597	114	1.32	*0.91;1.90*	1.33	*0.92;1.91*

[a] Adjusted for sex and birth year

Table 3. Baseline characteristics and mortality risk (follow-up until 2014) for the sample still alive at 2011 (T5), by response pattern

	Responders with complete follow-up	Responders with incomplete follow-up	Non-responders
Baseline characteristics			
Age 1968, mean	38.3	38.6	38.4
Women %	58.6	61.9	52.4
Education beyond compulsory school %	49.4	39.9	37.1
Married/cohabitating %	90.4	89.3	81.9
Manual workers %	34.9	45.8	50.5
Non-manual workers %	51.2	43.5	34.3
Mortality risk[a]			
Number of deaths	55	37	22
Crude HR (95 % CI)	1.00	1.35 (*0.89;2.05*)	1.26 (0.77;2.08)
Adjusted HR[a] (95 % CI)	1.00	1.36 (*0.90;2.07*)	1.27 (0.77;2.08)
N	324	168	105

[a] Adjusted for sex and birth year

Discussion

Summary of results

   Using longitudinal data (43 years of follow-up) on individuals born 1924-1934 (age 34-44 at baseline), we found that selective mortality changed the sample by successively increasing the proportion of women, people with non-manual occupations, younger birth cohorts, people with more than a compulsory education, and people who were married at baseline. When drop-outs were also excluded from the analyses, the sample selection found for the demographic variables did not change in any substantial way. After excluding both those who died during follow-up and drop-outs, the greatest change was found in two variables: civil status among men and mean birth year among women. That is, over time, a lower proportion of men who were unmarried at baseline participated than men who were married at baseline, and a lower proportion of women from older birth cohorts participated than women from younger birth cohorts.

   The effects of selective mortality on the socioeconomic characteristics of the sample were substantially compounded by the effects of selective survey participation. When we removed both the drop-outs and those who died during follow-up, the sample included substantially more people with non-manual occupations and a higher proportion of people with more than a compulsory education than when we only removed those who had died. Education in men was the only socioeconomic factor for which we did not find a compounding effect of survey non-participation.

   Finally, this study shows that individuals with incomplete follow-up have lower mortality risk than individuals with complete follow-up. This implies that analyses restricted to individuals with complete data not only skew the proportion individuals with low socioeconomic position. There is also an imminent risk of underestimate the occurrence of fundamental epidemiological events, such as mortality.

Strengths and limitations

   The data from the LNU and SWEOLD surveys provided a unique opportunity to investigate selection processes over time because of the rich information that was available for the total sample group. In 1968, LNU researchers pioneered the linkage of data from different sources when they combined register information with survey data to empirically study living conditions in Sweden for social policy purposes (Johansson, 1973). As a result, socio-demographic information is available on the total sample, both baseline responders and baseline non-responders. This contrasts with the situation in many other longitudinal studies, in which information on the baseline non-responders is usually limited (Cheshire, Ofstedal, Scholes & Schröder, 2011).The Swedish system, in which each resident is assigned a personal identification number that is used in nationwide registers, was also a prerequisite in this study, as it made it possible to follow mortality over time.

   One drawback of the data material, however, is the limited sample size. The sample size is too small to analyse the potential bias introduced by the selections in associational studies. Moreover, a larger sample would also have made it possible to distinguish between temporary non-response (i.e. individuals that come back at a later wave after missing one or more waves) and those who drop-out and never re-appear in the study again. These groups may very well differ in terms of demographics and socioeconomic conditions.

   The birth cohort included in the present study had high response rates in both the LNU and SWEOLD studies (between 73 and 91%). It is not clear whether the selective non-participation found in longitudinal studies with relatively high response rates can be assumed to be representative for longitudinal studies with lower response rates. It is possible that the selection processes would be stronger in a study with higher non-response, but it is also possible that the non-response is more selective in LNU and SWEOLD than in a study with lower response rate.

   Another limitation of this study is that we only investigated selection related to rather stable demographic and socioeconomic characteristics, not selection related to other individual resources, such as health, personality, and cognition. We made this choice because a major strength of our data was the availability of information about demographic and socioeconomic characteristics for the entire sample, not only those interviewed. Hence, we chose to include only factors about which information was available for the total sample.

   Further, the demographic and socioeconomic characteristics were measured only at baseline despite the time-dependent nature of level of education, social class, and civil status. Using only baseline information and not allowing the variables

to vary over time ignores any intra-individual changes that could have occurred over time. However, most of the variables in this study could be expected to remain relatively stable over time because they were measured when the sample members were in middle age; for instance, social mobility was low after middle age in these cohorts, especially mobility between manual workers and non-manual workers (Eriksson & Åberg, 1987; Sjögren Lindquist, 2006). During follow-up, however, civil status changed for some people in the sample. Our intention, however, was not to measure differences in mortality risk or the likelihood of survey participation that are associated with changes in civil status. Rather, our intention was to investigate how individuals that were not married/cohabiting in middle age were represented in a longitudinal survey over time (married/cohabiting was the predominant civil status in this cohort).

The substantial advantage of the time-invariant design was the isolation of the crude selections and the reduction of influence from confounding factors. Yet, age is still a potentially confounding factor as it is associated with mortality, as well as with social class and education. Individuals born earlier have a greater risk of having a less education and a less privileged social class, in addition to a higher mortality risk. However, as the age range in the sample is limited to 11 years, the confounding effect of age is likely to be limited. In fact, the analyses of differences in mortality between those with complete follow-up and the accumulated drop-out group showed that age and sex had a minor impact on the mortality differences between these groups.

Comparison with other studies

Few studies have explored how selective mortality and survey participation change the social composition of longitudinal samples over time. Our results are in line with the results of Zajacova and Burgard (2013), who also found that selective mortality changes the distribution of basic socio-demographic characteristics over time, making the sample successively contain more women, younger birth cohorts, people with a higher level of education, and a higher proportion of individuals who were married at baseline. Social class was not part of their study.

Zajacova and Burgard (2013) did test whether or not selective survey participation affected their results. In contrast to our results, they did not find any differences in the selections when they added

attrition for other reason than mortality to attrition caused by mortality. Selective survey participation was, however, outside the focus of their study, and they described their findings on attrition for other reason than mortality as preliminary. In addition, as previously noted, they did not include social class in their study – the characteristic in our study that was most influenced by selection.

Another difference between the studies was the follow-up time. Our results are based on longitudinal data from 43 years of follow-up. In Zajacova and Burgard's study, the follow-up time was 16 years. The degree of bias in estimates of health outcomes due to attrition has also been investigated in a UK study of individuals 50 years and older with six years of follow-up (Lacey, Jordan, & Croft, 2013). The result showed that only baseline non-response contributed to the bias; no further selection bias occurred because of non-response in subsequent waves. Although we did not investigate any biasing effect of the selection we found, it is noteworthy that our results suggest that selection becomes stronger over time. It would be of interest to further investigate differences between various longitudinal studies to see whether the compounded selection found in the present study are generalisable to other contexts and studies or are specific to the Swedish context and/or the long follow-up time.

Finally, our results contrast with the results of a 10-year follow-up study among 25 to 75-year-old Americans, which found that marital status better predicted survey participation among women than men (Radler & Ryff, 2010). The results of our study, on the other hand, suggest that marital status better predicts survey participation among men than women, meaning that non-responses are more likely to result in a sample in which non-married men are underrepresented to a greater extent than non-married women. We also found differences between women and men with regard to the compounding effect of education. Over time among women but not men, drop-out associated with low education (compulsory) compounded the attrition caused by mortality associated with low education. This indicates that non-response in longitudinal data are more likely to produce a sample in which the educational level is higher than in the population, and the difference will be more pronounced among women than men. Regarding social class, however, the patterns were very similar for women and men: significant sample selection was driven both by

selective mortality and selective survey participation.

Consequences of selection

Research suggests that selective attrition in longitudinal surveys may produce biased estimates of factors such as wealth, health, and labour force participation (Michaud, Kapteyn, Smith & Van Soest, 2011); cognition (Weir, Faul, & Langa, 2011); change in cognitive function (Rajan, Leurgans, Weuve, Beck & Evans, 2011); and the association between smoking and cognitive decline (Weuve et al., 2012). On the other hand, some researchers question whether there is a systematic relationship between health and attrition (Carter, Imlach-Gunasekara, McKenzie & Blakely, 2012), and several studies have also reported that selective attrition tends to cause only small biasing effects on estimates of longitudinal changes (Carter et al., 2012; Deeg, 2002; Salthouse, 2013).

Although the findings reported in the literature are inconsistent, there are good reasons for examining the potential effects of sample selection in longitudinal studies of old people, as attrition in this group tends to be related to disability and mortality (Deeg, 2002; Radler & Ryff, 2010). We should not assume that non-response among individuals 75 years and older is random, and the mortality rate is far from ignorable and random.

When considering the consequences of attrition, it is important to separate attrition that occurs as a result of mortality and attrition for reasons other than mortality. Attrition for reasons other than death (except for emigration) appears only in study samples and does not reflect real changes in the target populations. It might therefore be a greater source of bias than attrition caused by mortality (Brilleman, Pachana, & Dobson, 2010), which produces the same selections that occur in the target populations and thus reflects real changes in those populations.

However, any analytical research, including comparisons of social groups, must include a consideration of whether or not the results have been affected by selective attrition. That is, it is important to examine whether selective mortality has contributed to an underestimation of the problems experienced by disadvantaged groups because those who are worse off are already dead when the outcomes are measured – a 'healthy survivor' effect (Murphy et al., 2011).

The results of the present study show that the majority of the compositional changes in the LNU and SWEOLD cohort we examined were rather modest and mostly driven by mortality. However, for some characteristics, class in particular, the selection was not ignorable, and in addition, was substantially compounded by survey non-participation. The substantial changes in social class composition found in our study indicate that selection in longitudinal samples should be of particular concern to researchers in the field of social inequality in health. For example, higher mortality among manual than non-manual workers may lead to greater intra-group changes in the proportion of manual than non-manual workers with poor health. If disadvantaged individuals with poor health are more likely than others to die, and if disadvantaged individuals are more likely to be found among manual workers, health selection will be stronger among the manual workers; i.e., the proportion with poor health will decrease more than among non-manual workers. As a consequence, there will appear to no class-related health inequality or the inequality will appear to decrease over time. Any analysis of health inequality or cumulative advantages should address the issue of selective mortality (Willson et al., 2007), and according to our results, selective survey participation should also be considered.

Implications

Some attrition always occurs in longitudinal surveys. When analysing longitudinal data, it is important to perform sensitivity analyses, to test all possible selections processes (Geneletti, Mason, & Best, 2011; Philipson, Ho, & Henderson, 2008) and, if possible, to use the unbalanced sample (i.e. the sample that also includes those with incomplete follow-up) (Michaud, Kapteyn, Smith & Van Soest, 2011). There are now well-established statistical methods that address missing data in analyses of longitudinal data (Muniz-Terrera & Hardy, 2014; Palmer & Royall, 2010). However, few methods can be assumed to be completely robust to misspecification after selective attrition.

Keeping the attrition low should therefore be of utmost importance to anyone running a longitudinal study. Although the forces behind selective mortality are outside the researchers' control, selective survey participation may, to some extent, be avoidable (Watson & Wooden, 2014). The gaps between the mortality and the drop-out curves (figures 2 and 3) are not inevitable. The effort invested, and the

methods used, to keep sampled individuals in a longitudinal survey, in combination with decisions taken during analyses, partly determine the width of this gap.

Achieving low attrition may be particularly challenging when the sample includes older individuals. In addition to the selective mortality that substantially affects surveys of older people, longitudinal surveys of older people are usually also characterised by attrition related to disability and cognitive impairment (Deeg, 2002; Radler & Ryff, 2010). Thus, conducting surveys among old people is extremely challenging and resource-demanding. It requires fieldwork strategies adjusted to older respondents, such as the use of indirect interviews, the ability to include institutionalised individuals, and the use of different interview modes, such as telephone interviews or postal questionnaires as alternatives to face-to-face interviews. The results of analyses of health-related prevalence rates might otherwise be severely biased (Kelfve et al., 2013; Lundberg & Thorslund, 1996).

Conclusion

Longitudinal surveys are invaluable tools for studying the aging process. However, results must be interpreted in the light of different selection processes. Selective mortality changes the social composition of a sample over time. Selective survey participation might compound this selection by giving already disadvantaged groups even less representation. In the present study, we found that selection changed the demographic characteristics of our sample in rather modest ways, whilst it changed the socioeconomic characteristics of the sample considerably over time. In addition, the compounded selection that occurred because of survey non-participation substantially affected the distribution of social class and education among women in the sample over time. This might have widespread consequences for research into social inequalities in health.

Selective mortality is unavoidable in an ageing sample. However, other selections might be partly avoided or dealt with by keeping non-response low, allowing individuals to return to the sample, and when possible, using statistical methods that enable the use of incomplete data in analyses; i.e. that can handle respondents with partial non-response.

References

Andersson, L-G., Erikson, R., & Wärneryd, B. (1981). Att beskriva den sociala strukturen: utvärdering av 1974 års förslag till socio-ekonomisk indelning [A memo on the 1974 socioeconomic classification]. *Statistisk Tidskrift, 19*, 113-136.

Banks, J., Muriel, A., & Smith, J. P. (2011). Attrition and health in ageing studies: evidence from ELSA and HRS. *Longitudinal and life course studies, 2*(2), 101-126. https://doi.org/10.14301/llcs.v2i2.115

Barford, A., Dorling, D., Smith, G. D., & Shaw, M. (2006). Life expectancy: women now on top everywhere. *BMJ, 332*(7545), 808. https://doi.org/10.1136/bmj.332.7545.808

Bartlett, J. W., Carpenter, J. R., Tilling, K., & Vansteelandt, S. (2014). Improving upon the efficiency of complete case analysis when covariates are MNAR. Biostatistics, 15(4), 719–730. https://doi.org/10.1093/biostatistics/kxu023

Bihagen, E., Nermo, M., & Erikson, R. (2010). Social class and employment relations: Comparisons between the ESeC and EGP class schemas using European data. In D. Rose, & E. Harrison (eds.), *Social Class in Europe: An introduction to the European Socio-economic Classification* (pp. 89-113). London and New York: Abingdon(Oxon): Routledge.

Brilleman, S. L., Pachana, N. A., & Dobson, A. J. (2010). The impact of attrition on the representativeness of cohort studies of older people. *BMC medical research methodology, 10*, 71. https://doi.org/10.1186/1471-2288-10-71

Carter, K. N., Imlach-Gunasekara, F., McKenzie, S. K., & Blakely, T. (2012). Differential loss of participants does not necessarily cause selection bias. *Australian and New Zealand journal of public health, 36*(3), 218-222. https://doi.org/10.1111/j.1753-6405.2012.00867.x

Chatfield, M. D., Brayne, C. E., & Matthews, F. E. (2005). A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. Journal of Clinical Epidemiology, 58(1), 13-19. https://doi.org/10.1016/j.jclinepi.2004.05.006

Cheshire, H., Ofstedal, M. B., Scholes, S., & Schröder, M. (2011). A comparison of response rates in the English Longitudinal Study of Ageing and the Health and Retirement Study. *Longitudinal and life course studies, 2*(2), 127-144. https://doi.org/10.14301/llcs.v2i2.118

de Souto Barreto, P. (2012). Participation bias in postal surveys among older adults: The role played by self-reported health, physical functional decline and frailty. *Archives of Gerontology and Geriatrics, 55*(3), 592-598. https://doi.org/10.1016/j.archger.2012.03.008

Deeg, D. J. H. (2002). Attrition in longitudinal population studies: Does it affect the generalizability of the findings? An introduction to the series. *Journal of Clinical Epidemiology, 55*(3), 213-215. https://doi.org/10.1016/S0895-4356(01)00472-3

Erikson, R., & Åberg, R. (Eds.). (1987). *Welfare in transition: a survey of living conditions in Sweden, 1968-1981*. Oxford: Clarendon Press.

Erikson, R., & Goldthorpe, J. H. (Eds.). (1992). *The constant flux: A study of class mobility in industrial societies*. Oxford: Clarendon Press.

Fejer, R., Hartvigsen, J., Kyvik, O., Jordan, A., Christensen, H. W., & Høilund-Carlsen, P. F. (2006). The Funen Neck and Chest Pain study: analysing non-response bias by using national vital statistic data. *European journal of epidemiology, 21*(3), 171-180. https://doi.org/10.1007/s10654-006-0006-x

Ferraro, K. F., & Shippee, T. P. (2009). Aging and cumulative inequality: how does inequality get under the skin? *The Gerontologist, 49*(3), 333-343. https://doi.org/10.1093/geront/gnp034

Ferraro, K. F., Shippee, T. P., & Schafer, M. H. (2009). Cumulative inequality theory for research on aging and the life course. In V. L. Bengtson, D. Gans, N. M. Putney & M. Silverstein (Eds.), *Handbook of theories of aging - 2nd ed.* (pp. 413-433). New York: Springer Publishing Company, LLC.

Ferrie, J. E., Kivimäki, M., Singh-Manoux, A., Shortt, A., Martikainen, P., Head, J., Marmot, M., Gimeno, D., De Vogli, R., Elovainio, M., & Shipley, M. J. (2009). Non-response to baseline, non-response to follow-up and mortality in the Whitehall II cohort. *International journal of epidemiology, 38*, 831-837. https://doi.org/10.1093/ije/dyp153

Fors, S., Lennartsson, C., & Lundberg, O. (2011). Live long and prosper? Childhood living conditions, marital status, social class in adulthood and mortality during mid-life: a cohort study. *Scandinavian journal of public health, 39*(2), 179-186. https://doi.org/10.1177/1403494810395823

Fritzell, J., & Lundberg, O. (Eds.). (2007). *Health inequalities and welfare resources: continuity and change in Sweden*: Bristol: Policy Press.

Galea, S., & Tracy, M. (2007). Participation rates in epidemiologic studies. *Annals of epidemiology, 17*(9), 643-653. https://doi.org/10.1016/j.annepidem.2007.03.013

Geneletti, S., Mason, A., & Best, N. (2011). Adjusting for selection effects in epidemiologic studies: why sensitivity analysis is the only "solution". *Epidemiology, 22*(1), 36-39. https://doi.org/10.1097/EDE.0b013e3182003276

Hardie, J. A., Bakke, P. S., & Mørkve, O. (2003). Non-response bias in a postal questionnaire survey on respiratory health in the old and very old. *Scandinavian journal of public health, 31*(6), 411-417. https://doi.org/10.1177/140349480303100603

Johansson, S. (1973). The Level of Living Survey: A presentation. *Acta sociologica, 16*(3), 211-219. https://doi.org/10.1177/000169937301600304

Kelfve, S., Thorslund,M ., & Lennartsson, C. (2013). Sampling and non-response bias on health-outcomes in surveys of the oldest old. *European journal of ageing, 10*(3), 237-245. https://doi.org/10.1007/s10433-013-0275-7

Kjøller, M., & Thoning, H. (2005). Characteristics of non-response in the Danish Health Interview Surveys, 1987–1994. *European journal of public health, 15*(5), 528-535. https://doi.org/10.1093/eurpub/cki023

Klein, D. J., Elliott, M. N., Haviland, A. M., Saliba, D., Burkhart, Q., Edwards, C., & Zaslavsky, A. M. (2011). Understanding nonresponse to the 2007 Medicare CAHPS survey. *The Gerontologist, 51*(6), 843-855. https://doi.org/10.1093/geront/gnr046

Korkeila, K., Suominen, S., Ahvenainen, J., Ojanlatva, A., Rautava, P., Helenius, H., & Koskenvuo, M. (2001). Non-response and related factors in a nation-wide health survey. *European journal of epidemiol, 17*(11), 991-999. https://doi.org/10.1023/A:1020016922473

Kuh, D., Cooper, R., Hardy, R., Richards, M., & Ben-Shlomo, Y. (Eds.). (2014). *A life course approach to healthy ageing.* New York: Oxford University Press.

Lacey, R. J., Jordan, K. P., & Croft, P. R. (2013). Does Attrition during Follow-Up of a Population Cohort Study Inevitably Lead to Biased Estimates of Health Status? *PLoS one, 8*(12), e83948. https://doi.org/10.1371/journal.pone.0083948

Lennartsson, C., & Lundberg, O. (2007). What's marital status got to do with it? Gender inequalities in economic resources, health and functional abilities among older adults. In J. Fritzell & O. Lundberg (Eds.), *Health inequalities and welfare resources* (pp. 179-198). Bristol: Policy Press.

Lennartsson, C., Agahi, N., Hols-Salén, L., Kelfve, S., Kåreholt, I., Lundberg, O., Parker, M. G., & Thorslund, M. (2014). Data Resource Profile: The Swedish Panel Study of Living Conditions of the Oldest Old (SWEOLD). *International journal of epidemiology, 43*(3), 731-738. https://doi.org/10.1093/ije/dyu057

Sjögren Lindquist, G. (2006). Late careers and career exits in Sweden. In H. P. Blossfeld, S. Buchholz & D. Hofäcker (Eds.), *Globalization, uncertainty and late careers in society* (pp. 211-234)*.* London and New York: Routledge.

Lundberg, O., & Thorslund, M. (1996). Fieldwork and measurement considerations in surveys of the oldest old. *Social Indicators Research, 37*(2), 165-187. https://doi.org/10.1007/BF00315527

Lynn, P. (2011). A fishy story: the roles of rods and nets in maintaining representative longitudinal survey samples. *Longitudinal and Life Course Studies, 2*(2), 98-100. http://dx.doi.org/10.14301/llcs.v2i2.117

Mackenbach, J. P. (2012). The persistence of health inequalities in modern welfare states: the explanation of a paradox. *Social science & medicine, 75*(4), 761-769. https://doi.org/10.1016/j.socscimed.2012.02.031

Michaud, P.-C., Kapteyn, A, Smith, J. P., & Van Soest, A. (2011). Temporary and permanent unit non-response in follow-up interviews of the Health and Retirement Study. *Longitudinal and Life Course Studies, 2*(2), 145-169. http://dx.doi.org/10.14301/llcs.v2i2.114

Muniz-Terrera, G., & Hardy, R. (2014). Longitudinal data analysis in studies of healthy ageing. In D. Kuh, R. Cooper, R. Hardy, M. Richards & Y. Ben-Shlomo (Eds.), *A life course approach to healthy ageing* (pp. 79-90). New York: Oxford University Press.

Murphy, T. E., Han, L., Allore, H. G., Peduzzi, P. N., Gill, T. M., & Lin, H. (2011). Treatment of death in the analysis of longitudinal studies of gerontological outcomes. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 66A*(1), 109-114. https://doi.org/10.1093/gerona/glq188

Palmer, R.F., & Royall, D.R. (2010). Missing Data? Plan on It! *Journal of the American Geriatrics Society, 58*(S2), 343-348. https://doi.org/10.1111/j.1532-5415.2010.03053.x

Phelan, J. C., Link, B. G., & Tehranifar, P. (2010). Social Conditions as Fundamental Causes of Health Inequalities. *Journal of Health and Social Behavior, 51*(1 suppl), S28-S40. https://doi.org/10.1177/0022146510383498

Philipson, P. M., Ho, W. K., & Henderson, R. (2008). Comparative review of methods for handling drop-out in longitudinal studies. *Statistics in medicine, 27*(30), 6276-6298. https://doi.org/10.1002/sim.3450

Radler, B. T., & Ryff, C. D. (2010). Who participates? Accounting for longitudinal retention in the MIDUS national study of health and well-being. *Journal of Aging and Health, 22*(3), 307-331. https://doi.org/10.1177/0898264309358617

Rajan, K. B., Leurgans, S. E., Weuve, J., Beck, T. L., & Evans, D. A. (2011). Effect of Demographic Risk Factors on the Change in Cognitive Function in the Presence of Non-Participation and Truncation due to Death. *Journal of biometrics & biostatistics, 3*(1), 1-7. https://doi.org/10.4172/2155-6180.S3-001

Salthouse, T. A. (2013). Selectivity of attrition in longitudinal studies of cognitive functioning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 69*(4), 567-574. https://doi.org/10.1093/geronb/gbt046

Stringhini, S., Dugravot, A., Shipley, M., Goldberg, M., Zins, M., Kivimäki, M., Severine, S., & Singh-Manoux, A. (2011). Health behaviours, socioeconomic status, and mortality: further analyses of the British Whitehall II and the French GAZEL prospective cohorts. *PLoS medicine, 8*(2), e1000419. https://doi.org/10.1371/journal.pmed.1000419

Thorslund, M., Wastesson, J. W., Agahi, N., Lagergren, M., & Parker, M. G. (2013). The rise and fall of women's advantage: a comparison of national trends in life expectancy at age 65 years. *European journal of ageing, 10*(4), 271-277. https://doi.org/10.1007/s10433-013-0274-8

Torssander, J., & Erikson, R. (2010). Stratification and mortality—A comparison of education, class, status, and income. *European Sociological Review, 26*(4), 465-474. https://doi.org/10.1093/esr/jcp034

Umberson, D. (1992). Gender, marital status and the social control of health behavior. *Social science & medicine, 34*(8), 907-917. https://doi.org/10.1016/0277-9536(92)90259-S

Watson, N., & Wooden, M. (2014). Re-engaging with survey non-respondents: evidence from three household panels. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 177*(2), 499-522. https://doi.org/10.1111/rssa.12024

Vaupel, J. W., Zhang, Z., & van Raalte, A. A. (2011). Life expectancy and disparity: an international comparison of life table data. *BMJ Open, 1*, e000128. https://doi.org/10.1136/bmjopen-2011-000128

Weir, D. R., Faul, J. D., & Langa, K. M. (2011). Proxy interviews and bias in the distribution of cognitive abilities due to non-response in longitudinal studies: a comparison of HRS and ELSA. *Longitudinal and Life Course Studies, 2*(2), 170-184. https://doi.org/10.14301/llcs.v2i2.116

Weuve, J., Tchetgen, E. J., Glymour, M. M., Beck, T. L., Aggarwal, N. T, Wilson, R. S., Evans, D. A., & Mendes de Leon, C. F. (2012). Accounting for bias due to selective attrition: the example of smoking and cognitive decline. *Epidemiology, 23*(1), 119-128. https://doi.org/10.1097/EDE.0b013e318230e861

Willson, A. E., Shuey, K. M., & Elder, G. H. (2007). Cumulative advantage processes as mechanisms of inequality in life course health. *American Journal of Sociology, 112*(6), 1886-1924. https://doi.org/10.1086/512712

Zajacova, A., & Burgard, S. A. (2013). Healthier, Wealthier, and Wiser: A Demonstration of Compositional Changes in Aging Cohorts Due to Selective Mortality. *Population Research and Policy Review, 32*(3), 311-324. https://doi.org/10.1007/s11113-013-9273-x

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

SPECIAL SECTION: Transition to young adulthood
Editors Marlis Buchmann, Tina Malti, Heike Solga

PAPERS

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