VOLUME 4 | ISSUE 3 | **OCTOBER 2013**

ISSN 1757-9597

Longitudinal and Life Course Studies: International Journal

Published by

SLLS Society for Longitudinal and Life Course Studies

Inside this issue

Longview
Promoting Longitudinal and Life Course Studies

Co-sponsors:

VU university medical center

ISER INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

LLCS EDITORIAL BOARD

SUBSCRIPTIONS

For immediate access to current issue and succeeding 2013 Issues (Volume 4) and all three Issues published in 2012 (Volume 3):

Individual rate: **£20 – discounted rate £55 for 3 years** (Euro and US Dollar equivalents)

Library **rate: £200** (Euro and US Dollar equivalents) to register any number of library readers.

Society for Longitudinal and Life Course Studies (SLLS) Membership: **£65,** ***Student*** **£26,** ***Corporate*** **£300** (Euro and Dollar equivalents).

SLLS members have automatic free access to the journal among many other benefits. For further information about SLLS and details of how to join including the application form go to http://www.slls.org.uk/pages/membership.shtml

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

Print on demand

An attractive A5 printed version for each Issue of the journal, including all back numbers, is available at a price of £10 (Euro and US dollar equivalents) plus packaging and postage. If you want to take advantage of this opportunity of having hard copies to order, go to http://www.slls.org.uk/pages/publications.shtml where you will find a link to our online bookshop. Depending on distance, you will receive the order within two weeks of submitting it.

Please note

The reselling of personal registration subscriptions and hard copies of the journal is prohibited. Personal subscription and orders must be purchased with a debit or credit card - proof of identity may be requested.

SLLS disclaimer

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

CONTENTS

PAPERS

KEYNOTE LECTURE

RESEARCH NOTE

This journal can be accessed online at: www.llcsjournal.org

Published by
The Society for Longitudinal and Life Course Studies
Institute of Education, 20 Bedford Way, London, WC1H 0AL
info@slls.org.uk http://www.slls.org.uk

A double prevention: how maternal education can affect maternal mental health, child health and child cognitive development

Mariachiara Di Cesare, Imperial College London, UK
Ricardo Sabates, University of Sussex, UK
r.sabates@sussex.ac.uk
Keith M. Lewin, University of Sussex, UK

(Received January 2013 Revised August 2013) doi:10.14301/llcs.v4i3.233

Abstract

This study uses the longitudinal data of Young Lives for Peru to investigate the protective role that maternal education has for children whose mothers suffer from mental health problems. Our first set of findings confirms previous research in this area by showing that maternal education is associated with reduced risk of mental health problems for mothers, and with improved nutrition and cognitive development for their children. We further find that maternal education reduces the burden of maternal mental health problems on child development. This is particularly the case for children of mothers with high levels of education. Unfortunately, for children of mothers with low levels of education, maternal mental health problems continue to predict poor nutritional status and poor cognitive development for children. These results suggest that monitoring and support may be especially important for mothers with lower levels of education if inequalities across generations are to be reduced.

Keywords: maternal education, maternal mental health, child nutrition, child cognitive development, Young Lives Study

1. Introduction

Mental health represents an important cause of the burden of morbidity and disability in developed and developing countries (Harpham el al., 2003; Murray, Hipwell, Hooper, Stein, & Cooper, 1996; Stewart, 2007). It is estimated that around 450 million people suffer from mental health problems globally (WHO, 2010a), and that women have a higher susceptibility to mental disorders than men (WHO, 2010b). Poor maternal mental health not only affects women's overall wellbeing but has important inter-generational consequences, which are transmitted directly through the mother-child interactions (Murray, 1992) and indirectly through the family environment and parental relationships (Grace, Evindar, & Stewart, 2003; Harpham, Huttly, De Silva, & Abramsky, 2005; Rahaman, Iqbal, Bunn, Lovel, & Harrington, 2004). For instance, postpartum depression has been shown to have strong effects on child development, impairing both growth (Harpham et al., 2005; Patel, DeSouza, & Rodrigues, 2003; Rahaman, Harrington, & Bunn, 2002; Rahaman et al., 2004; Stewart, 2007), as well as cognitive, emotional, and behavioural outcomes (Black et al., 2007; Lyons-Ruth, Zoll, Connell, & Grunebaum, 1986; Murray, 1992; Murray & Cooper, 1997; Patel, Rahman, Jacob, & Hughes, 2004; Petterson & Albers, 2001).

Previous research has shown that the overall effect of maternal mental health on child

development depends on a number of risk and protective factors, of which education plays a key role (WHO, 2010a). Instability in partnership, as a risk factor, not only increases the likelihood of poor maternal mental health and impacts on the emotional development of children, but its overall effect has been observed to be different for mothers with high levels of education, compared with mothers with low levels of education (Wang, Anderson, & Florence, 2011). Social support, as a protective factor, is provided for mothers to overcome mental health difficulties and to assist with children's needs. Yet, for mothers with high levels of education, social support could be more effective at dealing with the consequences of maternal mental health problems on child development (Feinstein, Duckworth & Sabates, 2008).

Based on previous studies showing that education is a key protective factor for maternal mental health (Chevalier & Feinstein, 2006; Robertson, Celasun, & Stewart, 2003; Ross & Mirowsky, 1999), this paper proposes to further investigate the direct and indirect effects of mothers' education on maternal mental health and on child development. Two research questions are proposed: (i) Is maternal education directly associated with reduced likelihood of maternal mental health problems and improved developmental outcomes for their children? (ii) Can maternal education indirectly reduce the burden of maternal mental health problems on developmental outcomes for children? To empirically answer these questions, this paper uses the longitudinal data of Young Lives for Peru.

2. Setting of hypotheses: the role of maternal education

Education is considered to be one of the main instruments to reduce the transmission of social inequalities within and between generations (CSDH, 2008; Kerr & West, 2010). In particular, education is crucial for empowering women, giving them knowledge and skills not only to participate in the labour market and in society (Card, 1999), but also to be more efficient at bringing up their children (Feinstein et al., 2008). Research has shown that women with more education place a higher value to their own health and the health of those for whom they care (Duncan & Brooks-Gunn, 2000); they also have more influence on household decisions and higher levels of civic participation (Malik & Courtney, 2011). Education can also enable women to provide a cognitively stimulating environment for their children, use language which is developmentally enhancing and support the physical, social, and emotional development of their children (Carneiro, Meghir, & Parey, 2013).

However, in the delicate equilibrium of intra-household relations, the mother-child interactions can be heavily affected by the mental health of the mother (Knitzer, Theberge, & Johnson, 2008). Mental disorder during pregnancy and the months following childbirth can alter the home environment (Rahaman et al., 2002). It can also reduce mothers' engagement with, and alter mothers' speech towards, the child (Murray, 1992; Murray, Kempton, Woolgar, & Hooper, 1993) and impact on the overall quality of parenting provision (Murray & Cooper, 1997). As a result, children of mothers who suffer from mental disorders tend to be affected in their growth (Harpham et al., 2005; Patel et al., 2003; Rahaman et al., 2004), are likely to show emotional and behavioural problems (Field et al., 1988), as well as high rates of insecure attachment and depressive behaviours (Edhborg, Lundh, Seimyr, & Windstroem, 2001; Lyons-Ruth et al., 1986; Murray, 1992; Murray et al., 1996; Teti, Gelfand, Messinger, & Isabella, 1995). Poor maternal mental health also has important consequences on children's cognitive functioning and their overall school achievements (Black et al., 2007; Cogill, Caplan, Alexandra, Robson, & Kumar, 1986; Galler, Harrison, Ramsey, Forde, & Butler, 2000; Lyons-Ruth et al., 1986; Murray et al., 1993; Patel et al., 2003; Petterson & Albers, 2001).

Given that maternal education is a key factor for enhancing child development and that maternal mental health problems can hinder this process, the role of maternal education as a protective factor becomes crucial. In order to respond to the two research questions proposed in this paper we formulate the following three hypotheses. First, education can reduce the risk of mental health disorders, in particular depression. This is because education has direct effects on self-esteem, confidence and locus of control which impacts on the overall mental wellbeing (Ross & Mirowsky, 1999). Previous studies have found that education has a direct association with depression, in particular for women (Chevalier & Feinstein, 2006), that the impact of education on mental health strengthens with age (Miech & Shanahan, 2000), and that the

relationship between education and women's mental health is particularly important during motherhood (Mirowsky & Ross, 2002; Robertson et al., 2003).

Our second hypothesis is that maternal education can attenuate the negative effects of maternal mental health disorders on child development. Due to their education, mothers who suffer from minor psychiatric disorders can be empowered to better manage their health condition (Duncan & Brooks-Gunn, 2000; Paxson & Schady, 2005). Augustine and Crosnoe (2010) show a moderate effect of maternal education on the relationship between maternal depression and children's school achievements. In other words, the magnitude of the relationship between maternal depression and children's school attainment is reduced when maternal education is introduced as a control in the analysis.

Importantly, our third hypothesis is that education can change the nature of the relationship between maternal mental health problems and child development, moderating the consequences of maternal mental health for children and hence affecting the inter-generational transmission of disadvantage (Feinstein et al., 2008). Previous research has shown that, the association between maternal mental disorders in the post-birth period and child development, is stronger among families living in socio-economic deprivation (Stewart, 2007). This suggests that maternal mental illness does not affect children equally within the population. It is possible that maternal education can enhance or hinder existing inequalities (CSDH, 2008).

3. Methods

3.1 Data and sample

We use data from the first and the second round of Young Lives Longitudinal Study (YL) Peru. The main aim of YL cohort study is to analyse how poverty affects children's wellbeing (Boyden, 2006). The YL Peru study started in 2002 in 20 sentinel sites, selected on a multi-stage sampling protocol based on poverty status. In each geographical site, households were randomly selected for the study (Wilson, Huttly, & Fenn, 2006). The selection of households was based on the presence of children belonging to two age cohorts. The youngest cohort was children aged between 6 and 18 months (baseline sample 2,052) and the oldest cohort consisted of children between 7 to 8 years old (baseline sample 714). The second round of YL Peru was conducted in 2006-2007 when the two cohorts were 4 to 5 years old and 11 to 12 years old, respectively.

This paper uses data exclusively from the youngest cohort, as we have information on maternal mental health after birth, as well as on child development, for at least two points during childhood, at age 1 (nutritional status) and at age 5 (cognitive development). The second round of YL Peru included 1,963 children from round 1; of these only 136 had missing responses for one of the key variables (maternal mental health, child nutritional status at age 1 and child cognitive development at age 5). Since maternal mental health, child nutritional status at age 1 and child cognitive development at age 5 are the outcome variables, and for the rest of the factors used in the analysis, data loss due to missing responses was far less than 1 per cent, we opted not to undertake any imputation. Our final sample consists of 1,821 mothers and their respective children.

The sample strategy for YL has implications for empirical analysis and interpretation of results. First, results from YL cannot be generalised at the country level. The sample as a whole is not nationally representative, but households are representative within geographical sites. Secondly, social and economic gaps between rich and poor households are underestimations of existing inequalities in Peru as a whole. This is because geographical sites were selected based on poverty status, and hence social and economic inequalities reflect the situations of households already living in the selected poor areas. Thirdly, although attrition is a problem in longitudinal studies (3.5% for YL Peru between first and second rounds), Outes-Leon and Dercon (2008) and Sanchez (2009) have found that for YL the attrition is relatively small, and that there are no systematic differences between households who stopped participating in the study and the rest of the sample.

3.2 Selection of variables

Maternal mental health is measured at the interview by the Self-Reporting Questionnaire 20 items (SRQ20). The SRQ20 has been developed by the World Health Organisation (WHO) as an instrument to detect minor psychiatric disorders (WHO, 1994), but without the ability to separate anxiety from depression (Harpham et al., 2005). The

SRQ20 is based on a series of 20 items/symptoms scoring yes (1) or no (0) with the past 30 days as reference period (hence the maximum score is 20). The variable has been used in previous studies as a dichotomous indicator where more than 7 positive responses to this scale is a proxy measure for the risk of depression. In this paper, we utilise the whole scale of the variable to avoid possible errors of including cases of women at risk of depression when they actually are not. The SRQ20's validity and reliability has been tested in developing countries including some in Latin America such as Nicaragua, Colombia, Ecuador, Brazil and Chile (Harpham et al., 2005; Tuan, Harpham, & Huong, 2004; WHO, 1994), however it has not been validated for Peru. For the Peruvian case, one may contend that less educated women and women whose native language is Quechua may have problems understanding some of the questions in the SRQ20. Therefore, the SRQ20 may confound ability to understand complex sentences with minor psychiatric disorders (see the estimation method section for the strategy adopted to test for this issue). Additionally we used the Kuder-Richardson Formula 20 (KR-20) to check for the internal SRQ20 consistency. A value of 0.83 suggests that the SQR20 is highly reliable within the study population.

Child development is measured by two outcome variables collected at different stages of childhood to take into account short and long term impacts of maternal mental health. The first variable is child growth, which is measured using height-for-age Z score (HAZ) when children were around 1 year of age. HAZ is commonly used as a proxy measure of early nutritional status (Wisniewski, 2009). The second variable is child cognitive development, measured by the Peabody Picture Vocabulary Test (PPVT) at age 5. The PPVT has been used in other studies showing a strong correlation with the severity of maternal depression (Kiernan & Huerta, 2008; Murray et al., 1996). In this paper we postulate that maternal mental health has direct impacts on nutrition at age 1 and on cognitive development at age 5, and that cognitive development at age 5 is also affected through the impact of maternal mental health on nutrition.

Maternal education is measured by a dichotomous variable based on years of schooling. We differentiated between mothers with 6 or fewer years of schooling (primary education or less) versus women with more than 6 years of schooling (post-primary education). This cutting point has been selected due to the high enrolment and completion of primary schooling but lower transition into post-primary education in Peru (EPDC, 2009), as well as the potential social benefits of primary school completion (Psacharopoulos & Patrinos, 2002). We also selected this threshold based on the empirical distribution of years of schooling in the YL data, whereby slightly over half of mothers in the sample have achieved more than 6 years of primary education (see Table 1).[i]

Important covariates associated with maternal mental health and child development are: mothers' perceptions on the size of the child at birth, whether the child has had any serious health issues, whether the child resulted from an unwanted pregnancy, whether the child had attended preschool by the age of 5, whether the mother was in a relationship or not, whether the mother had a disability, household wealth, the number of years living in the current dwelling, number of children, financial difficulties which affected household wealth and regional controls. Table 1 shows descriptive statistics for these variables.

Table 1. Summary statistics: variable name, description, survey round, mean and standard deviation (s.d.)

Variable	Variables description	Round	Mean	s.d.
Mental Health (mother)	Maternal mental health status (SRQ20 score 0-20)	R1	5.60	4.15
Nutritional status (child)	Child height-for-age z-score	R1	-0.75	1.29
PPVT score (child)	Child PPVT score (scale from 0 to 95)	R2	29.05	17.69
Age (child)	Child's age in months	R1	11.54	3.51
Size at birth (child)	Mother perception of child's birth size: normal to large	R1	0.73	0.45
Overall Health (child)	Child never had a serious illness	R1	0.69	0.46
Pregnancy (mother)	Mother had an unwanted pregnancy	R1	0.45	0.50
Pre-school (child)	Child had attended pre-school by age 5	R2	0.82	0.39
Marital status (mother)	Mother is not in union with partner	R1	0.13	0.34
Education (mother)	Mothers with more than 6 years of education	R1	0.56	0.50
Disability (mother)	Mothers with the presence of a disability	R1	0.06	0.23
Residential stability (household)	Number of years living in the current place of residence	R1	15.93	11.07
Wealth index (household)	Household wealth index (based on three components: housing quality, consumer durable, services. Value range between 0 and 1).	R1	0.47	0.23
Children (household)	Number of children in the household	R1	2.64	2.05
Bad event (household)	Shock that decreased household wealth since the mother was pregnant with child	R1	0.39	0.49
Place of residence (household)	Household lives in a rural area	R1	0.33	0.47

3.3 Estimation method

According to our proposed hypotheses, maternal mental health is modelled directly as a function of education together with other risk factors (test of hypothesis 1). Maternal education can have a direct impact on child nutritional status at age 1 and both direct and indirect impacts on cognitive development at age 5. Education can attenuate the relationship between maternal mental health and child development, and hence it is included as an important factor associated with maternal mental health and with both our defined outcomes for child development (test of hypothesis 2). Finally, we estimate the parameters of the path model separately for women of different educational levels. In doing so, we test whether maternal education changes the dynamics between maternal mental health and child development (test of hypothesis 3).

We construct a pathway model for the possible relations between maternal mental health, maternal education and child development, which takes into account the timing of these events (see Figure 1). We use structural equation models, maximum likelihood estimation, to analyse the relations between these key variables.

In order to develop a path model that better fits the data, preliminary independent linear regression models to predict maternal mental health and the two child development outcomes were conducted, and only covariates with a statistically significant relationship were included in the analysis. This was done to improve power analysis. In addition to socio-demographic factors associated with maternal depression (child health status, unwanted pregnancy, marital status, maternal health, number of siblings, bad event since pregnancy, place of residence), specific socio-demographic variables associated with child growth (wealth index, child age, place of residence, size at birth) and with child cognitive development (wealth index, child age, child pre-school attendance) have been included. Omission of these factors can strongly bias the estimation of parameters in the model. Covariance paths have been constrained to be zero when non-significant to improve power analysis. The final model includes an error term, e_i, for each of the key outcome variables (maternal mental health, child nutrition and child cognitive development).

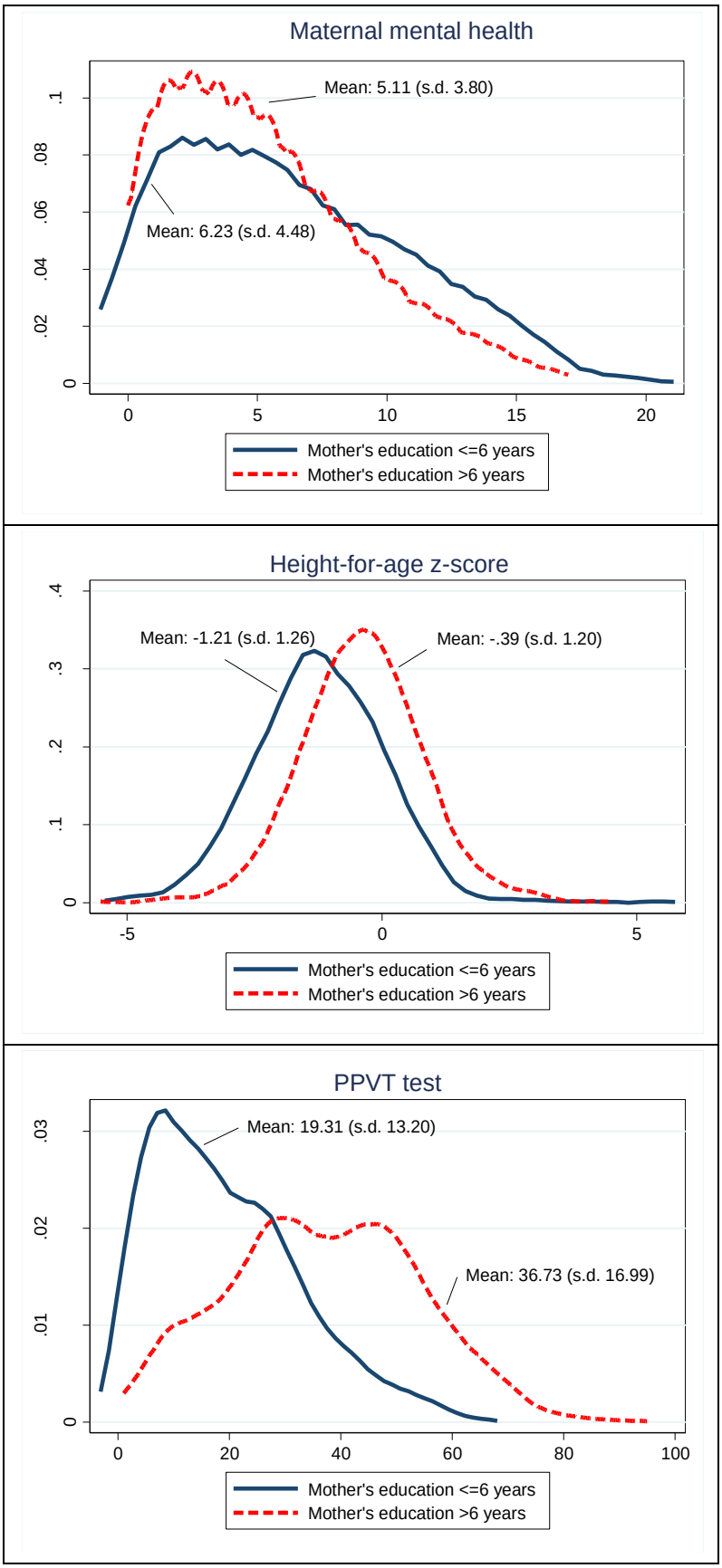
We use a multi-group path analysis to investigate the model described in Figure 1 by mother's level of education. The fit of all models is determined by the Chi-square, Goodness of Fit Index (GFI), Comparative Fit Index (CFI), and Root Mean Square Error of Approximation (RMSEA). The chi-square test measures the comparability between models. The GFI gives the proportion of variance in the variance-covariance matrix explained by the model, and when its value exceeds 0.9 it ensures a good fit. The CFI compares the model with the independence model, where all paths among variables are considered to be zero, and it should also exceed 0.9. Finally the RMSEA estimated the lack of fit comparing our model with a saturated model (which includes all possible correlations within and between exogenous and endogenous variables) and its value should not exceed 0.05.

For a sensitivity analysis, we estimated the same model on the subsample of women using Spanish as their first language, to check possible biases in the SRQ20 score due to women's ability to understand complex sentences. All the analyses were performed using SPSS AMOS.

4. Results

Figure 1 shows the Kernel density distributions for mother mental distress, child growth, and child cognitive development by mother's educational level. Mother level of education is a protective factor towards all the three indicators. Women with education below primary level achieve higher scores in the SRQ20 than women with levels of education above primary schooling (this result partially confirms hypothesis 1). The middle chart shows higher levels of HAZ among children whose mothers have more than primary education. Finally, we found that the distribution of PPVT scores, for children whose mothers have primary schooling or below, is skewed to the left in correspondence with lower PPVT scores.

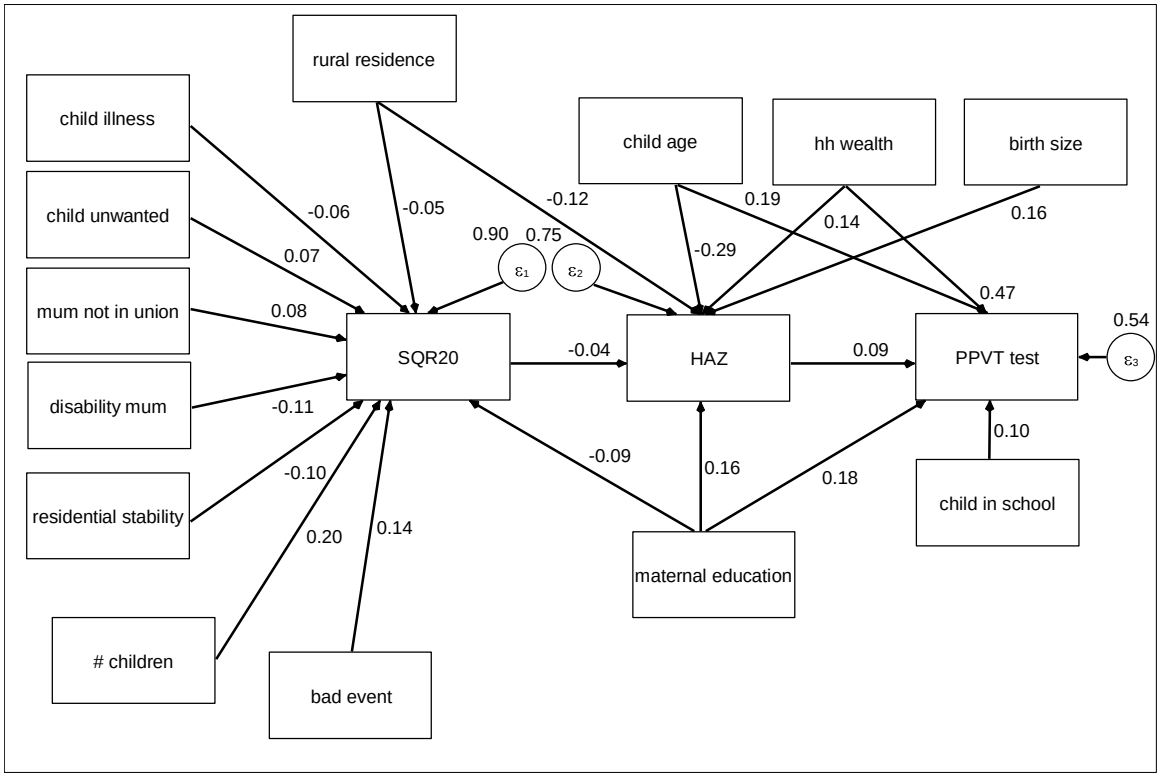
Figure 1. Kernell density function maternal mental health, HAZ, PPVT scores by mother's level of education



The path diagram where hypotheses 1 and 2 are tested is shown in Figure 2 (see Appendix Table A1, Model 1, for detailed results). Maternal education shows a direct and statistically significant association with mental health status (hypothesis 1). Maternal education is also associated with child nutritional status and child cognitive development, after controlling for the effect of other covariates.

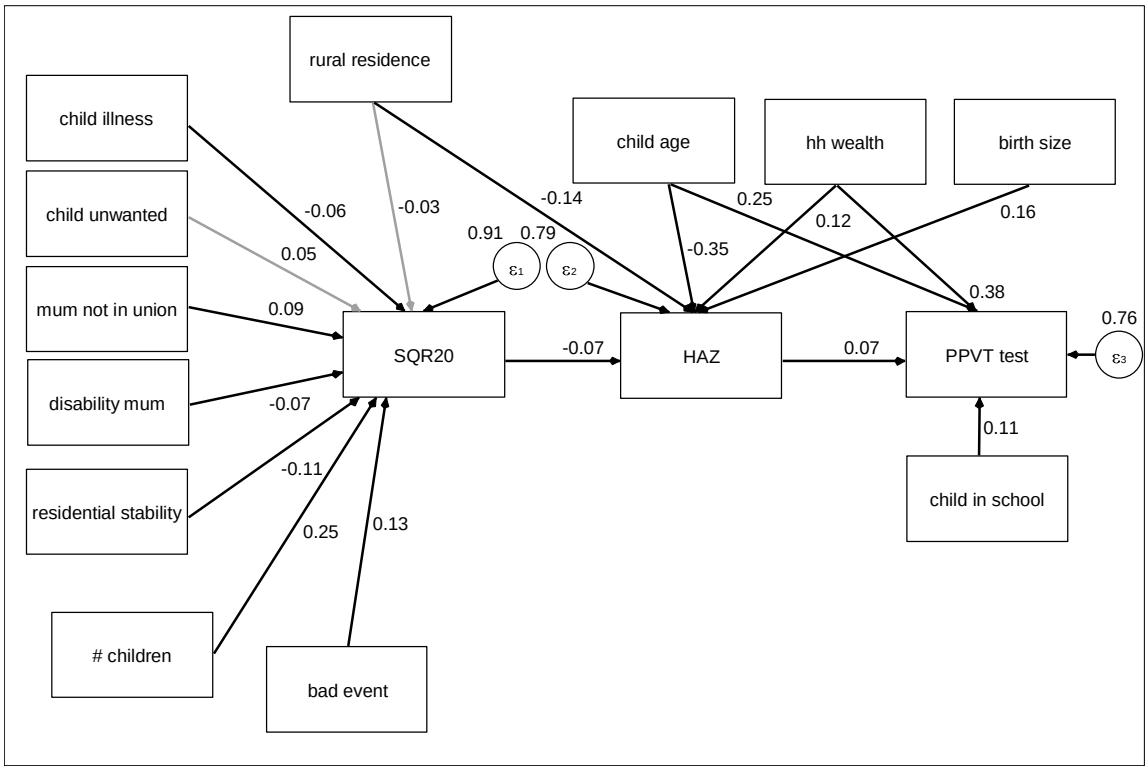
Secondly, although there was a prior significant correlation between maternal mental health and child development, the final path model does not show a direct significant effect from maternal mental health (SRQ20) to children's cognitive development (PPVT) at age 5. Nonetheless, maternal education is found to be significantly associated with both of these factors, and its inclusion reduced the statistical significance level between maternal mental health and child cognitive development. This result confirms the role of maternal education in weakening the relationship between maternal mental health and child cognitive development (hypothesis 2).

Figure 2. Path diagram Model 1: standardized path coefficients



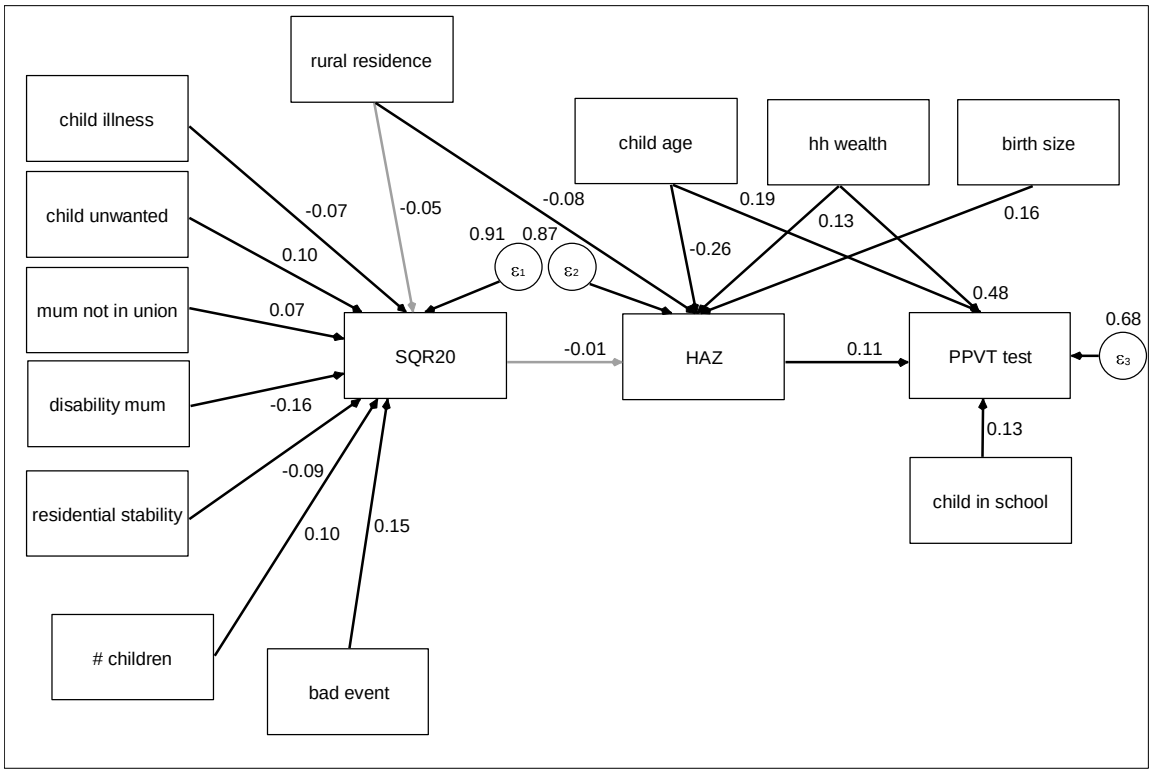
In order to test hypothesis 3, results from the path analysis were performed separately for women with 6 years or less schooling and for women with more than 6 years of schooling. Results are shown in Figure 3a and 3b (see Appendix Table A1, Model 2 and Model 3 for detailed results). We maintain identical models to let the non-significant paths emerge. Looking specifically at the pathway "*maternal mental status-child nutritional status-child cognitive development*", our results show that children of mothers with high SRQ20 scores tend to have lower nutritional status than children of mothers with low SRQ20 scores. However, this result is only supported for children of mothers with low education (st.est=-.067, p<.05) and not for children of mothers with high levels of education (st.est.=-.006, p>0.1). This result suggests that the nutritional status of the children tends to decrease by 0.067 standard units for each unit increase in their mothers' SRQ20 scores, only for children whose mothers have lower levels of education. Furthermore, the indirect effects of maternal mental health status on child cognitive development are statistically significant for the group of less educated mothers (st.est.=-.005, p<.05), but not for children whose mothers have high levels of education (st.est.=-.003, p>0.1).

Figure 3a. Path diagram Model 2: standardized path coefficients. Mother education<=6 years



Notes. (1) Black arrows indicate statistical significance <0.1, grey arrows indicate statistical significance >=0.1
(2) Residual and covariance between covariates are included in the model but not shown in the diagram.

Figure 3b. Path diagram Model 3: standardized path coefficients. Mother education>6 years



Notes. (1) Black arrows indicate statistical significance <0.1, grey arrows indicate statistical significance >=0.1
(2) Residual and covariance between covariates are included in the model but not shown in the diagram.

Table 2 shows the measures of model fit for the two independent groups, showing good fit for both subsamples. In addition, a multi-group model has been estimated to test for measurement invariance across the two groups. Our unconstrained model has been compared to a constrained model, where all the parameters (paths) between covariates and outcomes have been constrained to be equal between groups. Looking at the chi-square difference between unconstrained and constrained models, we can conclude if the same model applies across groups or not. The measure of fit for this final procedure is shown in Table 3. The unconstrained model has a slightly better fit than the constrained model, but more importantly, the chi-square difference test (chi-square=39.746, df=17) between the unconstrained and constrained model is significant (p<.001), suggesting that the hypothesis of invariance between the two groups is rejected. This therefore allows us to conclude that a difference exists in the path analysis between the two groups.

Table 2. Model fit statistics

	df	X2	GFI	CFI	RMSEA
Model 1	63	123.965	.992	.987	.023
Model 2	58	115.355	.982	.951	.035
Model 3	58	115.795	.985	.952	.031

Table 3. Model fit statistics for multi-group path analysis (mother's level of education)

	df	X2	GFI	CFI	RMSEA
Unconstrained	116	231.154	.984	.023	.951
Constrained	133	270.899	.981	.024	.942

There are other important results to highlight, which were obtained from the inclusion of covariates in the model (Table A1 in Appendix, Model 1). Focusing on *maternal mental status*, the strongest protective factor was residential stability (length of the time spent by the mother living in the same place). This dimension can be considered as a proxy of social capital, hypothesizing that the more the mother has been living in the same place, the more embedded she may be within the community, and the more protected by her own social network. The other contextual variable included in the model, i.e. rural place of residence at the time of the interview, has a direct positive effect on maternal mental status. As expected, child health affects mother mental wellbeing, with mothers of children who have never had a serious illness, being more likely to have a lower score in the SRQ20 test, hence better mental health. Among the strongest negative direct effects we detected was larger family size, having experienced a bad event, having some kind of disability (referred to the mother), being a single mother, and having had an unwanted pregnancy.

In addition to the association of maternal education with *child nutritional status* (which shows the strongest direct effect), age of the child, place of residence and maternal mental health status are associated with direct negative effects on child nutritional status. Living in a wealthier household and having been born into a normal to large size family is associated with higher HAZ.

Finally we find the expected association of child nutrition at age 1 on *child cognitive development* at age 5. We also find that children who attended pre-school, those who were relatively older, whose mothers had higher levels of education, and those living in richer households, were more likely to achieve higher scores in the PPVT at age five than the rest of the children. The sensitivity analysis shows that results remained unchanged when only Spanish-speakers mothers are included in the analysis.

5. Discussion

This paper set out to investigate whether maternal education was directly associated with reduced likelihood of maternal mental health problems and improved developmental outcomes for their children, and whether maternal education could reduce the burden of maternal mental health problems on developmental outcomes for children.

This last research question was hypothesised and investigated in terms of the potential moderating effect of maternal education.

Our results based on the Young Lives data for Peru confirmed previous empirical studies indicating the protective role of maternal education on maternal mental health (Chevalier & Feinstein, 2006; Robertson et al., 2003; Ross & Mirowsky, 1999) and child development (Black et al., 2007; Rahaman et al., 2002; Stewart, 2007). In particular we found that women with more education were less likely to self-report mental health problems. We also found that women with more education were likely to have better nourished children with higher levels of cognitive development than the children of mothers with less education.

In terms of the effects of maternal education in reducing the burden of maternal mental health on child development, we found a strong protective effect. The inclusion of maternal education remains one of the key factors predicting improved developmental outcomes for children (Feinstein et al., 2008). This is not the case for maternal mental health. This result is consistent with previous studies which have found weakening relationships between maternal mental health and children's outcomes upon the inclusion of maternal education as a key protective factor (Augustine & Crosnoe, 2010; Murray et al., 1996).

In terms of the moderating effect of education, our results show that the negative impact of maternal mental health problems on nutritional status and, indirectly on cognitive development, hold only for children whose mothers had low levels of education. Among children whose mothers had high levels of education, no significant direct or indirect effect of maternal mental health problems on child development were detected. These results are in line with Augustine and Crosnoe (2010) who found a negative relation between maternal mental illness and child cognitive development only for less educated women in the United States. It seems possible, therefore, that for women living in deprived areas of Peru, who managed to achieve levels of education beyond basic primary school, education plays a protective role not only for the general wellbeing of the mother, but also for the wellbeing of her child. However, less than half of women living in deprived areas of Peru failed to reach the basic level of primary education. These women lack the education to protect their children's physical and mental development against the effects of poor maternal mental health. This combination of disadvantage, as indicated by maternal mental health problems, economic deprivation and lack of maternal education, amongst other factors, is potentially responsible for the transmission of inequalities between generations (Exworthy, Blane & Marmot, 2003; Palloni, Milesi, White, & Turner, 2009).

With all its limitations in terms of availability of indicators and precision of measurements, this paper shows a clear role for maternal education, both as a protective factor and more importantly as a possible barrier to the transmission of inequalities between generations. If indeed women with low levels of education are those who are more likely to suffer from mental health problems, and more importantly, that their children are the ones more likely to lag behind nutritionally and academically, then the widening of inequalities between generations is likely to get even larger. In fact, these results are obtained for a relatively homogenous group in the Young Lives sample of Peru, so it is expected that for the whole population, inequalities in early life are likely to be even larger and the protective role of mother's education even more important. Therefore, this paper suggests that more analysis is needed in different communities so that particular interventions can be designed to help to tackle the transmission of these inequalities. We believe that monitoring of maternal health and child health, as well as provision of support for mothers with low levels of education, seem to be especially important if inequalities across future generations are to be reduced.

Acknowledgements

This research was funded by a grant from The BUPA Foundation (TBF-S09-034) entitled: "Access to education and child health." Research from this paper was presented at the Reproductive Morbidity and Poverty Workshop Announcement, 6th November 2010, London School of Economics, UK. Ethical approval was granted by the University of Sussex. The authors are grateful to the UK Data Archive and the Young Lives Team for their efforts in making the Young Lives Data available for research. We are especially thankful to Dr Rod Bond for his support and comments on the estimation of SEM models.

References

Augustine, J.M., & Crosnoe, R. (2010). Mothers' depression and educational attainment and their children's academic trajectories. *Journal of Health and Social Behavior, 51*, 274 –290. doi: 10.1177/0022146510377757

Boyden, J. (2006). Young Lives Project: conceptual framework. Project document http://www.younglives.org.uk/publications/project-documents

Black, M., Baqui, A.H., Zaman K., McNary, S.W., Le, K., Arifeen, SE., Hamadani, J.D., Parveeb, M., Yunus, M., & Black, R.E. (2007). Depressive symptoms among rural Bangladeshi mothers: Implications for infant development. *Journal of Child Psychology and Psychiatry, 48*, 764–772.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (pp. 1801-1859). Amsterdam: Elsevier-North Holland.

Carneiro, P., Meghir, C., & Parey, M. (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, *11*(s1), 123–160. doi: 10.1111/j.1542-4774.2012.01096.x

Chevalier, A., & Feinstein, L. (2006). Sheepskin or Prozac: The Causal Effect of Education on Mental Health. *IZA Discussion Paper No. 2231*. Available at SSRN: http://ssrn.com/abstract=923530

Cogill, S., Caplan, H., Alexandra, H., Robson, K., & Kumar, R. (1986). Impact of postnatal depression on cognitive development in young children. *British Medical Journal*, 292, 1165–1167.

CSDH. (2008). Closing the gap in a generation: health equity through action on the social determinants of health. World Health Organization, Geneva.

Duncan, G., & Brooks-Gunn, J. (2000). Family poverty, welfare reform, and child development. *Child Development*, *71*, 188–196. doi: 10.1111/1467-8624.00133

Edhborg, M., Lundh, W., Seimyr, L., & Widstroem, A. (2001). The long-term impact of postnatal depressed mood on mother-child interaction: A preliminary study. *Journal of Reproductive & Infant Psychology*, *19*, 61-71. doi: 10.1080/02646830123255

Education Policy and Data Center, EPDC. (2009). Pupil performance and age: A study of promotion, repetition and drop out rates among pupils in four age groups in 35 developing countries. Working Paper EDPC-09-02. Washington DC: Education Policy and Data Center.

Exworthy, M., Blane, D., & Marmot, M. (2003). Tackling health inequalities in the United Kingdom: the progress and pitfalls of policy. *Health Services Research, 38*, 1905-1922, Part II. doi: 10.1111/j.1475-6773.2003.00208.x

Feinstein, L., Duckworth, K. & Sabates, R. (2008). Education and the Family: Passing Success Across Generations. In Peter Aggleton & Sally Power (Eds.), *Foundations and Futures of Education* (pp. 1-222). Oxford: Routledge.

Field, T., Healy, B., Goldstein, S., Perry, S., Bendell, D., Schanberg, S., Zimmerman, E.A., & Kuhn, C. (1988). Infants of depressed mothers show depressed behaviours even with non-depressed adults. *Child Development*, *59*, 1569-1579.

Galler, J., Harrison, R., Ramsey, F., Forde, V., & Butler, S.C. (2000). Maternal depressive symptoms affect infant cognitive development in Barbados. *Journal of Child Psychology and Psychiatry, 41*, 747–757.

Grace, S.L., Evindar, A., & Stewart, D.E. (2003). The effect of postpartum depression on child cognitive development and behavior: a review and critical analysis of the literature. *Archives of Women's Mental Health, 6*, 263-274.

Harpham, T., Reichenheim, M., Oser, R., Thomas, E., Hamid, N., Jaswal, S., Ludermir, A., & Aidoo, M. (2003). Measuring mental health in a cost-effective manner. *Health Policy and Planning, 18*, 344-349. doi: 10.1093/heapol/czg041

Harpham, T., Huttly, S., De Silva, M.J., & Abramsky, T. (2005). Maternal mental health and child nutritional status in four developing countries. *Journal of Epidemiology and Community Health, 59*, 1060-1064. doi:10.1136/jech.2005.039180

Kerr, K. & West, M. (2010). *BERA insight: schools and social inequality*. London: British Educational Research Association.

Kiernan, K., & Huerta, C. (2008). Economic deprivation, maternal depression, parenting and children's cognitive and emotional development in early childhood. *The British Journal of Sociology, 59*, 783-806. doi: 10.1111/j.1468-4446.2008.00219.x

Knitzer, J., Theberge, S., & Johnson, K. (2008). *Reducing Maternal Depression and Its Impact on Young Children: Toward a Responsive Early Childhood Policy Framework*. Issue Brief 2. New York: National Center for Children in Poverty.

Lyons-Ruth, K., Zoll, D., Connell, D. & Grunebaum, H. U. (1986). The depressed mother and her one-year-old infant: Environment, interaction, attachment and infant development. In E. Z. Tronick & T. Field (Eds.), *Maternal Depression and Infant Disturbance. New Directions for Child Development*, 34, (pp. 61–82). Jossey-Bass : San Francisco.

Malik, S., & Courtney, K. (2011). Higher education and women's empowerment in Pakistan. *Gender and Education*, *23*, 29-45. doi: 10.1080/09540251003674071

Miech, R.A., & Shanahan, M.J. (2000). Socioeconomic status and depression over the life course. *Health and Social Behavior, 41*, 162-176.
Mirowsky, J., & Ross, C. (2002). Depression, parenthood, and age at first birth. *Social Science & Medicine, 54*, 1281-1298. doi: /10.1016/S0277-9536(01)00096-X
Murray, L. (1992). The impact of postnatal depression on infant development. *Journal of Child Psychology and Psychiatry, 33*, 543-561. doi: 10.1111/j.1469-7610.1992.tb00890.x
Murray, L., Kempton, C., Woolgar, M., & Hooper, R. (1993). Depressed mothers' speech to their infants and its relation to infant gender and cognitive development. *Journal of Child Psychology and Psychiatry, 34*, 1083-1101.
Murray, L., Hipwell, A., Hooper, R., Stein, A. & Cooper, P. (1996). The cognitive development of 5-year-old children of postnatally depressed mothers. *The Journal of Child Psychology and Psychiatry, 37*, 927-935.
Murray, L., & Cooper, P.J. (1997). Postpartum depression and child development (Editorial). *Psychological Medicine*, *27*, 253-260.
Outes-Leon, I., & Dercon, S. (2008). Survey attrition and attrition bias in Young Lives. *Young Lives Technical Note*, *5*.
Palloni, A., Milesi, C., White, R.G., & Turner, A. (2009). Early childhood health, reproduction of economic inequalities and the persistence of health and mortality differentials. *Social Science and Medicine, 68*, 1574-1582. doi: 10.1016/j.socscimed.2009.02.009
Patel, V., DeSouza, N., & Rodrigues, M. (2003). Postnatal depression and infant growth and development in low-income countries: A cohort study from Goa, India. *Archives of Disease in Childhood*, *88*, 34–37. doi: 10.1136/adc.88.1.34
Patel, V., Rahman, A., Jacob, K.S., & Hughes, M. (2004). Effect of maternal mental health on infant growth in low income countries: new evidence from South Asia. *British Medical Journal*, *328*, 3, 820-823. doi: 10.1136/bmj.328.7443.820
Paxson, C., & Schady, N. (2005). Child health and economic crisis in Peru. *World Bank Economic Review, 19*, 203-223. doi: 10.1093/wber/lhi011
Petterson, S.M., & Albers, A.B. (2001). Effects of Poverty and Maternal Depression on Early Child Development. *Child Development, 72*, 1794–1813.
Psacharopoulos, G., & Patrinos, H. (2002). Returns to investment in education: a further update. *World Bank Policy Research Working Paper*, No. 2881. Washington DC: The World Bank.
Rahaman, A., Harrington, R., & Bunn, J. (2002). Can maternal depression increase infant risk of illness and growth impairment in developing countries? *Child: Care, Health & Development*, *28*, 51–56. doi:10.1136/bmjopen-2011-000523
Rahaman, A., Iqbal, Z., Bunn, J., Lovel, H., & Harrington, R. (2004). Impact of maternal depression on infant nutritional status and illness. *Archives of General Psychiatry, 61*, 946-952. doi:10.1001/archpsyc.61.9.946
Robertson, E., Celasun, N., & Stewart, D.E. (2003). Risk factors for postpartum depression. In D.E. Stewart, E. Robertson, C.L. Dennis, S.L. Grace, S.L., & T. Wallington (Eds.), *Postpartum depression: Literature review of risk factors and interventions*. Available at: *www.who.int/mental_health/.../lit_review_postpartum_depression.pdf*
Ross, C. & Mirowsky, J. (1999). Refining the association between education and health: The effects of quantity, credential, and selectivity. *Demography, 36*, 445-60.
Sanchez, A. (2009). Early nutrition and later cognitive achievement in developing countries. Background paper for the Education for All global monitoring report 2010: *Reaching the marginalized.*Paris: UNESCO.
Stewart, R.C. (2007). Maternal depression and infant growth – a review of recent evidence. *Maternal and Child Nutrition, 3*, 94-107.
Teti, D.M., Gelfand, D.M., Messinger, D.S., & Isabella, R. (1995). Maternal depression and the quality of early attachment: An examination of infants, preschoolers, and their mothers. *Developmental Psychology, 31*, 364–376.
Tuan, T., Harpham, T., & Huong, N.T. (2004). Validity and reliability of the self-reporting questionnaire 20 items in Vietnam. *Hong Kong Journal of Psychiatry, 14*, 15-18.
Wang, L., Wu, T., Anderson, J.L., & Florence, J.E. (2011). Prevalence and risk factors of maternal depression during the first three years of child rearing. *Journal of Women's Health*, *20*, 711-718. doi:10.1089/jwh.2010.2232
Wilson, I., Huttly, S.R.A., & Fenn, B. (2006). A case study of sample design for longitudinal research: Young Lives. *International Journal of Social Research Methodology*, *9*, 351-365.
Wisniewski, S.L.W. (2009). Child nutrition, health problems, and school achievement in Sri Lanka. *World Development*, 38, 315-332. doi:10.1016/j.worlddev.2009.09.009.
WHO (1994). *A user's guide to the self reporting questionnaire (SRQ)*. Division of Mental Health, Geneva: WHO. www.whqlibdoc.who.int/hq/1994/WHO_MNH_PSF_94.8.pdf
WHO (2010a). *Mental health and development: targeting people with mental health conditions as a vulnerable group*. Geneva: WHO. http://www.who.int/mental_health/policy/mhtargeting/en/
WHO (2010b). *Gender and women's mental health. Gender disparities and mental health: The facts*. Geneva: WHO. http://www.who.int/mental_health/prevention/genderwomen/en/

Appendix Table A1. Estimations: standardized parameters (st.est.) with standard error (s.e.), and p-value

	Model 1			Model 2			Model 3		
				Mother's education <=6 years			Mother's education>6 years		
	st. est.	s.e.	p-value	st. est.	s.e.	p-value	st. est.	s.e.	p
Maternal mental health									
Place residence: Rural	-.048	.025	0.06	-.035	.034	0.31	-.049	.030	0.11
Child health: Ill	-.062	.022	<.01	-.058	.034	0.09	-.071	.030	0.02
Unwanted pregnancy	.070	.022	<.01	.046	.035	0.19	.101	.030	<.01
Mother not in union	.080	.023	<.01	.087	.035	0.01	.073	.031	0.02
Mother disable	.107	.022	<.01	.066	.034	0.06	.157	.030	<.01
Residential stability	-.099	.023	<.01	-.109	.037	<.01	-.092	.030	<.01
Number of Children	.196	.025	<.01	.247	.037	<.01	.097	.031	<.01
Experienced a bad event	.141	.022	<.01	.131	.034	<.01	.154	.030	<.01
Mother education	-.092	.026	<.01						
Child nutritional status									
Child age (months)	-.287	.019	<.01	-.354	.029	<.01	-.257	.028	<.01
Place residence: Rural	-.116	.026	<.01	-.137	.036	<.01	-.075	.033	0.02
Wealth Index	.143	.028	<.01	.123	.036	<.01	.133	.033	<.01
Birth size: Normal/Big	.155	.020	<.01	.163	.031	<.01	.162	.029	<.01
Mother education	.162	.025	<.01						
Mother mental health	-.037	.022	0.07	-.067	.033	0.03	-.006	.029	0.84
PPVT									
HAZ	.089	.019	<.01	.070	.033	0.04	.114	.027	<.01
Child age (months)	.188	.018	<.01	.246	.032	<.01	.193	.027	<.01
In pre-school	.096	.018	<.01	.110	.032	<.01	.126	.026	<.01
Mother education	.181	.021	<.01						
Wealth Index	.469	.019	<.01	.381	.029	<.01	.476	.022	<.01

Notes. Covariance between exogenous variables are included but not shown

Endnote

[i] A threshold effect may occur higher than grade six, but this issue is only relevant if there are no differences with respect to the association between maternal education and maternal mental health, as well as between maternal education and children's nutritional and cognitive outcomes, which is not the case as it is demonstrated in the paper.

Childhood friendships and the clustering of adverse circumstances in adulthood - a longitudinal study of a Stockholm cohort

Ylva B Almquist
Centre for Health Equity Studies, Stockholm University/Karolinska Institutet
yba@chess.su.se
Lars Brännström
Swedish Institute for Social Research, Stockholm University

(Received June 2012 Revised June 2013) doi: 10.14301/llcs.2013.3.2

Abstract

Friendships constitute a central feature of childhood, yet little is known about the developmental significance extending beyond childhood and adolescence. Based on longitudinal data for more than 14,000 individuals born in Stockholm in 1953, this study investigates the association between childhood friendships and groupings of adult outcomes, here conceptualised as conditions related to education, employment, economic hardship and health. Children's outcome profiles as adults were identified by means of latent class analysis. Multinomial logistic regression analysis was used to analyse whether childhood friendships predicted adult outcome profiles net of observed family background characteristics. The results indicated that children who lacked leisure time friends and a best friend in the school class, had increased risks of ending up in the more adverse clusters as adults, whereas the opposite association was found for those who reported being solitary. The effect of childhood friendships was rather consistent across both single and multiple problems, suggesting that the disadvantages of being without friends in childhood may not accumulate over the life course to any large extent. Generally, the findings were the same for males and females. The results were not explained by various family-related circumstances during upbringing. It is concluded that childhood friendships are important for adverse circumstances in adulthood, for both genders. As far as the long-lasting effects of children's friendships involve varying access to social support, school-based interventions should compensate for the scarcity of support following the lack of childhood friends.

Keywords: Childhood, friendship, living conditions, life course, cohort, latent class analysis

Introduction

It is widely acknowledged that friendships constitute a central feature of childhood, closely related to children's development. Taking their starting-point in a life-course perspective, scholars have argued that research into children's friendships needs to highlight the developmental significance extending beyond childhood and adolescence (Almquist, 2011a; Hartup & Stevens, 1997, 1999). The aim of the present study is therefore to examine the links between childhood friendships and adult outcomes. However, rather than focusing on various adult outcomes analysed in isolation, the present study adopts a person-oriented approach where the emphasis is put on the pattern of outcomes as a whole.

The importance of childhood friendships

While the interest in children's friendships has grown rapidly over the last decade or two (Berndt, 2004), the measurement issues surrounding this field of research remain difficult to solve. Most

would however, agree to the definition of friendship as a bilateral construct, since it emerges from the interaction between two individuals who have reciprocated positive feelings for one another (Bukowski & Hoza, 1989). For this reason, researchers have typically focused on pairs, or dyads, of children as the unit of analysis. These friendship dyads have commonly been identified through the use of peer nomination procedures, where children are asked to select their best friends or the classmates they like the most (Erdley, Nangle, Newman, & Carpenter, 2001). Assessments of friendships have, however, also been based on more lenient criteria, such as self-reports of the number of friendships (Gifford-Smith & Brownell, 2003; van der Horst & Coffé, 2012). Although the significance of quantitative aspects of friendship has been stressed (i.e. having friends), more qualitative measures (e.g. identity of friends and friendship quality) are being studied to an increasing extent (Hartup & Stevens, 1997). In the present study, three indicators of childhood friendships will be examined: hanging out with friends during leisure time, experiencing solitude, and having a best friend in the school class. While the two former are self-assessed, the latter is based on peer nominations. These three indicators are assumed to capture different, but perhaps similarly important, dimensions of friendship. For example, best friend reflects the situation in the school class; a context which is known to constitute one of the most significant arenas for peer interaction (Cairns & Cairns, 1994; Hartup, 1984). It also captures the deep structure of friendship, i.e. reciprocity (Hartup & Stevens, 1997). Hanging out with schoolmates outside school is a measure of companionship (Heaney & Israel, 2008) and could provide important opportunities to maintain and 'nurture' friendships obtained at school (Flora & Segrin, 1998). Finally, solitude is likely to reflect a more qualitative aspect of friendlessness, namely the subjective feeling of being without a close friend (Inderbitzen-Pisaruk, Clark, & Solano, 1992).

The scientific literature has identified a number of important provisions that are obtained through friendships (Furman & Robbins, 1985). For example, children's friendships provide affection (experiencing strong and positive ties); intimacy (sharing secrets and personal aspects); and a sense of reliable alliance (being loyal and sharing resources). Friendships also reflect inclusion (feeling a sense of belonging and acceptance). Other provisions are instrumental aid (getting help if needed); nurturance (feeling competent and needed); enhancement of worth (feeling proud and valued); and companionship (sharing activities). Considering this, it is not surprising that previous research has found friendships to be strongly correlated to a variety of indicators of adjustment (e.g. Zettergren, Bergman, & Wångby, 2006). For example, children with friends have been found to be more socially out-going and pro-social; have higher self-worth and self-esteem; and manage difficult transitions more smoothly (Bagwell, Newcomb, & Bukowski, 1998; Berndt, 1996; Berndt, Hawkins, & Jiao, 1999; Gest, Graham-Bermann, & Hartup, 2001). Children without friends, on the other hand, have been found to be more shy, timid, withdrawn, sensitive and unsuccessful at solving conflicts (Parker & Seal, 1996; Rubin, Bukowski, & Parker, 1998; Shantz & Shantz, 1985).

There are some longitudinal studies that have examined various aspects of friendship as predictors of subsequent developmental outcomes. However, most of these studies have had short follow-up periods and included rather small sample sizes. The results indicate, for example, that the lack of a (best) friend predicts depression (Erdley et al., 2001); suicidal thoughts (Bearman & Moody, 2004); internalising and externalising problems (Ladd & Troop-Gordon, 2003); emotional distress (Wentzel, Caldwell, & Barry, 2004); and academic achievement (Wentzel & Caldwell, 1997). Since relevant longitudinal data materials with a lengthy follow up are rare, studies focusing on whether the effects of children's friendships could last into young adulthood and beyond are not as numerous. Some do nevertheless exist: one study by Bagwell and colleagues (1998) indicated that those who were friendless in 5th grade were more likely to have psychopathological symptoms at age 23. This pattern was also demonstrated when these individuals were additionally followed-up at age 28 (Bagwell, Schmidt, Newcomb, & Bukowski, 2001). Another study (Almquist, 2011a) showed that having fewer friendships in childhood (age 8-12) was linked to an increased risk of poor self-reported health in mid-life (age 45-52).

In sum, the overall conclusion that may be drawn from research into children's friendships is that children who have less favourable experiences with friends also have increased risks of

developmental problems and, moreover, these risks do not appear to be confined to childhood and adolescence.

The clustering of adverse circumstances in adulthood

Given the fact that certain aspects of children's friendships, such as being without a friend, have shown persistent effects across a wide range of developmental outcomes in childhood and adolescence, it is reasonable to expect a similar consistency in the associations with adult outcomes. Moreover, it is well-known that difficulties in adulthood may go hand in hand (Bask, 2011; Bäckman & Nilsson, 2011; Fritzell, Gähler, & Nermo, 2007; Fritzell & Lundberg, 2000). It could thus be the case that a small group of individuals have a general susceptibility toward a wide range of adversities and thereby make up for a large proportion of the associations across outcomes. Consequently, it may be misguided to study a catalogue of adult outcomes in isolation.

One way of dealing with this issue is to instead examine the clustering of outcomes. Here, some parallels may be drawn to the person-oriented approach which originates from the field of developmental psychology (Bergman & Trost, 2006; Eye & Bogat, 2006). This approach, among other things, puts emphasis on the fact that it is the outcome pattern as a whole that carries the information rather than the parts regarded separately (Bergman, Magnusson, & El-Khouri, 2003; Bergman & Trost, 2006). Through the use of cluster analytical methods, it is possible to create outcome profiles where individuals are being categorised according to how similar or dissimilar they are to one another: for some individuals, one type of problem in adulthood may dominate whereas specific combinations or accumulation of problems may apply to other groups. In the present study, the aim is thus to investigate the potential influences of childhood friendships on the grouping of various outcomes in adult life. The outcomes are chosen to reflect a broad picture of adverse circumstances in adulthood and, as such, they include health-related, economic and social indicators. A similar strategy has been applied in another study using the same data material (Almquist & Brännström, 2012). Based on information about educational level, social assistance benefits, unemployment as well as mental and behavioural disorders, they distinguished four clusters of which one represented 'average' individuals, two reflected single problems, and the last comprised the most problem-burdened individuals.

Due to the scarcity of studies focusing on the long-lasting effects of childhood friendships, little is known about the pathways linking friendship to adult outcomes. Some inspiration may be drawn from research into the importance of family-related conditions as childhood precursors of future life chances, where it has been maintained that the individual's access to resources determines the level of opportunity at various stages across the life course. The lack of resources at one stage may result in limited resources at the next stage, thus bringing about a concatenation of disadvantages (Elder, 1998). As previously discussed, friendships involve important provisions. The lack of these provisions in childhood may impinge on circumstances (and hence the access to resources) at subsequent stages of life: first, through psychosocial mechanisms such as expectations, emotions, and ambitions. This is assumed to further influence the individual's educational choices, health-related behaviours, membership of networks, and coping strategies, which could ultimately have impacts on adult outcomes (cf. Almquist, 2011a; Almquist, 2011b; Östberg & Modin, 2007) in terms of e.g. educational career, income, labour market attachment and health.

Gender differences

Children tend to choose friends who are similar to themselves, particularly in terms of gender (McPherson, Smith-Lovin, & Cook, 2001; Shrum, Cheek, & Hunter, 1988). Gender is thus an underlying segregating principle for children's social interaction. A vast amount of research has found that females' friendships provide higher levels of closeness, affection, nurturance, trust and security (for a review, see Rose & Rudolph, 2006). It is therefore reasonable to assume that the lack of friendships is more disadvantageous for females' subsequent life chances. This has, however, not yet been fully confirmed in empirical studies.

Study aim

Based on a longitudinal cohort study of children born in Stockholm in 1953, the overall aim of the study is to examine the importance of childhood friendships for the clustering of adverse circumstances in adulthood. Indicators of childhood

friendship are: leisure time friends, solitude, and best friend. The adult circumstances, based on which various outcome profiles will be determined, include educational level, social assistance benefits, unemployment, and mental and behavioural disorders. Furthermore, various possible confounders reflecting social conditions related to the family will be examined (e.g. parental social class, income and education as well as family type, number of siblings and parental mental health problems). The role of gender for the studied associations will be investigated as well.

The following research questions will be investigated:

1) Are childhood friendships, in terms of a) leisure time friends, b) solitude, and c) best friend, associated with the clustering of adverse circumstances in adulthood?

2) Does the effect of childhood friendships vary across outcome profiles?

3) Is the association between childhood friendships and the clustering of adverse circumstances in adulthood confounded by social conditions during upbringing?

4) Are there any gender differences in any of the studied associations?

Based on previous research and theory, it is tentatively hypothesised that those who lack friends in childhood (in terms of each of the three friendship indicators) have increased risks of experiencing adverse circumstances in adulthood in general, and multiple difficulties in particular. Moreover, the strength of these associations is expected to be more pronounced among females compared to males.

Methods

The data material used is the Stockholm Birth Cohort Study (SBC), which was created in 2004/2005 by a probability matching of two longitudinal data sets: The Stockholm Metropolitan Study (SMS) and The Swedish Work and Mortality Data Base (WMD). The SMS cohort was initially defined as all children born in 1953 and living in the Stockholm metropolitan area in 1963 (n=15,117). All data was de-identified in 1986. In 2004–2005, the SMS was linked by a probability matching to the WMD, which is a temporary, population-based and anonymous database of Swedish residents who were born before 1985 and alive in 1980 and/or 1990. Of the original 15,117 individuals, approximately 95% (n=14,294) were positively matched and thereby included in the SBC (for a detailed description of the matching procedure, see Stenberg & Vågerö, 2005). Ethical permission for the SBC has been obtained from the Stockholm Regional Ethics Committee (No 739-03-629).

Variables

The present study uses four types of adult circumstances to calculate the outcome profiles: educational level social assistance benefits, unemployment, and mental and behavioural disorders. Concerning childhood friendships, three indicators were included: leisure time friends, solitude and best friend. Variables reflecting parental social class, parental income, parental education, family type, number of siblings, and parental mental health problems were included as controls (Table 1).

Information about educational level, social assistance benefits and unemployment in adulthood was derived from a Swedish registry database called Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA). To begin with, information about educational level in 2007 was based on the Swedish Educational Terminology (SUN), consisting of seven categories: pre-primary education; primary and lower secondary education, less than 9 years; primary and lower secondary education, 9 (or 10) years; upper secondary education; post-secondary education, less than two years; post-secondary education, two years or longer; and post-graduate education. In the present study, this variable was transformed into three categories: compulsory; upper secondary; and university. Secondly, based on the registered income from social assistance benefits during the period 1992-2007, dichotomised measures for each year were calculated (receiving benefits in 1992, yes or no; receiving benefits in 1993, yes or no; and so forth). The sum of years was subsequently recoded into the following categories: 0 years; 1-2 years; and 3 or more years. Thirdly, yearly information about the number of days in full-time unemployment was derived for the period 1992-2007. This variable was re-coded and collapsed in the same way as for social assistance benefits, thus indicating varying degrees of unemployment experience.

Table 1. Descriptive statistics for all study variables (n=10,694)

Variable	Year/time period	Distribution			
		Males (n=5,022)		Females (n=5,672)	
		n	%	n	%
Leisure time friends	1966				
Yes		4,850	96.6	5,481	96.6
No		172	3.4	191	3.4
Solitude	1966				
No		4,810	95.8	5,347	94.3
Yes		212	4.2	325	5.7
Best friend	1966				
Yes		4,625	92.1	4,953	87.3
No		397	7.9	719	12.7
Educational level	2007				
Compulsory		857	17.1	652	11.5
Upper secondary		2,200	43.8	2,525	44.5
University		1,965	39.1	2,495	44.0
Unemployment	1992-2007				
0 years		3,438	68.5	3,724	65.7
1-2 years		678	13.5	850	15.0
3 or more years		906	18.0	1,098	19.4
Social assistance benefits	1992-2007				
0 years		4,159	82.8	4,428	78.1
1-2 years		472	9.4	660	11.6
3 or more years		391	7.8	584	10.3
Mental and behavioural disorders	1992-2007				
0 admissions		4,740	94.4	5,408	95.4
1-2 admissions		156	3.1	153	2.7
3 or more admissions		126	2.5	111	2.0
Parental social class	1963				
Upper/upper middle		841	16.8	910	16.0
Non-manual, officials		1,775	35.3	1,997	35.2
Non-manual, entrepreneurs		378	7.5	465	8.2
Manual, skilled		1,121	22.3	1,277	22.5
Manual, unskilled		788	15.7	902	15.9
Other		119	2.4	121	2.1
Parental income (thousands of SEK)	1964	Mean=13.6		Mean=13.9	
Parental education	1960				
No graduate		3,520	70.1	4,001	70.5
One graduate		1,005	20.0	1,128	19.9
Two or more graduates		306	6.1	314	5.5
No information		191	3.8	229	4.0
Family type	1964				
Two-parent household		4,586	91.3	5,122	90.3
Other		436	8.7	549	9.7
Number of siblings	1964	Range=0-9		Range=0-8	
Parental mental health problems	1953-1965				
No		4,832	96.2	5,455	96.2
Yes		190	3.8	217	3.8

Finally, information about mental and behavioural disorders was derived from the Hospital discharge register. This measure was based on the number of discharges from Swedish hospitals (in-patient care) due to mental and behavioural disorders, during the period 1992-2007. Using the 9th Revision (for the period 1992–1996) and the 10th Revision (for the period 1997–2007) of the International Classification of Diseases (ICD), diagnoses in the ICD 9 chapter 'Mental disorders' (290-319) and the ICD 10 chapter 'Mental and behavioural disorders' (F00-F99) were coded as mental and behavioural disorders (including e.g. depression, anxiety and substance abuse). The variable was collapsed into three categories: 0 admissions; 1-2 admissions; and 3 or more admissions.

The three measures of childhood friendships – leisure time friends, solitude and best friend – were based on information from the School Study of 1966. All school classes in Stockholm were included in the School Study, with the exception of classes of children with learning disabilities. The measure of leisure time friends was derived from the question: "*How often do you spend your leisure time with some of your school mates?*" The response options were: 'Every day'; 'A few times a week'; 'Once a week'; 'Seldom'; and 'Never'. Individuals who had responded 'Never' were put into one category (3%), whereas the rest formed another category. The measure of solitude was based on the question: "*With whom do you spend most of your time?*" The response options were: 'Mostly with boys'; 'Mostly with girls'; 'Mostly with a group of boys and girls'; 'Mostly with adults'; and 'I am mostly on my own'. Those who answered 'I am mostly on my own' were categorised as being solitary (5%), whereas the remaining individuals were put into another category. Lastly, based on a peer nomination procedure, the measurement of best friend could be derived. All students in each school class were asked: "*Who are your three best friends in class?*" Individuals who did not receive any nominations from their classmates were assumed to lack a best friend (10%), whereas the rest were categorised as having at least one best friend. The choice of distinguishing relatively extreme categories was made, in order to more clearly discriminate individuals who could be considered to have friendship problems, from those who did not. The results from a correlation analysis showed positive but weak correlations between the three measures (leisure time friends and solitude: r=0.14; leisure time friends and best friend: r=0.05; and solitude and best friend: r=0.06).

Six types of family-related circumstances during upbringing were included as control variables in the analysis (for the distribution of these variables, see Table 1. Parental social class (1963) was based on pre-coded occupational data concerning the head of the household (in most cases the father). Parental income (1964) was measured through the mean of the combined earned income of both parents. Information about parental education (1960) was based on the total number of household members who graduated from secondary school or equivalent. Concerning family type (1964), two categories were distinguished: children who were recorded as living in a two-parent household (biological parents or reconstituted families) and those who lived in any other type of families (single-parent household, foster parents or widow/widower). Furthermore, a continuous measure of the number of siblings (1964) was included. With regard to parental mental health problems (1953-1965), it had originally been divided into three categories: psychiatric problems or depressed; receiving psychiatric treatment; and committed to an institution for care. All these categories were, in this study, used as indicators of parental mental health problems. The six indicators of family-related circumstances were weakly correlated, also with the indicators of friendship. Most correlations were lower than 0.20 and many below 0.10, with the following exceptions (ranging between 0.20 and 0.40): parental income and parental social class; parental education and parental social class; and family type and parental income.

Data analysis

Based on the information about circumstances in adulthood, 'outcome profiles' were generated to identify similarities between individuals in the present study. Here, latent class analysis was applied, using Mplus 6.1. The basic notion underlying latent class analysis is that there are unobserved relationships among observed variables (Hagenaars & McCutcheon, 2002). The number of clusters was determined through different indicators of model fit (see Table 2). For example, the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) can be used to assess the relative goodness of fit of different models, where lower values indicate better fit (Kuha, 2004). In the present study, BIC was lowest for the three-cluster solution, whereas the five-cluster solution had the lowest AIC. The four-cluster solution

had the second lowest values in both instances. A further analysis showed that the five-cluster model had better scores for entropy (reflecting the quality of the classification, where a value of 0 reflects randomness and a value of 1 corresponds to a perfect classification). Moreover, the models with four and five clusters respectively did not have problems with correlated bivariate residuals. The choice between the models was further based on a Vuong-Lo-Mendell-Rubin likelihood ratio test (Henson, Reise, & Kim, 2007) which showed that the four-cluster solution had significantly better model fit compared to the three-cluster solution, whereas the five-cluster solution did not fit the data better than the four-cluster solution. The model with four clusters was therefore preferred. Performing a gender-specific analysis yielded virtually the same four-cluster solution (data not presented).

Table 2. Latent class analysis: model fit statistics (n=10,694)

Number of classes/clusters	AIC	BIC	Entropy/Quality of classification	Uncorrelated bivariate residuals	Likelihood ratio test[a]
1	58428.371	58486.591	1.000	No	-
2	56408.556	56532.272	0.734	No	p<0.000
3	56287.673	56476.886	0.589	No	p<0.000
4[b]	56250.974	56505.684	0.631	Yes	p<0.000
5	56250.284	56570.491	0.658	Yes	p=0.304

Notes.
BIC = Bayesian Information Criterion
AIC=Akaike Information Criterion
[a] Vuong-Lo-Mendell-Rubin likelihood ratio test
[b] Preferred model

The next step was to analyse the association between childhood friendships and adverse circumstances in adulthood. Since the outcome was nominal (i.e. non-ordered response categories), multinomial regression analysis was used, producing relative risk ratios (RRR). The RRRs were obtained by exponentiating the multinomial logit coefficients. Three models were generated, of which the first demonstrates the effects of each friendship indicator separately, adjusted for gender. Here, the results from the interaction analysis were also included. Each interaction term was calculated by multiplying a friendship variable with the gender variable, of which both had a binary coding (zero and one). The new variable could thus also have the possible values of zero and one. The second model incorporates the effects of all three indicators of childhood friendships simultaneously, as well as gender. Finally, the third model controls for the various family-related circumstances (social class, income and education of parents, as well as family type, number of siblings and parental mental health problems). The results from the multinomial regression analysis are displayed in Table 3.

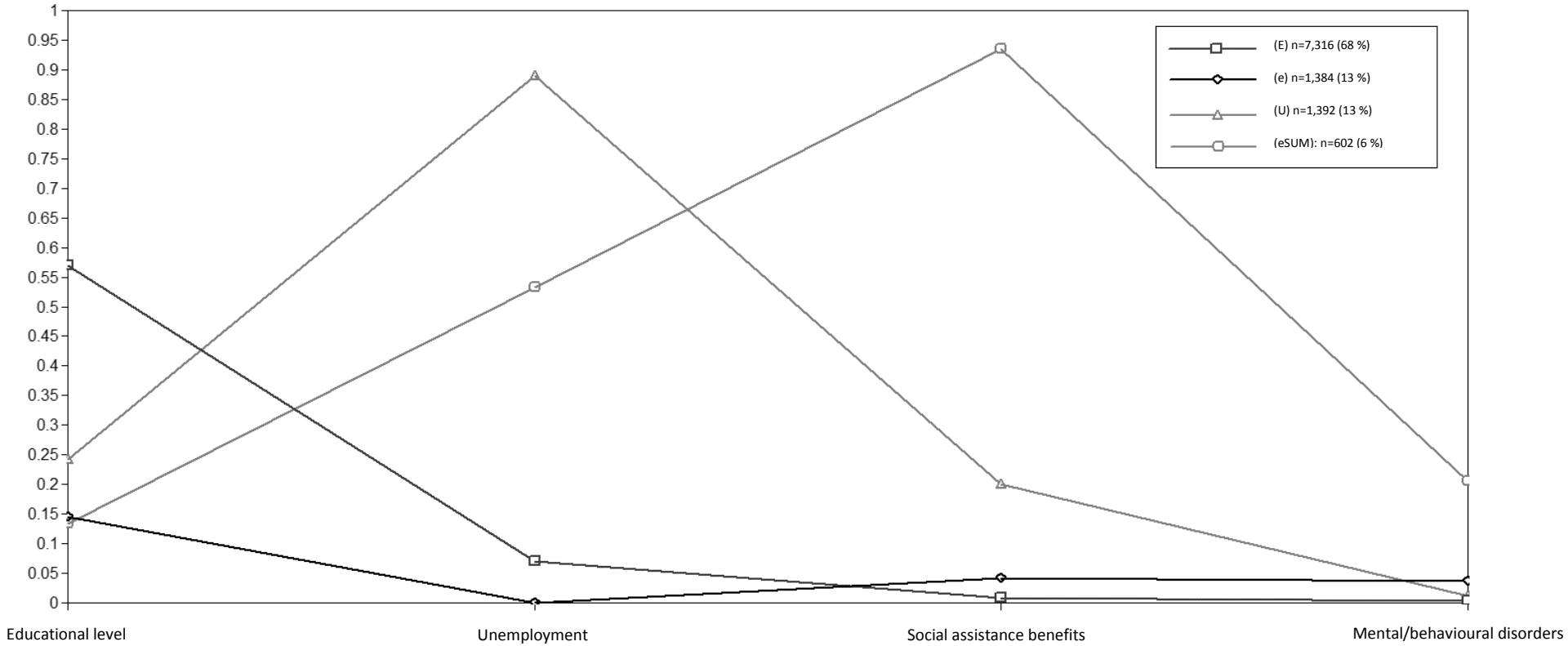
Study sample

The analyses performed in the present study are based on individuals with information about all study variables. Of the original SBC sample (n=14,294), 10,694 individuals were thereby included (to some extent positively selected in terms of childhood conditions, data not presented).

Results

In Figure 1, the four-cluster solution produced by the latent class analysis is demonstrated, using a plot of the estimated probabilities. The labeling of the clusters is based on the level of the probabilities. The first cluster contains 68% of the individuals and includes those with comparably More education (E). The second cluster is characterised by relatively Less education (e), containing 13% of the sample. In the third cluster, which contains 13% of the individuals, More unemployment (U) is evident. Finally, the fourth cluster, containing 6% of the sample, is more problem-burdened compared to the others, with comparably Less education, more social assistance benefits, more unemployment, and more mental and behavioural disorders (eSUM).

Figure 1. Description of the outcome profiles; estimated probabilities derived using Mplus, number of cases and relative cluster size within brackets (n=10,694)



Key

E: More education; e: Less education; U: More unemployment; eSUM: Less education, more social assistance benefits, more unemployment, and more mental and behavioural disorders

Table 3 presents the association between childhood friendships and adverse circumstances in adulthood. The gender-adjusted results from Model 1 indicate, firstly, that those who lack leisure time friends have a 37% higher risk of ending up in the cluster characterised by Less education (e) rather than in the More education cluster (E). These individuals also have an increased risk (68% and 89%, respectively) of being found in the cluster characterised by More unemployment (U) as well as in the most problem-burdened cluster (eSUM). In the lower part of the table, the results from the interaction analysis are presented. There are no statistically significant gender differences in the association between leisure time friends and adult circumstances. In Model 2, where the three indicators of childhood friendships are included simultaneously, the estimates remain largely unchanged. No statistically significant interaction effects between the indicators of friendship were found here (data not presented). Model 3 controls for the various family-related circumstances, reducing the risk: e.g. the risk for individuals without leisure time friends to end up in the Less education (e) cluster decreases from 1.37 to 1.26 and is no longer statistically significant.

When it comes to solitude, Model 1 shows that those who have reported that they are solitary have decreased risks of ending up in the clusters characterised by Less education (e), More unemployment (U) and More unemployment, more social assistance benefits, more mental and behavioural disorders, and less education (eSUM). Thus, contrary to what was expected, solitude shows a negative association with the clustering of adverse circumstances in adulthood. These results are however not statistically significant. The interaction analysis presented in the lower part of the table indicates that females who experience solitude in childhood have decreased risks to end up in the Less education (e) cluster and increased risks of ending up in the More unemployment (U) cluster compared to males. The estimates in the mutually adjusted column (Model 2) become slightly increased compared to the gender-adjusted results (Model 1). When the control variables are added to Model 3, the strength of the association between solitude and adult circumstances is marginally reduced.

Concerning the third indicator of childhood friendships, the gender-adjusted results (Model 1) suggest that those who lack a best friend have an increased risk of adverse circumstances in adulthood. For example, individuals who did not receive any nominations for best friend from their classmates have a 51% increased risk of later ending up in the cluster characterised by Less education (e), rather than in the More education cluster. Higher risks are also found for the two remaining clusters: 51% for the More unemployment cluster (U) and 85% for the most problem-burdened cluster (eSUM). No statistically significant interaction between best friend and gender is found. When the effects of all three friendship measures are included simultaneously (Model 2), only marginal changes occur. Model 3, which incorporates the effect of the control variables, shows that a limited part of the association between having a best friend and adult circumstances is accounted for by family-related circumstances.

Finally, although not displayed in Table 3, it should be noted that there were some ‘crude’ effects of gender on adverse circumstances in adulthood. More specifically, while no gender differences in the risk of ending up in the Less education (e) cluster were demonstrated, females had a 16 % higher risk of being found in the More unemployment cluster (U) and a 29 % higher risk to end up in the most problem-burdened cluster (eSUM). These results are also largely reflected in Model 2, although these estimates are adjusted for the three friendship indicators.

Table 3. The association between childhood friendships and the clustering of adverse circumstances in adulthood. Results from multinomial regression analysis (n=10,694)

	e			U			eSUM		
	Model 1[a]	Model 2[b]	Model 3[c]	Model 1[a]	Model 2[b]	Model 3[c]	Model 1[a]	Model 2[b]	Model 3[c]
	RRR (95 % CI)	RRR (95 % CI)	RRR (95 % CI)	RRR (95 % CI)	RRR (95 % CI)	RRR (95 % CI)	RRR (95 % CI)	RRR (95 % CI)	RRR (95 % CI)
Leisure time friends									
Yes (ref.)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
No	1.37 (1.01-1.86)	1.37 (1.01-1.87)	1.26 (0.92-1.72)	1.68 (1.27-2.23)	1.67 (1.25-2.23)	1.56 (1.16-2.08)	1.89 (1.29-2.78)	1.90 (1.29-2.80)	1.68 (1.13-2.49)
Solitude									
No (ref.)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Yes	0.85 (0.64-1.12)	0.79 (0.60-1.04)	0.82 (0.62-1.09)	0.92 (0.71-1.20)	0.84 (0.64-1.09)	0.86 (0.66-1.13)	0.81 (0.54-1.21)	0.70 (0.46-1.06)	0.74 (0.49-1.12)
Best friend									
Yes (ref.)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
No	1.51 (1.27-1.80)	1.51 (1.27-1.81)	1.46 (1.22-1.75)	1.51 (1.27-1.80)	1.50 (1.25-1.78)	1.45 (1.21-1.73)	1.85 (1.46-2.33)	1.83 (1.45-2.32)	1.74 (1.37-2.22)
Gender									
Males (ref.)		1.00	1.00		1.00	1.00		1.00	1.00
Females		1.07 (0.95-1.20)	1.06 (0.94-1.19)		1.14 (1.02-1.28)	1.13 (1.01-1.27)		1.25 (1.06-1.48)	1.15 (1.05-1.48)
Leisure time friends * gender	n.s.			n.s.			n.s.		
Solitude * gender	p<0.01			p<0.05			n.s.		
Best friend * gender	n.s.			n.s.			n.s.		

Notes.
More education (E) is the reference group.
[a] Adjusted for gender
[b] Adjusted for gender + mutual adjustment for friendship indicators
[c] Adjusted for gender, parental social class, parental education, parental income, family type, number of siblings, and parental mental health problems + mutual adjustment for friendship indicators
RRR=Relative risk ratio
E: More education; e: Less education; U: More unemployment; eSUM: Less education, more social assistance benefits, more unemployment, and more mental and behavioural disorders

Discussion

Childhood friendships as a predictor of adverse circumstances in adulthood

The aim of the present study was to examine the association between various indicators of childhood friendships and the clustering of adverse circumstances in adulthood. The results suggested that such links do exist: children who did not hang out with school friends during leisure time and did not had a best friend in their school class, had increased risks of adversities as adults. This is consistent with previous studies identifying harmful long-term effects among those who were without friends in childhood (Almquist, 2011a; Bagwell et al., 1998; Bagwell et al., 2001). Moreover, the association between childhood friendships and adult circumstances was not fully accounted for by any of the family-related circumstances included in the study (e.g. the socioeconomic status and mental health of the parents, as well as family type and number of siblings).

The explanations for the link between childhood friendships and the clustering of adversities in adulthood may reside in the functions of friendships, where social support previously has been identified as a key aspect (Heaney & Israel, 2008). Four types of social support, linked to various types of provisions, are commonly differentiated: emotional (love, trust and caring), instrumental (aid and services), informational (advice, suggestions and information), and appraisal (constructive feedback and affirmation). It is reasonable to assume that individuals who lack friends in childhood have a lower level of support, which in turn could make them less equipped to deal with various choices and events occurring across the life course (e.g. getting an education, entering the labour market, and adopting healthy behaviours).

While some have argued that social support is beneficial only for individuals under stress and, as such, 'buffers' the adverse effects of stressful events, others maintain that social support is linked to the individual's outcomes irrespective of whether a stressor is present (Cohen & Wills, 1985). The present study did not attempt to separate between buffering effects and main effects of childhood friendships. In other words, although a main effect of social support on adult outcomes was observed, this does not rule out childhood friendships as a potentially buffering factor.

The case of solitude

All indicators of childhood friendships did not display the same association to adult circumstances. On the contrary of what was expected, individuals who reported being solitary in childhood seemed to have lower risks of ending up in outcome profiles characterised by high levels of adverse outcomes. Thus, it appears as if solitude is protective against later problems. This should however be interpreted with the low correlations between the indicators of childhood friendships in mind: solitude seems to be composed primarily of aspects not related to hanging out with classmates during leisure time and having a best friend in the school class. Rather, it could be the case that children who report being solitary still have high levels of social support originating from other people (e.g. friends outside school, family and siblings) as well as other important resources, which in turn could counterbalance the negative effects of experiencing solitude. A sensitivity analysis (data not presented) showed, for example, that those who reported being solitary had higher average marks compared to those who did not experience solitude (whereas the opposite pattern was shown for the other two indicators of childhood friendships). Thus, this group of children may still have the abilities required to end up in more advantageous outcome profiles later in life.

Relative importance of friendship indicators

The results revealed independent effects of all three indicators of childhood friendships on adverse circumstances later in life. Again, this could be ascribed to the low correlations between these indicators. Having a best friend in the school class seems to involve a process that is different from hanging out with classmates during leisure time and also different from being solitary. For example, having a best friend is likely to reflect reciprocity, i.e. mutual liking and acceptance, to a greater extent. As such, it is reasonable to expect that emotional support is the primary type of provision gained by these individuals. Spending time with classmates outside of the school is, on the other hand, strongly related to the presence of companionship, which is another function commonly mentioned in friendship research. Companionship involves, for example, the sharing

of activities during leisure time (Heaney & Israel, 2008). This may imply joint membership in organisations or networks (e.g. youth or sports clubs), which in turn could further boost the provision of various types of social support. Lastly, solitude was initially assumed to reflect loneliness since it entailed being mostly on one's own. Loneliness has previously been described as a multi-dimensional phenomenon that encompasses the individual's experience of an unpleasant lack of relationships (de Jong Giervald, 1998). Young people are believed to be especially vulnerable to loneliness due to the increased importance of friendship during this period in life (Inderbitzen-Pisaruk et al., 1992). However, as previously mentioned, children who reported being solitary still enjoyed academic success, something which provides resources that are perhaps even more salient for adult living conditions than social support.

Accumulation of problems over the life course

Based on the indicators of adverse circumstances in adulthood, the latent class analysis revealed four outcome profiles: the first two reflected higher and lower education, respectively. The third was characterised by higher levels of unemployment whereas the fourth reflected adversities in all four arenas: less education, more social assistance benefits, more unemployment, and more mental and behavioural disorders. The corresponding cluster solution has been found in a previous study of children's status position in the peer group as a predictor of subsequent living conditions (Almquist & Brännström, 2012). Additionally, in that study, low peer status showed the strongest link to the most problem-burdened cluster, suggesting that adverse circumstances may accumulate over time (cf. DiPrete & Eirich, 2006). In the present study, contrary to what was initially expected, there were no stronger effects of childhood friendships on profiles characterised by multiple problems: with the cluster characterised by comparably higher levels of education as the reference, the risks were rather evenly distributed across the other profiles. This could indicate that the negative effects of being without friends in childhood may not accumulate over the life course to any large extent.

Gender differences

In the cohort studied here, males had slightly lower levels of unemployment and social assistance benefits as well as a higher number of hospitalisations due to mental and behavioural disorders, whereas females were somewhat more educated. However, the cluster analysis revealed the same outcome profiles for both genders (data not presented), which suggests that males and females are similar to one another when it comes to the 'mix' of adverse problems in adulthood. The regression analysis largely confirmed this, although females had a higher risk of ending up in the cluster characterised by higher levels of unemployment, as well as the most problem-burdened cluster compared to males. Concerning the association between childhood friendships and the clustering of adverse circumstances in adulthood, no gender differences were found, with the exception of solitude, where solitary females had a decreased risk of being found in the cluster characterised by less education, and an increased risk of ending up in the cluster with higher levels of unemployment. In sum, however, males and females appear to be more similar than different when it comes to the issues examined in the present study.

The use of a person-oriented approach

In the present study, a person-oriented approach was favoured over a variable-oriented approach. The person-oriented approach focuses on the pattern of outcomes as a whole, targeting the ways in which individuals are similar or dissimilar to one another in terms of e.g. adverse living conditions. A major advantage of applying this approach was that the researchers could empirically identify different combinations of adverse living conditions. Another advantage was that the person-oriented approach did not assume that linear associations exist between the variables or for all individuals. In order to highlight these important differences between the person-oriented approach and the variable-oriented approach, the analysis in the present study was also performed by using the latter strategy (Appendix 1). First, a summary measure of adverse living conditions was created (0 problems; 1 problem; 2 problems; 3-4 problems). At first glance, the distribution of this variable roughly corresponded to the percentages of individuals in the various clusters (Table A). However, a more scattered picture emerges when cross-comparing the two distributions (Table B). The greatest

consistency is shown for the More education (E) cluster as well as the most problem-burdened cluster (eSUM) which is mainly represented by individuals who also were identified as having multiple problems in the summary measure (2-4 problems). The remaining clusters correspond considerably less to the summary measure. The association between childhood friendships and adverse circumstances in adulthood (using the summary measure) was subsequently analysed using ordinal regression analysis (Table C). The results indicate that individuals who lack of friendships in childhood have more problems as adults. However, in comparison to the previous analyses, this approach masks the combinations of problems that exist in the data, and (more or less incorrectly) assumes that the effect of the predictor is linear on the outcome. The conclusion drawn here is that the person-oriented approach, through its emphasis on outcome patterns, may extend the knowledge previously established by more traditional research.

Strengths and limitations

With rich information on childhood friendships and adult circumstances, the Stockholm Birth Cohort study provided an excellent opportunity to investigate these matters in detail. There are, however, some limitations that need to be recognised. First of all, the indicators of childhood friendships were predominantly school-based; the measure of best friend referred to classmates, whereas the question regarding leisure time friends referred to schoolmates. Given that the question about solitude was posed in a battery of school-related items, it is possible that the respondents gave their answers as primarily based on conditions in school. Since the school constitutes a central arena for the establishment and maintenance of friendships, it is fully reasonable to focus on school-based friendships. It should nevertheless be acknowledged that children may also have supportive friendships outside of school. Furthermore, the three chosen indicators were mostly concerned with the quantitative aspects of friendships. Previous research has highlighted the importance of also taking into account the quality of friendships as well as who the friends are (Hartup & Stevens, 1997). For example, having a friend may not be as beneficial for subsequent outcomes if the relationship is characterised by conflict and distrust, or if the friend has adopted adverse behaviours. These dark sides of friendships could offset the positive effects to some extent. A second issue has to do with causality. The present study controlled for several factors reflecting social conditions during childhood (i.e. parental social class, income, education and mental health problems as well as family type and number of siblings). This did not lead to any substantial changes in the results. However, other circumstances, which could not be taken into account here, may be equally or even more important, such as the child's own health status. Moreover, no mechanisms were investigated. It is plausible that factors such as the individual's educational choices and entrance into the labour market may be of relevance (Bäckman & Nilsson, 2011).

Finally, it should be pointed out that the results of the present study are based on a cohort of Stockholm children born in 1953. The clustering of living conditions should be viewed in the light of its context: Sweden is generally considered a generous welfare state with a well-developed 'safety net'. Hence, it could be argued that the very existence of clustering of adult adversity, diverges from the idea of a welfare state which aims to compensate for the unequal distribution of opportunity across groups of people. Regardless of the content and structure of the clusters, however, it is reasonable to expect that the association between childhood friendships and adult outcomes would look similar in other geographical and cultural contexts. More research is needed to verify this.

Concluding remarks

Childhood friendships matter for the risk of adverse circumstances in adulthood, among males and females alike. It is suggested that the provisions gained through childhood friendships have long-lasting effects on individual outcomes. Friendship is based on mutual liking and, therefore, it would be impracticable to try to enforce such relationships among children. Rather, school-based intervention programmes should aim at finding ways to compensate for the lack of social support among those who are without friends in childhood.

Acknowledgements

The creation and maintenance of the Stockholm Birth Cohort Data Base is a collaboration between the Centre for Health Equity Studies (CHESS) and the Swedish Institute for Social Research (SOFI), financed by the Swedish Research Council. Sten-Åke Stenberg of SOFI prepared the original Metropolitan Data Base, Denny Vågerö of CHESS prepared the follow-up data, and Reidar Österman of CHESS organised the probability matching of the two data sets. The present study was financially supported by the Swedish Research Council for Health, Working Life and Welfare (grant number 2012-0201).

References

Almquist, Y. (2011a). Childhood friendships and adult health: findings from the Aberdeen Children of the 1950s Cohort study. *European Journal of Public Health*, doi:10.1093/eurpub/ckr1045.

Almquist, Y. (2011b). *A class of origin: The school class as a social context and health disparities in a life-course perspective.* Stockholm: Stockholm University.

Almquist, Y. B., & Brännström, L. (2012). Childhood peer status and the clustering of adverse living conditions in adulthood. SOFI Working papers 1/2012. Stockholm: Stockholm University.

Bäckman, O., & Nilsson, A. (2011). Pathways to social exclusion - a life-course study. *European Sociological Review, 27*, 107-123. Bagwell, C. L., Newcomb, A. F., & Bukowski, W. M. (1998). Preadolescent friendship and peer rejection as predictors of adult adjustment. *Child Development, 69*, 140-153.

Bagwell, C. L., Schmidt, M. E., Newcomb, A. F., & Bukowski, W. M. (2001). Friendship and peer rejection as predictors of adult adjustment. *New Directions for Child and Adolescent Development, 91*, 25-49.

Bask, M. (2011). Cumulative disadvantage and connections between welfare problems. *Social Indicators Research, 103*, 443-464.

Bearman, P. S., & Moody, J. (2004). Suicide and friendships among American adolescents. *American Journal of Public Health, 94*, 89-95.

Bergman, L. R., Magnusson, D., & El-Khouri, B. (2003). *Studying individual development in an interindividual context. A person-oriented approach*. London: Lawrence Erlbaum Associates.

Bergman, L. R., & Trost, K. (2006). The person-oriented versus the variable-oriented approach: Are they complementary, opposites, or exploring different worlds? *Merrill-Palmer Quarterly, 52*, 601-632.

Berndt, T. J. (1996). Friendship quality affects adolescents' selfesteem and social behavior. In W. M. Bukowski, A. F. Newcomb & W. W. Hartup (Eds.), *The company they keep: Friendship during childhood and adolescence* (pp. 346–365). New York: Cambridge University Press.

Berndt, T. J. (2004). Children's friendships: shifts over a half-century in perspectives on their development and their effects. *Merrill-Palmer Quarterly, 50*, 206-223.

Berndt, T. J., Hawkins, J. A., & Jiao, Z. (1999). Influences of friends and friendships on adjustment to junior high school. *Merrill-Palmer Quarterly, 45*, 13-41.

Bukowski, W. M., & Hoza, B. (1989). Popularity and friendship: Issues in theory, measurement, and outcome. In T. J. Berndt & G. W. Ladd (Eds.), *Peer relationships in child development* (pp. 15-45). New York: Wiley.

Cairns, R. B., & Cairns, B. D. (1994). *Lifelines and risks*. Cambridge: Cambridge University Press.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98*, 310-357.

de Jong Giervald, J. (1998). A review of loneliness: concept and definitions, determinants and consequences. *Reviews in Clinical Gerontology, 8*, 73-30.

DiPrete, T. A., & Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: a review of theoretical and empirical developments. *Annual Review of Sociology, 32*, 271-297.

Elder, G. H. J. (1998). Life course as developmental theory. *Child Development, 69*, 1-12.

Erdley, C. A., Nangle, D. W., Newman, J. E., & Carpenter, E. M. (2001). Children's friendship experiences and psychological adjustment: theory and research. In D. W. Nangle & C. A. Erdley (Eds.), *New directions for child and adolescent development: the role of friendship in psychological adjustment* (pp. 5-23). San Francisco: Jossey-Bass.

Flora, J., & Segrin, C. (1998). Joint leisure time in friend and romantic relationships: the role of activity type, social skills and positivity. *Journal of Social and Personal Relationships, 15*, 711-718.

Fritzell, J., Gähler, M., & Nermo, M. (2007). Vad hände med 1990-talets stora förlorargrupper? Välfärd och ofärd under 2000-talet [What happened to the disadvantaged groups of the 1990s? Welfare and disadvantage during the first years of the new millennium]. *Socialvetenskaplig tidskrift, 2-3*, 110-133.

Fritzell, J., & Lundberg, O. (2000). Ofärdens ackumulering och samvariation [The accumulation and covariation of illfare]. *Delrapport från Kommittén Välfärdsbokslut [Interim report by the Welfare Commission]. SOU 2000:41*. Stockholm: Fritzes.

Furman, W., & Robbins, P. (1985). What's the point? Issues in the selection of treatment objectives. In B. H. Schneider, K. H. Rubin & J. E. Ledingham (Eds.), *Children's peer relations: Issues in assessment and intervention*(Insert page numbers here please). New York: Springer-Verlag.

Gest, S. D., Graham-Bermann, S. A., & Hartup, W. W. (2001). Peer experience: Common and unique features of number of friendships, social network centrality, and sociometric status. *Social Development, 10*, 23-40.

Gifford-Smith, M. E., & Brownell, C. A. (2003). Childhood peer relationships: social acceptance, friendships, and peer networks. *Journal of School Psychology, 41*, 235-284.

Hagenaars, J. A., & McCutcheon, A. L. (2002). Applied latent class analysis. New York: Cambridge University Press.

Hartup, W. W. (1984). The peer context in middle childhood. In W. A. Collins (Ed.), *Development during middle childhood* (pp. 240-282). Washington D. C.: National Academy Press.

Hartup, W. W., & Stevens, N. (1997). Friendships and adaption in the life course. *Psychological Bulletin, 121*, 355-370.

Hartup, W. W., & Stevens, N. (1999). Friendships and adaptation across the life span. *Current Directions in Psychological Science, 8*, 76-79.

Heaney, C. A., & Israel, B. A. (2008). Social network and social support. In K. Glanz, B. K. Rimer & K. Viswanath (Eds.), *Health behavior and health education: theory, research, and practice* (Insert page numbers here please). San Francisco: Jossey-Bass.

Henson, J. M., Reise, S. P., & Kim, K. H. (2007). Detecting mixtures from structural model differences using latent mixture modeling: A comparison of relative model fit statistics. *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 202-226.

Inderbitzen-Pisaruk, H., Clark, M. L., & Solano, C. H. (1992). Correlates of loneliness in midadolescence. *Journal of Youth and Adolescence, 21*, 151-167.

Kuha, J. (2004). AIC and BIC. Comparisons of assumptions and performance. *Sociological Methods Research, 33*, 188-229.

Ladd, G. W., & Troop-Gordon, W. (2003). The role of chronic peer difficulties in the development of children's psychological adjustment problems. *Child Development, 74*, 1344-1367.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: homophily in social networks. *Annual Review of Sociology, 27*, 415-444.

Östberg, V., & Modin, B. (2007). Status relations in school and their relevance for health in a life course perspective: Findings from the Aberdeen children of the 1950's cohort study. *Social Science & Medicine, 66*, 835-848.

Parker, J. G., & Seal, J. (1996). Forming, losing, renewing, and replacing friendships: Applying temporal parameters to the assessment of children's friendship experiences. *Child Development, 67*, 2248-2268.

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin, 132*, 98-131.

Rubin, K. H., Bukowski, W., & Parker, J. G. (1998). Peer interactions, relationships, and groups. In W. Darnon & N. Eisenberg (Eds.), *Handbook of child psychology* (pp. 619-700). New York: John Wiley.

Shantz, C. U., & Shantz, D. W. (1985). Conflict between children: Social cognitive and sociometric correlates. In M. W. Berkowitz (Ed.), *Peer conflict and psychological growth* (pp. 3-21). San Francisco: Jossey-Bass.

Shrum, W., Cheek, N. H. J., & Hunter, S. M. (1988). Friendship in school: gender and racial homophily. *Sociology of Education, 61*, 227-239.

Stenberg, S.-Å., & Vågerö, D. (2005). Cohort Profile: The Stockholm birth cohort of 1953. *International Journal of Epidemiology, 35*, 546-548.

van der Horst, M., & Coffé, H. (2012). How friendship network characteristics influence subjective well-being. *Social Indicators Research, 107*, 509-529.

Wentzel, K. R., & Caldwell, K. A. (1997). Friendships, peer acceptance, and group membership: Relations to academic achievement in middle school. *Child Development, 68*, 1198-1209.

Wentzel, K. R., Caldwell, K. A., & Barry, C. N. (2004). Friendships in middle school: Influences on motivation and school adjustment. *Journal of Educational Psychology, 96*, 195-203.

Zettergren, P., Bergman, L. R., & Wångby, M. (2006). Girls' stable peer status and their adulthood adjustment: A longitudinal study from age 10 to age 43. *International Journal of Behavioral Development, 30*, 315-325.

Appendix 1

Table A. Description of a summary measure* of problems in adulthood (n=10,694)

	Distribution			
	Males (n=5,022)		Females (n=5,672)	
	n	%	n	%
0 problems	2,642	52.6	2,812	49.6
1 problem	1,463	29.1	1,839	32.4
2 problems	654	13.0	815	14.4
3-4 problems	263	5.2	206	3.6

* A 'problem' is defined in the following way:
Educational level: Less than secondary education in 2007; Social assistance benefits: Receiving social assistance benefits at some point during the period 1992-2007; Unemployment: Being unemployed at some point during the period 1992-2007; Mental and behavioural disorders: At least one admission from the hospital during the period 1992-2007

Table B. Comparison between the cluster solution and the summary measure (n=10,694)

	E	e	U	eSUM
0 problems	5,454	0	0	0
1 problem	2,378	273	532	120
2 problems	53	471	556	389
3-4 problems	0	69	86	314

E: More education; U: More unemployment; E: Less education; eSUM: Less education, more social assistance benefits, more unemployment, and more mental and behavioural disorder

Table C. The association between childhood friendships and the clustering of adverse circumstances in adulthood. Results from ordinal regression analysis (n=10,694)

	Problems in adulthood (high score = more problems)		
	Model 1[a]	Model 2[b]	Model 3[c]
	OR (95 % CI)	OR (95 % CI)	OR (95 % CI)
Leisure time friends			
Yes (ref.)	1.00	1.00	1.00
No	1.64 (1.35-1.99)	1.62 (1.33-1.98)	1.54 (1.26-1.87)
Solitude			
No (ref.)	1.00	1.00	1.00
Yes	0.92 (0.78-1.08)	0.84 (0.71-0.99)	0.86 (0.73-1.01)
Best friend			
Yes (ref.)	1.00	1.00	1.00
No	1.62 (1.35-1.70)	1.51 (1.34-1.69)	1.45 (1.29-1.63)
Gender			
Males (ref.)		1.00	1.00
Females		1.06 (0.99-1.14)	1.04 (0.97-1.12)

[a] Adjusted for gender

[b] Adjusted for gender + mutual adjustment for friendship indicators

[c] Adjusted for gender, parental social class, parental education, parental income, family type, number of siblings, and parental mental health problems + mutual adjustment for friendship indicators

OR=Odds ratio

Longitudinal and Life Course Studies 2013 Volume 4 Issue 3 Pp 196 –217 *ISSN 1757-9597*

Work and family over the life course. A typology of French long-lasting couples using optimal matching

Ariane Pailhé, Ined, Paris, France
pailhe@ined.fr
Nicolas Robette, Université Versailles-St-Quentin and Ined, Paris, France
Anne Solaz, Ined, Paris, France

(Received March 2013 Revised August 2013) doi:10.14301/llcs.v4i3.250

Abstract

Decisions regarding the division of labour are part of a dynamic process of negotiation between partners and thus develop throughout the life cycle, in relation to family events such as successive childbirths. This article investigates the degree of interaction between work and family of both partners in the long run over the life course. Using an innovative methodology, optimal matching analysis, and data from the French Family and Employers Survey (2004-05), it defines a typology of work-family strategies for about 950 long-lasting couples observed from 3 years before couple formation to 18 years later, and identifies related key life-course stages. Finally, it analyses the factors leading to the various trajectories, and assesses whether preferences or opportunities and constraints greatly influence couples' profiles. Results bring to light a wide variety of work-family patterns, where the number of children and the woman's employment trajectory are the key determinants of these couple profiles. In spite of the trend towards equal opportunities, only women adapt their work patterns, except in the most "work-oriented couples". They use several strategies, by adjusting sequence and timing of births. In line with the standard human capital approach, partners' initial relative endowments influence couples' histories. A more traditional division of work is observed among less educated men and women, while women with greater human capital are more likely to remain employed through the transition to parenthood, whatever their partner's level of education.

Keywords: Work, family, couples, gender, life course, sequence analysis, optimal matching

1. Introduction

There has been a marked decline over decades in the degree to which men and women specialise in distinct economic activities. The couple model with male breadwinner is being replaced by the dual earner model in most developed countries (Blossfeld & Drobnic, 2001). However, childbirths remain a crucial stage in work division within couples. They often imply a re-setting of gender relations within the family, new trade-offs for spouses between career and family life and, often, a modification of the gender division of labour within and outside the family. The gender gap in the allocation of time usually increases after childbirth and with the number of children (Pailhé & Solaz, 2006; Bonke, Deding, Lausten & Stratton, 2008), with women devoting more time to housework and parental activities and men to market work (Van

der Lippe, 1994; Fine-Davis, Fagnani, Giovannini, Hojgaard, & Clarke, 2004; Regnier-Loillier, 2009).

Decisions regarding the division of labour are therefore not static; they are part of a dynamic process of negotiation between partners (Jansen & Liefbroer, 2006). Couples adapt their work-family strategies to their life stage (Becker & Moen, 1999; Anxo, Mencarini, Pailhé, Solaz, Tanturri et al., 2011), and the birth of a child is a key stage. For this reason, most studies analyse employment transitions after childbearing, generally at parity one. These studies focus on a particular moment of the life cycle and do not link the labour market consequences of the various demographic events, e.g. couple formation or subsequent childbirths. Nor do they account for the fluidity of employment status, i.e. movements into and out of the labour force, for many parents along the life course (Hynes & Clarkberg, 2005). Moreover, they focus on one direction of the shift, usually from family to work, i.e. the consequences of childbearing for labour market participation. Thus the reverse relationship, i.e. strategies dealing with childbearing decision and the timing of fertility as a result of employment opportunities, are rarely taken into account, although some couples may reduce the number of children they have or postpone childbearing decisions, during economic uncertainty on the labour market (Matysiak & Vignoli, 2009; Pailhé & Solaz, 2012).

Furthermore, little research has been devoted to describing these various strategies regarding work and family using both a longitudinal and a couple perspectives. Usually, studies linking professional careers and family histories adopt an individual point of view, often the female's, since women bear the main part of post-birth adjustments (Gutiérrez-Domènech, 2005; Del Bocca, Pasqua & Pronzato, 2009; Fouarge, Manzoni, Muffels & Luijkx, 2010). With the generalisation of the dual-earner couple model, male and female working patterns are increasingly inter-dependent. Since male and female labour participation decisions are made jointly, couples rather than individuals should be used as analytical units (Becker & Moen, 1999; Blossfeld & Drobnic, 2001).

This article investigates the degree of interaction between work and family of both partners in the long run across the life course, without assuming a direction of causality. It analyses the diversity of couples' work-family strategies, and identifies related key life course stages for long-lasting couples, i.e. couples in partnership for at least 15 years. Finally, it analyses the factors leading to the various trajectories, and assesses whether preferences or opportunities and constraints greatly influence couples' profiles.

Both the lack of adequate data and the limitations of standard quantitative techniques explain the scarcity of studies on couple trajectories. Hence, few longitudinal socio-demographic surveys interview both partners. Getting information on the partner from the respondent, limits retrospective inquiries and potentially biases the data due to the lack of accuracy and the specific position of the respondent. The second problem is technical and is related to the methods that can be used to analyse data at a couple level. As fertility and work choices may be highly inter-dependent, using standard regression models has generally failed because they necessarily make an *a priori* choice of direction of causality, i.e. family to work, or work to family. For this reason, they are both theoretically limited and statistically biased. Additional econometric techniques may be used to solve endogeneity problems, but finding an appropriate instrumental variable, that explains work outcomes and not family ones, remains difficult. Another challenge is taking into account the many transitions which compose the unfolding life course.

This article overcomes these methodological difficulties. First, the French "Enquête Familles et Employeurs" collected by INED (Institut National d'Etudes Démographiques) in 2004-2005 provides retrospective employment and family histories for each partner. These longitudinal data can be used to link family dynamics with employment trajectories. Second, it uses an innovative methodology to describe the work and family histories of both partners, without making any causal assumption about the relationship between family and work, or between a partner's characteristics and the other's behaviour. Sequence analysis, and more specifically optimal matching analysis, is used to examine the diverse ways that couples balance work and family in the long run.

2. Background

Two potential mechanisms may affect couples' work-family strategies. According to the standard neoclassical economic theory, women are assumed

to work for pay if the opportunity cost of working in the market is lower than the opportunity cost of staying at home (Becker, 1965). Women with higher human capital are more likely to remain at work across the life cycle, because they have a higher opportunity cost if they stop working. The presence of children is assumed to decrease female labour-force participation because it raises the value of women's time at home. The number of children has two opposite effects. On the one hand, the need for care-giving increases with the number of children, and home productivity is higher. On the other hand, financial needs increase with the number of children, so there is more need for the woman's income. The age of children is also important: as children grow older they require less supervision but more financial expenditure, which decreases the value of a woman's time spent at home. The impact of children on career path concerns the partner who invests in home production, either the man or woman in theory, but is usually studied from the female point of view.

At the couple level, spouses are assumed to specialise in market work and domestic production according to their comparative advantage. Women are supposed to have a comparative advantage in domestic work, as they often earn less on average than their husbands for many reasons. A more recent and growing body of the literature on the female labour supply is moving away from this human capital model. These alternative economic models dealing with household decision-making processes and strategic interactions between spouses have given an alternative explanation of labour supply (Chiappori, 1988). In that framework, each spouse maximises their own welfare, and takes into account their current and future bargaining position in the decision to participate in the labour force or not. The individual bargaining positions depend on relative spouse resources such as age, income, level of education and occupational position.

The sociological literature rather emphasises the influence of norms and gender roles in explaining the division of labour between spouses. It argues that, despite women's massive entry into the labour force, men and women often play different roles in their partnerships. For instance, according to the "doing gender" approach, the greater involvement of women in housework and of men in market work, is a way of reaffirming their gender identities (Goffman, 1977). Work can be seen as an instrument enabling men to exercise their role as providers, and housework as an instrument enabling women to exercise their role as family carers. This relationship of dependency in the couple is in line with normative expectations. The strong social pressure that allocates child rearing to women and disapproves of working mothers explains the negative effect of children on the female labour supply (Shelton & John, 1996). This specialisation of women in child-rearing tasks is explained in "role theory" by the substantial social rewards they receive when they participate in mothering, whereas men receive far fewer such rewards for fathering (Van der Lippe, 1994).

Conversely, preference theory focuses on lifestyle preferences as a predictor of women's choices regarding fertility and employment patterns (Hakim, 2003). This theory emphasises the heterogeneity of women in each country and distinguishes three ideal-types of women, based on preferences related to combinations of market work and family work: they are home-centred, work-centred, or adaptive women. Adaptive women, who are the majority, combine employment and family work without giving priority to either, and often opt for part-time work after childbearing. Work-centred women prioritise their jobs and many remain childless, while home-centred women are focused on family life and avoid market work as long as economically possible.

Considerable empirical research using cross-sectional data shows that the current employment statuses of both partners and the fact of having children both play a role in the household division of labour (Craig & Killian, 2010). Fewer studies focus on women's life courses, and even fewer on men's (Ravanera, Rajulton, Burch & Le Bourdais, 2002). However, the timing and sequence of events are important. Indeed, the timing of parenthood is affected by the position of both partners in the labour market, especially the stability of contracts (Vignoli, Drefahl & De Santis, 2012). Similarly, the timing of marriage plays a role in the division of household labour: the amount of housework is lower for men and women who entered marriage at a later age (Pittman & Blanchard, 1996). Parenthood transition has also a significant effect on women's housework, while there are no effects on men's (Gupta, 1999; Baxter, Hewitt & Haynes, 2008). Most studies examine the short-term effects

of parenthood, mainly due to limited data. However, two recent studies show that it has long-term effects. This highly gendered division of labour caused by parenthood continues more than 10 years after childbirth in Germany (Kühhirt, 2012). Even if there are cross-country differences, the mothers' participation rates do not return to pre-birth levels in Germany, the Netherlands and the UK (Fouarge et al., 2010). Finally, the few existing studies are relatively consistent in noting that children affect mothers' career paths and non-market work, more than fathers'. These studies focus on the impact of the first child and devote no attention to subsequent childbirths.

We analyse the long-term inter-relations between work and family in France using a life course perspective, paying attention to both the career path and subsequent family events. Special attention is paid to the human capital of both partners, such as age and educational attainment in the process of work division throughout the life-cycle. We also analyse how educational and cultural background, i.e. the family values inherited from parents, religious feeling, ethnic origin, working history of the mothers and number of siblings, might be related to the various profiles of couples.

France is particularly interesting since it has a high maternal full-time labour force participation rate. Moreover, France is one of Europe's most fertile countries, with 2.01 children per woman on average in 2012. The quite generous and diversified family benefit system and the encouragement of early socialisation of children through nursery schools[i], enable various strategies regarding fertility and a range of adjustments after childbearing.

3. Data and Method

3.1. Data

The data used for this study is the "Enquête Familles et Employeurs" (Families and Employers Survey, EFE hereafter), a French nationwide representative survey conducted by INED in 2004-2005[ii]. A sample of 9,547 individuals (5,107 women and 4,440 men) aged from 20 to 49 were interviewed, with a maximum of two persons per household. In total, data on 3,279 couples were gathered. This survey is one of the few in which retrospective information on family and professional life course is available for both members of couples.

Via a precise description of socio-demographic characteristics and current employment situation, the survey gives detailed information on partnership and fertility histories, i.e. the timing of cohabiting or married unions lasting at least 6 months[iii] and of each birth[iv]. It also includes retrospective work histories: respondents were asked for their employment situation in every year from the age of 18. This information, taking the form of an event-history calendar, is based on the respondents' situations that have lasted at least 6 continuous months. For each spell longer than 6 months, six main types of situation are possible: employment (distinguishing between part-time and full-time work)-, unemployment, studies or training, military service, parental leave, other economic inactivity. An additional "status" was included: combination of short spells of employment or training with short spells of unemployment. For each of these seven states, the respondents were asked to point out the corresponding years of occurrence. It might be that more than one situation was identified for a given year[v]. As in all retrospective surveys, the answers on job and family history may be at risk of memory problems or lapses (Auriat, 1991). However, for this purpose, special care was taken in the collecting process, by limiting contents to the most basic events such as couple formation, births, moves or changes in employment status, and by linking these events with a computerised calendar to check response coherence.

3.2. Sample selection

As the study is intended to analyse the relationship between couple's employment situation and successive births, the couple's life course is observed over the long run. In order to take into account the significant proportion in France of couples with at least 3 children[vi] and childbearing postponements, couples in partnership for at least 15 years are selected[vii]. Men and women were observed over 18 years, i.e. from three years before couple formation to 15 years after[viii]. These long-lasting couples are specific in a context of increasing divorce and separation, but a long period is necessary to observe couple specialisation. Moreover, the longer histories are followed, the more complete fertility history we observe. A shorter duration would prevent us from examining how the work-family balance occurs after the birth

of high-parity children. Hence, 15 years after couple formation, almost all third births within the couple are observed (i.e. 97%) and more than 90% of the fourth births.

Our sample is made up of individuals belonging to the 1954-1968 cohort, i.e. aged from 36 to 49 in 2004. Among them, 5% have never been in partnership and 12% have had one or more partnerships for less than 15 years. Long-lasting couples are the majority since 83% have had at least one partnership for 15 years or more: 32% were single at the time of the survey, 16% whose partner was not surveyed[ix] and 35% were in partnership and we collected information for both partners. Our sample selects the latter and includes 941 cohabiting or married couples (1,882 individuals)[x].

Table 6 in the Appendix compares people in our sub-sample to people of the same cohort in the four different groups presented above. Of course, individuals who remain in partnership for a long time are different from individuals who do not: they form their first and current couple partnership earlier and have more children than individuals of the same cohort with shorter or no partnerships. There are small dissimilarities between those two populations regarding the distribution by educational level – they less frequently hold a university diploma – but, in terms of occupational careers, individuals in the sample have had longer periods of stable employment than the population at large, and shorter periods of unemployment, short-term jobs and inactivity. Thus, the analysis is restricted to quite stable couples, in terms of both family situation and employment profiles.

3.3. Definition of couples' activity and parental status

Six activity states are taken into account: studies, inactivity, unemployment, short jobs, part-time stable jobs and full-time stable jobs. Short jobs are employment spells lasting less than 6 months. Some states have been grouped together: studies and national service have been put together, since national service was entered just after the end of schooling, and also inactivity and parental leave. These states are not equally spread throughout the couples' histories (Table 1). During the union, while men spend most time in full-time stable employment, the situation is highly different for women. Full-time stable employment represents on average less than half of their life-course time, inactivity more than 20% and part-time stable employment about 13%.

Concerning the parental trajectory, at the time of the survey, 4% of the couples had no child, 15% had one, 49% two, 25% three and only 8% four or more. Therefore the parental variable is coded in four dynamic states: 0 child; 1 child; 2 children; 3 children or more[xi].

Table 1. Percentage of time spent in the various occupational states during the union, for women and men

Occupation	Women	Men
Studies	3.9	4.2
Inactivity	23.0	0.5
Unemployment	3.8	1.1
Short jobs	5.0	3.3
Part-time stable employment	14.2	1.4
Full-time stable employment	50.1	89.5
Total	100.0	100.0
N	941	941

Source: EFE, 2004-2005
Sample: Couples of the 1954-1968 cohort in partnership for at least 15 years.

3.4. A sequence analysis approach

We use optimal matching analysis (OMA) to define couple trajectories. This method is based on a set of dynamic algorithms mostly used in molecular biology to study DNA strings. It was introduced into the field of social sciences by Andrew Abbott in the 1980s (Abbott & Forrest, 1986). Since then, optimal matching has been successfully applied to life courses and family histories (Billari, 2001; Schoon, McCullough, Joshi, Wiggins & Bynner, 2001; Aassve, Billari & Piccarreta, 2007; Pollock, 2007) and careers (Stovel, Savage & Bearman , 1996; Blair-Loy, 1999; Han & Moen, 1999; Robette & Thibault, 2008; Anyadikes-Danes & McVicar, 2010), as well as many other topics such as couples' time-schedules (Lesnard, 2010). Despite a few criticisms (Wu, 2000; Elzinga, 2003), it has now become widely accepted as a useful tool for life course scholars (Aisenbrey & Fasang, 2010).

Optimal matching is based on measuring similarities between pairs of sequences. More precisely, it measures the dissimilarity between two sequences by calculating the cost of the transformation of one sequence into the other. The transformation is carried out by means of three elementary operations: insertion (one element is inserted into the sequence), deletion (one element is deleted from the sequence) and substitution (one element is replaced by another). Each elementary operation can be assigned a specific cost. A series of operations costs the equivalent of the sum of the elementary operations involved, and eventually the dissimilarity between two sequences is equal to the minimal cost of transformation of one sequence into the other. Optimal matching of each pair of sequences in the data leads to the creation of a distance matrix, which can then be used as an input for clustering or scaling techniques, for instance.

The dissimilarity measuring procedure can be illustrated by the following example. Consider two sequences of letters, e.g. two words: "family" and "marriage" (Figure 1). The transformation of "family" into "marriage" can be carried out in this way: F is replaced by M, R in inserted between A and M, M is replaced by R, L by A, Y by G and finally E is inserted at the end of the word. If substitution, insertion and deletion operations all have the same unitary cost, then matching "family" and "marriage" will cost 6: the dissimilarity between these two sequences is equal to 6.

Figure 1. Optimal matching of "marriage" and "family"

M	A	R	R	I	A	G	E
F	A	M	I	L	Y	_	_

The couple sequences we want to analyse can be considered as multi-dimensional: they include the woman's occupational history, the man's occupational history and the partners' parental history. This aspect requires a specific strategy. From a methodological point of view, several alternatives emerge in the life course literature based on sequence analysis (Blanchard, 2010; Gauthier, Widmer, Bucher & Notredame, 2010). A first strategy involves first using optimal matching to compute one distance matrix for each dimension and then to carry out a linear combination of these matrices into one by means of linear combination (Han & Moen, 1999). A second strategy involves creating a new state variable combining the single states associated to each dimension (see for instance Blair-Loy, 1999; Aassve, Billari & Piccarreta, 2007; Pollock, 2007; Robette, 2010): for example, a combined state might be "female with a part-time job, male with a full-time job, one child". The advantage of this second alternative – the one we chose – is to take into account the interdependency of the dimensions right from the coding stage. Moreover, although the combined state variable potentially has 6*6*4=144 states[xii] which may seem a large number, previous studies have emphasised the robustness of OMA even in cases of large state spaces (Robette, 2010).

The choice of substitution, insertion and deletion costs is an important stage in OMA – and the main problem in the application of optimal matching (Wu, 2000): it involves a trade-off between favouring the contemporaneousness of states and emphasising their order within sequences (Lesnard, 2010). As far as substitution costs are concerned, since there are no clear theoretical assumptions about the relative resemblance between states, they are set to a constant value of 2. The insertion and deletion cost (called *indel*[xiii] cost) is set to slightly more than half the substitution cost (i.e. 1.1), which avoids an excessive use of *indel* operations and retains the information concerning the time when transitions occur (MacIndoe & Abbott, 2004)[xiv]. Computations

are done with the TraMineR package in R (Gabadinho, Ritschard, Müller & Studer, 2011).

Our sequences are built from an event, i.e. couple formation. A given age, for instance 18 years old, would have been an alternative, but clusters would have been driven mostly by the duration of studies and so by educational level, which was not our main point of interest. Besides, while computing for each cluster median ages at couple formation for women and men, we observe that the variations between clusters are quite limited.

Each year, a couple's status is defined by the cross-combination of parity (0; 1; 2; 3 children or more), the woman's employment status (inactivity; unemployment; studies; part-time; short jobs; stable job) and the man's employment status (ditto). This means our approach takes into account the differences in association between fertility and employment. OMA, as we apply it, deals very well with the relationship between family and employment. It shows graphically how childbearing and women and men's employment are related, for each parity separately. So this method offers a very good description of life courses; it does not aim at understanding underlying mechanisms or establishing any causal relationship.

4. A variety of couple histories

From these similarities of couples' professional and family histories, an 11-cluster typology is built, using Hierarchical Cluster Analysis with Ward's criterion on the distance matrix produced by OMA[xv]. Table 2 summarises the main characteristics of these different clusters, according to the women's and men's activity profiles, and transitions to parenthood, and the level of intra-class dissimilarity, which indicates how heterogeneous the class is. The number of couple profiles of is quite high, reflecting the various forms of work-family adjustment in France. None of these groups dominates clearly, showing that work-family patterns are very heterogeneous in France, in terms of both employment history and childbearing history. The two main groups are couples with 2 children and full-time employment for both partners, and the last cluster called "*multiple states*", the most heterogeneous one according to the intra-cluster dissimilarity index, which seems to be the remaining cluster[xvi] containing males and females who change professional status several times during their life course. These two groups each represent about 15% of couples.

The number of children after 15 years of union and the pace of births – i.e. quantum and tempo - differ between clusters. But almost all couples have at least one child, even the most work-oriented ones. This result comes from both the fact that France remains a family-oriented country with strong parental norms and generous family policy, but also from the selection of our sample, which over-represented long-lasting couples. The few childless couples (after 15 years of life together) are almost all in the *0 or 1 postponed child* cluster that represents 10% of couples.

Table 2. 11 couple profiles

Pattern name	Woman's activity	Man's activity	Children	N	%	Heterogeneity (average intra-cluster dissimilarity)
Mother works full time				***463***	***49.2***	
Work-oriented, 1 child	Stable employment	Stable employment	1 (or 2)	90	9.6	12.9
0 or 1 postponed child	Stable employment	Stable employment	0. 1 or 2 late	93	9.9	16.3
Spaced births	Stable employment	Stable employment	2 spaced	71	7.5	19.2
2 children and full-time employment	Stable employment	Stable employment	2	133	14.1	18.9
3 children and full-time employment	Stable employment	Stable employment	3 or more	76	8.1	20.7
Mother works part-time				***121***	***12.8***	
2 children and part-time employment	Shift to part-time	Stable employment	2	67	7.1	20.1
3 children and part-time employment	Shift to part-time	Stable employment	3 or more	54	5.7	22.1
Shift to inactivity				***209***	***22.3***	
1 child and inactivity	Shift to inactivity	Stable employment	1 or 2	60	6.4	25.0
2 children and inactivity	Shift to inactivity	Stable employment	2	61	6.5	19.3
3 children and inactivity	Shift to inactivity	Stable employment	3 or more	88	9.4	21.7
Multiple states	Different status	Different status	Various	***148***	**15.7**	33.1
TOTAL				*941*	*100*	

Source: EFE, 2004-2005.
Sample: Couples of the 1954-1968 cohort in partnership for at least 15 years

Male profiles are very homogeneous: most of them are in stable employment all along their trajectory whatever the cluster. Thus, men's occupational careers are not impacted by childbirths. This result, which suggests that men make few trade-offs for the sake of family life, is in line with studies conducted in Germany, the UK and the US (Kaufman & Uhlenberg, 2000; Kühhirt, 2012) and confirms previous results for France (Pailhé & Solaz, 2006). On the other hand, female profiles are very diverse. Some women experience inactivity or part-time job spells, of varying length and/or frequency, sometimes equally spread and in other cases mostly at the beginning or at the end of the union (see tables 7 and 8 in the Appendix, which give average duration of spells and frequencies of participation). Some women stay inactive or work full-time all the time. The variety of work-family

adjustments mainly depends on women and on strategies regarding childbearing. However, dual earners working full-time prevail: about half of couples follow this egalitarian pattern. Couples for whom full-time employment is the most continuous belong to the *One child work-oriented* cluster (10% of couples). In this group, for men as for women, births do not involve any visible change in the work history. The first child is born during the first eight years but it is not followed by a second birth. Among the other full-time dual-earner clusters, a large share has at least two children.

On the other hand, some women shift to long-term inactivity (22% of the couples) or to part-time jobs (13%) after childbearing. Part-time usually takes place after the birth of the second child and continues thereafter. The situation is particularly frequent for two-child families or when the first three children are very close in age. However, one small but interesting cluster, called *One child inactivity*, is made up of couples for whom inactivity begins from, or even before, the first child. A large proportion of women in this group work for some years before couple formation but some stop working at this particular moment. Some of them face long-term unemployment at the moment of integration into the labour market and may feel discouraged and choose to invest more in family life.

The computation of median spacing between events[xvii] (Table 3), state distribution plots and index plots (see Figures 2 and 3 in the Appendix), provide more precise and original information about the timing and rhythm of family and occupational transitions. For instance, it brings to light a *Spaced births* cluster (7.5% of couples). The time interval between the first two births in this cluster is noticeably larger than in the other clusters with at least two children, i.e. 6 years on average compared to 2-3 years in the other groups. Other couples postpone the birth of the first child. The first child of the couples in the *0 child or 1 postponed child* cluster, if born at all, arrives very late, after more than 7 years of cohabiting union. One reason is that a significant proportion of these couples began living together while they were still students and some of them faced several spells of unemployment or short-job periods. Young people entering the French labour market face a high risk of unemployment and job insecurity, even more pronounced during an economic downturn. One way to limit the deleterious effects of reconciling motherhood and work is to deliberately limit the number of children: that is the compromise chosen by couples in the *Work-oriented 1 child* and *0 or 1 postponed child* clusters.

More generally, a lower number of children is associated with more spaced family events, particularly in clusters with part-time jobs or inactivity. For instance, couples in the *3 children and inactivity* group have their first child 1.5 years after couple formation, while those of the *2 children and inactivity* group 2.1 years after and those of the *1 child and inactivity* group 5.5 years after.

Lastly, the final cluster called "*multiple states*" contains males and females who both change professional status several times during their life course. Even when both partners have discontinuous careers, the woman's trajectory is much more discontinuous than the man's. The alternation of unemployment spells, short-term jobs and training spells even after the end of initial studies, reveals these women's strong commitment to work. As a consequence, their rhythm of family events is very similar to that of the *Spaced births* cluster.

Table 3. Median duration between demographic events, by cluster (years)

	Cluster										
	Work-oriented, 1 child	0 or 1 postponed child	Spaced births	2 children and full-time job	3 children and full-time job	2 children and part-time job	3 children and part-time job	1 child and inactivity	2 children and inactivity	3 children and inactivity	Multiple states
Couple → marriage	0.8	3.7	2.5	1.0	0.8	1.3	0.5	0.9	0.3	0.5	1.8
Couple → 1st birth	3.7	14.2	4.9	2.5	2.3	3.5	1.9	5.5	2.1	1.5	4.0
1st birth → 2nd birth	-	-	6.0	3.1	2.0	3.4	1.4	-	3.2	1.7	6.5
2nd birth → 3rd birth	-	-	-	-	2.4	-	2.2	-	-	2.3	-

Source: EFE, 2004-2005.
Sample: Couples of the 1954-1968 cohort in partnership for at least 15 years

5. Characteristics of various couples

Beyond their family and employment histories, groups of couples differ in their individual and family background characteristics. These characteristics are particularly interesting because they were not included in the clustering process, and did not contribute to its building. We observe relationships between background and couples' lifecycle patterns (see the results in Tables 4 and 5). As individual and couple variables, we analyse partners' relative endowment in terms of age[xviii], level of education, age and activity status at partnership formation, and marital status. Some background variables are also introduced, since individuals may adhere to behaviour, values, and norms that dominated during their childhood, and this may affect their preferences regarding work and family. Characteristics of the family of origin include whether the respondent had at least two siblings, and an indicator of immigrant background, separating French-born from European and non-European immigrants. An additional indicator of the cultural context is introduced through the individual's religiosity, i.e. if the respondent reported that religion is important in their daily life.

The two most career-oriented groups, i.e. the *Work-oriented one child* group and the *0 or 1 postponed child* group share common traits. They both have more egalitarian and non-traditional characteristics and background. They are more likely to be non-married, to have grown up with few siblings and a working mother, especially men (the two characteristics might be linked). Women are more often older in these couples and men are much less religious. These two groups are mainly distinguished by the female level of schooling, which is higher in the *0-1 postponed child* group. Women more frequently hold a tertiary level of education and their long period of studies makes them more likely to still be students at the time of couple formation.

In the two groups that combine full-time employment and a medium or high number of children, i.e. the *2 children full-time* and *3 children full-time* groups, women also have a heritage of continuous employment from their mothers and their mothers-in-law. Women holding a tertiary level of education are also over-represented. However, the mothers of the *3 children* group display more traditional characteristics. They come from larger families, religion is more important and the age gap between partners (M-F) is the highest of all clusters.

In the last career-oriented group, the *spaced births* group, men and women also have a weak religious feeling and a high level of education. But the men and women more often come from families where the mother never worked or interrupted her career path. Women that reconcile employment and family responsibilities through part-time work more frequently hold an intermediate level of education.

Conversely, couples where women are inactive, share both low levels of human capital and a more traditional background. Here, men and women have a low level of education. Men are more often older than their spouses, both partners' mothers have

more frequently never worked and they are more frequently non-European immigrants, particularly women. The main difference between these 3 groups is related to their number of siblings. In particular, both men and women of the *3 children and inactivity* group are more frequently born in large families.

Finally, the group *multiple states* includes couples where the spouses are close in age, partly because they more often met during studies. The share of highly educated is the highest and their long integration into the labour market comes from their long schooling, which implies a postponement of subsequent family stages.

From this description, it is clear that the division of labour within couples and fertility decision patterns are jointly determined by both the partners' human capital and cultural and family values. Couples' choices regarding the division of labour are mainly related to their human capital investment. In particular, people who studied for a long period are more likely to work continuously and thereby gain labour market returns on their human capital investment. So, couples in which women reconcile childbearing and work full-time with at least three children, are those with the most educated women on average. Their income probably lets them afford childcare expenses. On the other hand, couples in which the division of labour is more traditional, i.e. when the woman becomes a housewife, are those with the least educated women, given that their opportunity cost for giving up working is lower. Between these two extremes, women with an intermediate level of education more often opt for part-time. Thus, the female career path is clearly related to their level of education.[xix] Education level matters less regarding the number of children born. Men's level of education is more uniform across clusters than women's, except for men whose partners stay at home, who are less educated on average, like their partners, probably due to a strong homogamy effect in the matching process.

Fertility and labour market choices are also related to family background, with a noticeable inter-generational transmission of preferences regarding work and family. Having grown up in a large family is positively correlated with family size (as shown by Booth & Kee, 2006) since it may indicate that the respondent was raised in a family with strong family orientation (Michaël & Tuma, 1985). However, our taxonomy shows that the relationship between the number of siblings and the number of children is mediated by the type of division of labour within the couple. For instance, couples that come from smaller families and where women opt for part-time work have more children. Other cultural aspects affect couples' trajectories. So immigrants, whose values about family and mothers' work are influenced by norms from their country of origin, are more likely to belong to the clusters with the most traditional division of labour. Religiosity is also a key determinant of the number of children and the choice of female career path.

A key factor in the couples' division of labour over the life cycle is related to the work history of respondents' mothers, i.e. whether they have been working all their life long, interrupted their career to raise their children or remained out of the labour force. An inter-generational transmission of family models and values can be seen with both their daughters', and also their daughters'-in-law division of labour[xx]. For instance, the cluster with the highest proportion of sons of mothers with a continuous working path, is the most work-oriented one. Similarly, the cluster with the highest proportion of housewife mothers is made up of couples where the woman's mother is a housewife as well. However, one notes a development over generations of female career paths towards stronger preference to combine work and care. For instance, sons whose mothers have never worked are more likely to belong to couples with 2 children and a part-time partner, or to those who have spaced their births. This may reflect a negotiation between partners about the division of labour in this generation, for whom the inactivity model is not the major norm any more. In the same trend of modernisation of the traditional model in the following generation, couples in which the woman is working part-time, whatever the number of children, are a little more likely to have grown up with a housewife mother. Couples with several children and a woman working part-time may probably have similar values about motherhood roles as their mothers, but they may have adapted it by reducing the number of children in order to carry on working.

Table 4. Couple characteristics by cluster

Cluster	Work-oriented, 1 child	0 or 1 postponed child	Spaced births	2 children and full-time job	3 children and full-time job	2 children and part-time job	3 children and part-time job	1 child and inactivity	2 children and inactivity	3 children and inactivity	Multiple States	Total
Age gap between partners (%)												
Men 5 years +older	10	15	10	14	21	10	20	23	20	23	9	**15**
Men 3-4 years older	21	23	13	16	21	22	17	27	26	20	19	**20**
Equal age	58	53	72	63	51	64	61	42	51	47	67	**58**
Women 3years + older	11	10	5	7	7	2	2	8	3	11	6	**7**
Age at couple formation (median)												
Men	22.9	23.9	22.6	22.7	23.6	23.5	24.2	23.7	23.2	23,0	22.8	**23.2**
Women	22.1	21.8	21.4	21.9	21.4	22.3	22.3	21.1	20.8	20.8	21.7	**21.6**
Married couples (%)	84.4	78.5	87.3	88.7	88.1	91.0	92.6	93.3	96.7	97.7	89.2	**89.2**
Men's education level (%)												
None or primary	22	20	13	20	26	21	26	35	40	36	23	**25**
Secondary	64	65	68	59	51	58	50	52	42	45	40	**54**
University	13	15	20	21	22	21	24	13	18	19	37	**21**
Women's education level (%)												
None or primary	26	19	18	26	19	19	15	43	48	45	26	**27**
Secondary	54	52	58	43	51	60	59	35	38	38	41	**47**
University	20	29	24	32	31	21	26	22	13	17	32	**25**
Men's employment status at couple formation (%)												
Working	88	88	81	83	80	89	82	83	87	83	68	**82**
Unemployed	0	2	4	1	3	1	0	0	0	1	2	**1**
Student	10	10	15	16	14	10	18	15	13	16	28	**16**
Out of labour force	2	0	0	0	3	0	0	2	0	0	2	**1**
Women's employment status at the couple formation (%)												
Working	82	82	82	72	70	86	73	42	56	56	74	**71**
Unemployed	2	1	1	4	1	1	7	13	7	4	3	**4**
Student	13	17	13	20	26	9	18	8	11	12	18	**16**
Out of labour force	3	0	4	4	3	4	2	37	26	28	5	**9**

Table 5. Family and cultural background by cluster

Cluster	Work-oriented, 1 child	0 or 1 postponed child	Spaced births	2 children and full-time job	3 children and full-time job	2 children and part-time job	3 children and part-time job	1 child and inactivity	2 children and inactivity	3 children and inactivity	Multiple States	Total
Life cycle working status of men's mother (%)												
Mother always worked	48	43	30	42	39	36	44	38	31	27	28	**37**
Mother occasionally worked	20	18	20	20	13	12	9	17	16	14	22	**17**
Mother never worked	32	39	49	35	45	48	43	43	51	55	46	**43**
Other	0	0	1	3	3	4	4	2	2	5	5	**3**
Life cycle working status of women's mother (%)												
Mother always worked	39	37	30	39	39	31	39	32	31	32	34	**35**
Mother occasionally worked	27	24	30	29	26	25	26	15	25	17	18	**24**
Mother never worked	32	38	39	29	33	43	35	53	39	49	43	**39**
Other	2	2	1	3	1	0	0	0	5	2	4	**2**
Childhood in a large family (%)												
Man	67	66	75	80	79	75	72	77	70	90	68	**74**
Woman	64	76	58	68	78	63	70	72	69	82	68	**70**
Men's immigration status (%)												
Born in France	97	96	97	93	93	90	89	87	90	91	89	**92**
Immigrant from Europe	3	2	3	5	1	7	9	5	3	2	3	**4**
Immigrant from elsewhere	0	2	0	2	5	3	2	8	7	7	8	**4**
Women's immigration status (%)												
Born in France	93	96	97	95	96	97	91	90	84	90	93	**93**
Immigrant from Europe	7	3	1	2	0	1	7	3	7	3	5	**4**
Immigrant from elsewhere	0	1	1	2	4	1	2	7	10	7	2	**3**
Men's religious feeling (%)												
None	42	40	41	36	33	31	30	27	10	28	41	**34**
Little	37	32	37	35	37	36	30	32	48	26	34	**35**
Important or used to be important	21	28	23	29	30	33	40	42	42	45	25	**31**
Women's religious feeling (%)												
None	47	48	44	53	37	40	35	38	30	42	49	**44**
Little	31	35	39	27	32	36	31	30	39	32	35	**33**
Important or used to be important	21	17	17	21	32	24	33	32	31	26	16	**23**

6. Conclusion and discussion

The sequence analysis approach (optimal matching analysis) we adopted, led to the identification of the major patterns of couples' work and family histories in France. This paper produces evidence of the effectiveness of sequence analysis in describing jointly the trajectories of both partners' work and family histories. It can be used to analyse long term effects of subsequent childbirths on male and female labour market participation.

Although the analysis is restricted to quite stable couples in terms of family and occupational trajectories, who are a specific sub-sample of the overall population, couple patterns are far from being uniform. The results show a high heterogeneity among the couples and bring to light a wide variety of work-family patterns: 11 distinct profiles are recorded. As the number of children is a key determinant of these profiles, our methodological choice of observing a long period - and therefore more completed fertility - but fewer couples rather than a shorter period and more couples, is an appropriate one. These profiles differ mainly by family size and the woman's employment trajectory, whereas that of the man is very similar between profiles, except for those with multiple states. Male careers are influenced neither by family events nor by their partners' careers. The fact that the variability of couple histories is largely more connected to women's careers and hardly at all to men's careers may moderate the contribution of the couple-level longitudinal approach adopted here. Yet its strength is that it empirically demonstrates that decisions to work less after the birth of a child are still made by women, even in a society where working mothers are usual and dual-earner couples are frequent.

More detailed information about occupational situations would probably indicate some kinds of work arrangements by men, not visible with our six-activity indicator in. However, more detailed indicators such as working hours or a subjective indicator of work investment may well be more informative but they are rarely available on a longitudinal basis, and memory bias may be more serious if they are recorded retrospectively. The small variations among men may also be driven by the selection of the observed population, and particularly by the positive correlation between a union and professional stability. Men with long-term employment are more likely to be selected for partnership, and men in long-term relationships are also more likely to be professionally stable.

The division of work within couples between market and domestic activities develops over the life cycle. Whereas the activity statuses of men and women are very similar 3 years before couple formation in every profile, they start to diverge at couple formation. They differ significantly after the birth of the first child, except in the case of work-oriented couples.

However, there is no polarisation of women's profiles. In contrast to the typology of Hakim (2003) that distinguishes 3 ideal-types of women, this analysis identifies many more strategies. In particular, it shows that the sequence and timing of births seem to be endogenous to women's career preferences. Some couples prioritise their careers at some points in their life course by postponing first childbearing or spacing births. This lesser polarisation may be due to France's comprehensive family policy, which is both relatively generous and highly diverse in the tools available to help parents. This policy benefits various couple profiles, and perhaps gives more freedom in their way of balancing family and work. French parents are allowed to take parental leave on a part-time or full-time basis, and benefit from subsidised childcare and free school for children from 3, family allowances from the second child and particularly further tax reductions from the third child. However, socio-economic differences remain. In line with the standard human capital approach, partners' initial relative endowments influence couples' histories. A more traditional division of work is observed among less educated men and women, while women with greater human capital are more likely to remain employed through the transition to parenthood, whatever their partner's level of education. With the regular increase in the female level of education and permanent support to families from the State, the egalitarian model implemented by the most educated people might prevail in the future. But this effect may be attenuated by firm attitudes regarding the family, both regarding family life and the employment path. Hence, religiosity and being non-European immigrants are factors in favour of large families and focus on family life. Moreover, there is an inter-generational transmission of family models and values between mothers and daughters and daughters-in-law. One may note a shift of female career paths over generations towards a stronger preference for combining work and care.

References

Aassve, A., Billari, F.C., & Piccarreta, R. (2007). Strings of adulthood: a sequence analysis of young British women's work-family trajectories. *European Journal of Population, 23,* 369-388.

Abbott, A., & Forrest, J. (1986). Optimal Matching Methods for Historical Sequences. *Journal of Interdisciplinary History, 16,* 471-494.

Aisenbrey, S., & Fasang, A.E. (2010). New Life for Old Ideas: The ''Second Wave'' of Sequence Analysis Bringing the ''Course'' Back Into the Life Course. *Sociological Methods and Research, 38,* 420-462.

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

Anxo, D., Mencarini, L., Pailhé, A., Solaz, A., Tanturri, M.L., & Flood, L. (2011).). Gender Differences in Time Use over the Life Course in France, Italy, Sweden, and the US,, *Feminist economics*, *17*,, 159–195.Auriat, N. (1991.). Who forgets? An analysis of memory effects in a retrospective survey on migration history. *European Journal of Population, 7*, 311-342.

Baxter, J., Hewitt, B., & Haynes, M. (2008). Life Course Transitions and Housework: Marriage, Parenthood, and Time on Housework. *Journal of Marriage and Family, 70,* 259-272.

Becker, G.S. (1965). A Theory of the Allocation of Time. *The Economic Journal*, *299*, 493-517

Becker, P.E., & Moen, P. (1999). Scaling Back: Dual-Earner Couples' Work-Family Strategies. *Journal of Marriage and Family, 61*, 995-1007.

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

Blair-Loy, M. (1999). Career patterns of executive women in finance: an optimal matching analysis. *The American Journal of Sociology*, *104*, 1346-1397.

Blanchard, P. (2010). *Analyse séquentielle et carrières militantes*, Report. Accessed 13th March from http://halshs.archives-ouvertes.fr/hal-00476193/

Blossfeld, H.P., & Drobnic, S. (2001). *Careers of couples in contemporary societies. A cross-national comparison of the transition from male breadwinner to dual-earner families*. Oxford: Oxford University Press.

Bonke, J., Deding, M., Lausten, M., & Stratton, L.S. (2008). Intra-Household Specialization in Housework in the United States and Denmark. *Social Science Quarterly, 89*, 1023-1043.

Del Boca, D., Pasqua, S., & Pronzato, C. (2009). Motherhood and market work decisions in institutional context: a European perspective. *Oxford Economic Papers, 61*, 147-171.

Booth, A., & Kee, H.J. (2006). Intergenerational transmission of fertility patterns in Britain. *IZA Discussion Paper, 2437*.

Chiappori, P.A. (1988). Rational Household Labor Supply. *Econometrica*, *56*, , 63-90.

Craig, L., & Mullan, K. (2010). Parenthood, gender and work-family time in USA, Australia, Italy, France and Denmark. *Journal of Marriage and the Family*, *72*, 1344 - 1361.

Elzinga, C.H. (2003). Sequence similarity: a nonaligning technique. *Sociological Methods and Research, 32*, 3-29.

Fine-Davis, M., Fagnani, J., Giovannini, D., Hojgaard, L., & Clarke, H. (2004). *Fathers and mothers. Dilemmas of the work-life balance, A comparative Study in Four European Countries*. Dordrecht, Boston and London: Kluwer.

Fouarge. D., Manzoni, A., Muffels, R., & Luijkx, R. (2010). Childbirth and cohort effects on mothers' labour supply: a comparative study using life history data for Germany, the Netherlands and Great Britain. *Work Employment and Society, 24*, 487-507.

Gabadinho, A., Ritschard, G., Müller, N.S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software, 40*, 1-37.

Gauthier, J.-A., Widmer, E.D., Bucher, P., & Notredame, C. (2010). Multichannel sequence analysis applied to social science data. *Sociological Methodology, 40*, 1-38.

Goffman, E. (1977). The Arrangement between the Sexes. *Theory and Society*, *4*, 301-331.

Gupta, S. (1999). The effects of transitions in marital status on men's performance of housework. *Journal of Marriage and the Family*, *61*, 700–711.

Gutiérrez-Domènech, M. (2005). Employment Transitions after Motherhood in Spain. *Labour*, *19*, 123–148.

Han, S.-K., & Moen, P. (1999). Clocking out: temporal patterning of retirement. *American Journal of Sociology, 105,* 191-236.

Hakim, C. (2003). Models of the Family in Modern Societies: Ideals and Realities. Aldershot: Ashgate.

Hynes, K., & Clarkberg, M., (2005). Women's Employment Patterns During Early Parenthood: A Group-Based Trajectory Analysis, *Journal of Marriage and Family, 67*, 222-240.

Jansen, M., & Liefbroer, A. (2006). Couples' Attitudes, Childbirth, & the Division of Labor. *Journal of Family Issues, 27*, 1487-1511.

Kaufman, G., & Uhlenberg, P. (2000). The Influence of Parenthood on Work Effort of Married Men and Women. *Social Forces, 78*, 931-947.

Kühhirt, M. (2012). Childbirth and the Long-Term Division of Labour within Couples: How do Substitution, Bargaining Power, and Norms affect Parents' Time Allocation in West Germany? *European Sociological Review, 28*, 565-582.
Lesnard, L. (2008). Off-Scheduling within Dual-Earner Couples: An Unequal and Negative Externality for Family Time. *American Journal of Sociology, 114,* 447-490.
Lesnard, L. (2010). Setting Cost in Optimal Matching to Uncover Contemporaneous Socio-Temporal Patterns. *Sociological Methods and Research, 38, 389-419*.
Macindoe, H., & Abbott, A. (2004). Sequence analysis and optimal matching techniques for social science data. In M. Hardy and A. Bryman (Eds.), *Handbook of Data Analysis* (pp. 387-406). London: Sage.
Matysiak, A. and Vignoli, D. (2009). Finding the "right moment" for the first baby to come: a comparison between Italy and Poland. *MPIDR Working Papers, WP-2009-011*.
Michaël R.T., & Tuma N.B. (1985,). Entry into Marriage and Parenthood by Young Adults. *Demography*, *22,* 515-544.
Pailhé, A.**,** & Solaz, A. (2006). Employment and childbearing: women bear the burden of the work-family balance. *Population and Societies*, *426.*
Pailhé, A., & Solaz, A. (2012). The influence of employment uncertainty on childbearing in France: A tempo or quantum effect? *Demographic Research*, *26*, 1-40.
Pittman, J.F., & Blanchard, D. (1996). The effects of work history and timing of marriage on the division of household labor: a life course perspective. *Journal of Marriage and Family, 58*, 78-90.
Pollock, G. (2007). Holistic trajectories: a study of combined employment, housing and family careers by using multiple-sequence analysis. *Journal of the Royal Statistical Society, 170,* 167-183.
Ravanera, Z.R., Rajulton, F. Burch, T.K., & Le Bourdais, C. (2002). The early life courses of Canadian men: Analysis of timing and sequences of events. *Canadian Studies in Population*, *29*, 293-312.
Regnier-Loillier, A. (2009). Does the birth of a child change the division of household tasks between partners? *Population & Societies, 461*.
Robette, N. (2010). The diversity of pathways to adulthood in France: evidence from a holistic approach. *Advances in life course research, 15*, 89-96.
Robette, N., & Thibault, N. (2008). Comparing qualitative harmonic analysis and optimal matching. An exploratory study of occupational trajectories. *Population-E, 64,* 533-556.
Schoon, I., McCullough, A., Joshi, H., Wiggins, R., & Bynner, J. (2001). Transitions from school to work in a changing social context. *Young, 9*, 4-22.
Shelton, B., & John, D. (1996). The division of household labour. *Annual Review of Sociology, 22*, 299-322.
Stovel, K., Savage, M., & Bearman, P. (1996). Ascription into achievement: models of career systems at Lloyds Bank, 1890-1970. *American Journal of Sociology, 102*, 358-399.
Van der Lippe, T. (1994). Spouses and their Division of Labour. *Kyklos, 30*, 43-62.
Vignoli, D., Drefahl, S., & De Santis, G. (2012). Whose job instability affects the likelihood of becoming a parent in Italy? A tale of two partners. *Demographic Research, 26*, 41-62.
Wu, L.L. (2000). Some comments on "Sequence analysis and optimal matching methods in sociology: Review and prospect". *Sociological Methods and Research, 29*, 41-64.

Appendix

Table 6. Main characteristics of people in and outside the sample, 1954-1968 cohort

	Sample: in partnership for 15 years and more, partner surveyed	In partnership for 15 years and more, partner not surveyed	Not in partnership, past partnership for 15 years and more	In partnership for less than 15 years	Always single	Total population
Age	42.7	44.3	44.4	39.9	41.4	41.9
Number of siblings	3.1	3.7	3.3	3.1	3.8	3.3
Number of children	2.2	2.3	2.4	1.8	0.2	2.0
No children	0.04	0.04	0.05	0.15	0.89	0.13
1 child	0.15	0.16	0.14	0.25	0.06	0.18
2 children	0.49	0.42	0.41	0.38	0.04	0.41
3+ children	0.33	0.38	0.40	0.22	0.02	0.28
Number of partnerships	1.1	1.2	1.2	1.5	0.0	1.2
Age at first partnership	22.5	21.2	21.0	25.4		21.8
Age at current partnership	22.8	23.5		30.0		24.8
Duration of current partnership	20.1	21.0		9.6		12.6
Age at end of studies	18.9	18.4	18.5	19.7	19.6	19.1
Unqualified	0.17	0.25	0.23	0.15	0.28	0.18
Vocational diploma	0.45	0.41	0.40	0.41	0.32	0.42
Secondary education	0.14	0.13	0.15	0.13	0.12	0.13
Tertiary education	0.23	0.18	0.20	0.31	0.27	0.26
Number of years of unemployment	0.6	0.8	0.9	0.9	1.2	0.8
Number of years of short-term jobs	0.9	1.1	1.3	1.4	1.6	1.2
Number of years of long-term jobs	17.5	15.9	16.0	14.7	15.2	16.0
Number of years out of the labour force	3.0	5.7	5.1	2.3	3.2	3.3
Number of jobs (> 6 months)	2.7	2.9	3.1	3.1	2.9	2.9
N	1882	889	1730	664	285	5450
%	34.5	16.3	31.7	12.2	5.2	100.0

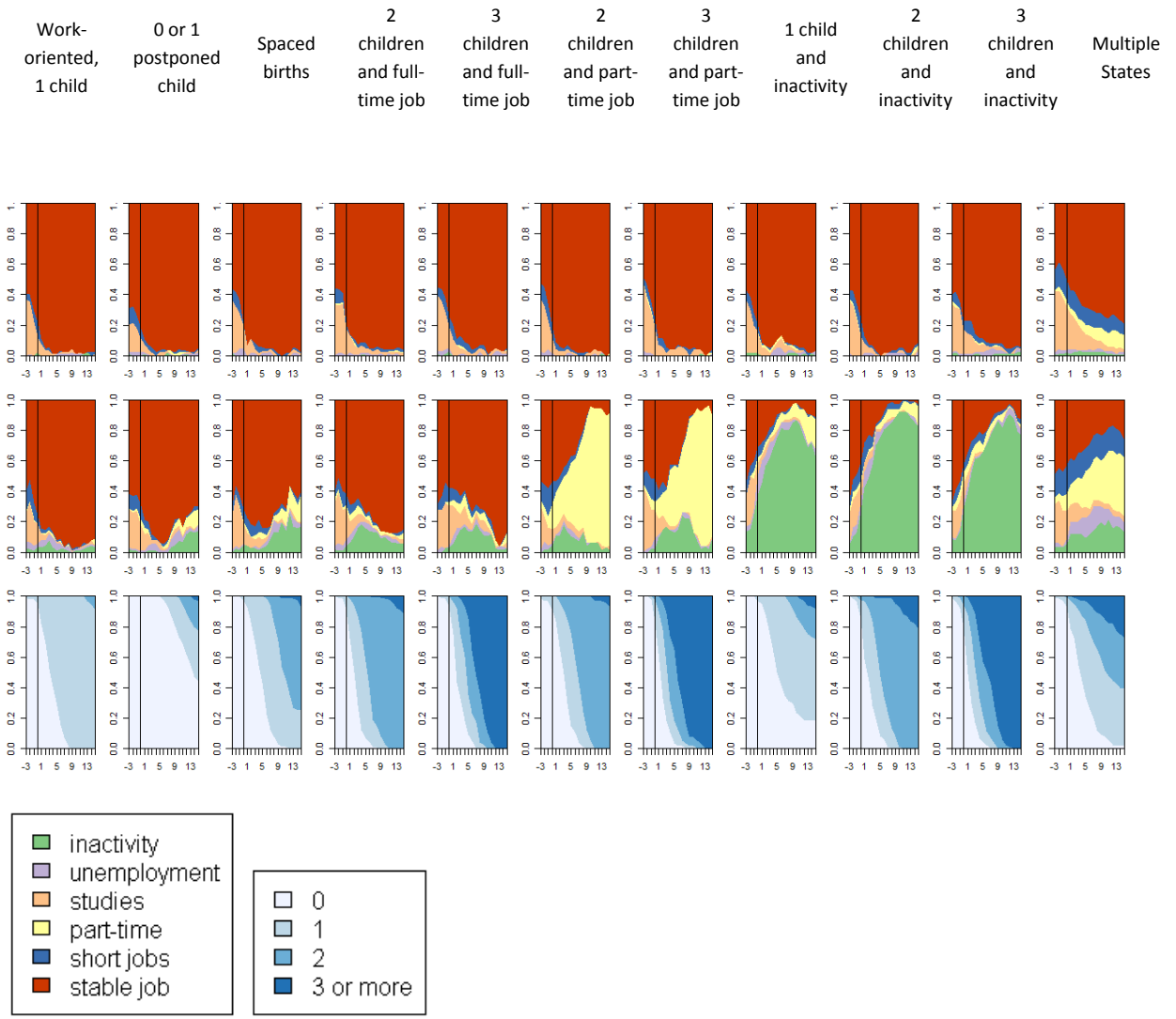
Table 7. Description of woman's and man's activity according to couples' career typology (duration in years)

	Cluster										
	Work-oriented, 1 child	0 or 1 postponed child	Spaced births	2 children and full-time job	3 children and full-time job	2 children and part-time job	3 children and part-time job	1 child and inactivity	2 children and inactivity	3 children and inactivity	Multiple States
MEN											
inactivity	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.2	0.3
unemployment	0.1	0.2	0.2	0.2	0.2	0.2	0.1	0.2	0.2	0.3	0.3
studies	1.2	0.8	1.3	1.6	1.5	1.1	1.6	1.3	1.2	1.6	2.6
part-time job	0.0	0.1	0.0	0.1	0.0	0.0	0.1	0.1	0.1	0.1	1.2
short job	0.3	0.6	0.4	0.5	0.7	0.6	0.4	0.4	0.5	0.6	2.0
stable job	16.3	16.4	16.0	15.7	15.6	16.1	15.8	15.8	16.0	15.2	11.6
WOMEN											
inactivity	0.5	0.8	1.6	1.6	1.4	1.2	1.8	11.1	11.5	11.0	2.2
unemployment	0.5	0.5	0.5	0.5	0.4	0.3	0.3	0.9	0.7	0.5	1.5
studies	0.9	1.3	1.3	1.4	1.9	1.0	1.5	1.0	0.9	0.8	1.7
part-time job	0.1	0.7	0.9	0.5	0.6	9.0	8.0	1.3	0.8	0.8	4.2
short job	0.6	0.7	0.8	0.6	0.8	1.0	0.7	0.6	0.9	0.7	2.8
stable job	15.5	14.1	12.9	13.4	12.9	5.5	5.7	3.2	3.2	4.2	5.5

Table 8. Frequency of labour market participation (at least one spell)

	Cluster										
	Work-oriented, 1 child	0 or 1 postponed child	Spaced births	2 children and full-time job	3 children and full-time job	2 children and part-time job	3 children and part-time job	1 child and inactivity	2 children and inactivity	3 children and inactivity	Multiple States
MEN											
inactivity	4.4	5.4	1.4	1.5	3.9	1.5	1.9	5.0	0.0	10.2	5.4
unemployment	3.3	9.7	12.7	10.5	10.5	11.9	5.6	10.0	13.1	17.0	16.9
studies	61.1	44.1	57.7	58.6	59.2	56.7	55.6	53.3	60.7	53.4	68.9
part-time job	0.0	2.2	0.0	3.8	2.6	1.5	5.6	3.3	6.6	4.5	14.2
short job	20.0	25.8	22.5	27.8	23.7	29.9	24.1	18.3	18.0	29.5	40.5
stable job	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	91.2
WOMEN											
inactivity	20.0	32.3	47.9	36.8	46.1	40.3	53.7	96.7	100.0	100.0	55.4
unemployment	21.1	22.6	23.9	21.8	23.7	16.4	20.4	41.7	29.5	18.2	37.2
studies	44.4	48.4	53.5	51.9	57.9	43.3	50.0	35.0	41.0	36.4	53.4
part-time job	8.9	25.8	36.6	22.6	27.6	100.0	100.0	33.3	31.1	23.9	60.1
short job	26.7	28.0	38.0	27.8	32.9	35.8	31.5	21.7	26.2	23.9	51.4
stable job	100.0	100.0	100.0	100.0	100.0	91.0	90.7	75.0	88.5	79.5	85.8

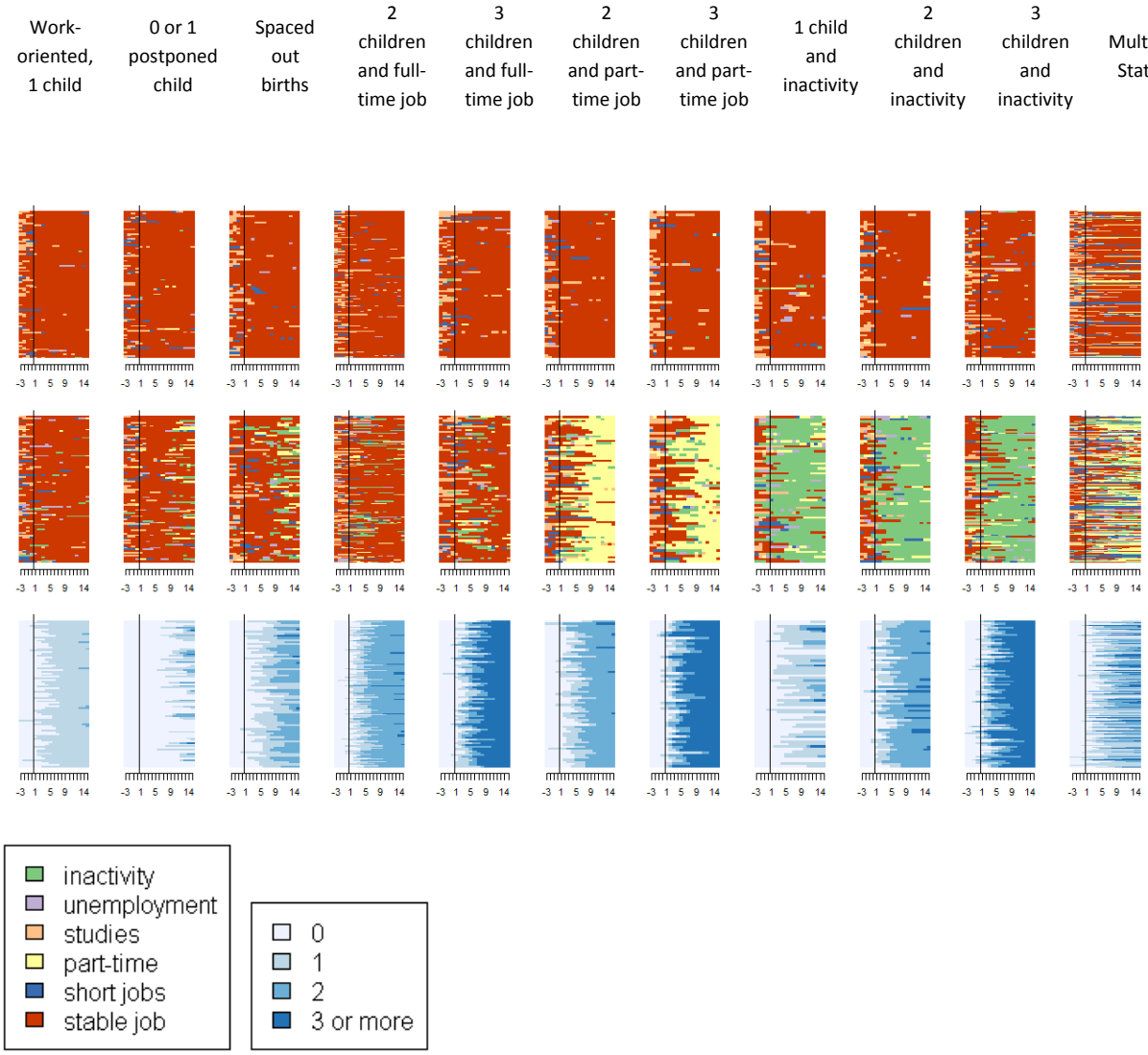
Figure 2. State distribution plots of the typology of couples' work and family histories



State distribution plots represent, for each cluster, each dimension of couples' life courses and any given year from 3 years before the entry into partnership to 15 years after it, the share of couples who belong to the various states. Upper panel plots stand for men's occupational careers, middle panel plots for women's careers and lower panel plots stand for couples' parental trajectories. Each cluster is represented by a column of 3 plots. Black vertical lines give the year of entry into partnership.

For instance, in the "work-oriented, 1 child" cluster (first column), 3 years before entering into partnership, about 35% of men (upper panel) are students and 60% have a stable job, while a few years later, almost all of them have a stable job. 98% of couples from this cluster are childless when they form a family, while they all have one child 8 years later.

Figure 3 - Index plots of the typology of couples' work and family histories



In index plots, each horizontal line stands for a couple's trajectory. Each coloured segment composing the line represents a spell in a state, the length of the segment corresponding to the spell duration. As with state distribution plots, upper panel plots stand for men's occupational careers, middle panel plots for women's careers and lower panel plots stand for couples' parental trajectories. Each cluster is represented by a column of 3 plots. Black vertical lines give the year of entry into partnership.

For instance, in the "3 children and part-time job" cluster (seventh column), women (middle panel) have in common to have worked part-time for a large part of the 18 years represented, although they started part-time at various periods and some of them have also had short spells of inactivity or short jobs. Almost every couple in this cluster had its first child very soon after the entry into partnership, rapidly followed by two more births.

Endnotes

[i] For instance, 97% of children aged three are enrolled at pre-school, even though schooling is not compulsory at that age.

[ii] There are in fact two surveys, one at the individual level and one at the employer's level, the employers being those of the individuals employed at the time of the survey. Here we only use information from the individual level data and but we intend to capitalize on the matched employer-employee data in a further step.

[iii] Because of the retrospective frame of the survey and the risk of memory bias, only information for these at-least-6-months partnerships were collected.

[iv] Including births of children who died

[v] First, some situations are not exclusive (for example, studies and unemployment); second, when a 6-month period started in year t and ended in year t+1, the interviewers were instructed to tick both years. In that case, we have divided the year by the number of situations identified and attributed to each situation a duration equal to the corresponding proportion of year.

[vi] In metropolitan France, the completed fertility rate of the 1957 birth cohort is 2.14 children. Among women born in 1950, 10% are childless, 20% have one child, 40% two children, 20% three children, and 10% have four or more children (Toulemon et al., 2008).

[vii] The duration of couple histories in our original sample varies greatly: from less than a year to 31 years with a median duration of 11 years.

[viii] Couples' histories are set to the same duration since optimal matching analysis remains relatively complex to implement with heterogeneity of lengths (Stovel et al., 1996).

[ix] The partner is not surveyed if they are out of the age range of the survey, not selected or not available to answer.

[x] Same-sex couples are excluded from the analysis because they were too few.

[xi] Marital status is not included in our sequences. The reason is that, while couples who get married may differ from those who do not, the time when they get married is not of foremost importance in the study of work-family balance as compared to timing of child births. Moreover, adding a marital dimension would excessively complicate sequence analysis, by notably enlarging the number of potential states in the sequences.

[xii] The women's and men's occupational dimensions have 6 states and the parental one has 4.

[xiii] Since inserting an element in one sequence is the same as deleting an element from the other, insertion and deletion have the same cost, which is called the *indel* cost.

[xiv] Other cost settings were tested, which led to quite convergent results.

[xv] The number of clusters is chosen according to a balance between parsimony on the one hand, and meaningfulness and homogeneity of clusters on the other. In this respect, summary statistics (as in Appendices 2 and 3) and graphical representations of clusters (Appendix 4) are helpful guidelines. Eleven is the minimal threshold to have distinct, homogeneous and substantively meaningful clusters. Divisions into fewer groups (say, 4 or 6) lead to one or two large

and very heterogeneous clusters (that contain more than half of couples) and a set of small homogeneous remaining clusters, which is not satisfying.

[xvi] As usual, in such classification methods, one cluster (here called “multiple states”) comprises all the remaining observations that do not fit into any other cluster.

[xvii] The median durations between couple formation and marriage are quite small. This may be due to the fact that, despite a clear definition of couple formation in the questionnaire, marriage may represent the actual time of couple formation for a large proportion of individuals.

[xviii] Five categories are formed: woman 5 years or more older, woman between 3 and 5 years older, partners of equal age within 2 years, man between 3 and 5 years older, man 5 years or more older. Couples do not differ much between clusters regarding the average age at union formation, which stands at around 22 for women and 23.5 for men.

[xix] The relationship can operate both ways, i.e. schooling choice may also be determined by priorities related to family or employment.

[xx] Generally, there is a correlation between mothers’ and mothers-in-law’s occupational status because of homogamy in the couple formation process. Almost half of women whose mother always worked match with a man whose mother worked continuously too, and the same tendency to partner similarity appears with inactive mothers.

Educational pathways and dropout from higher education in Germany

Sophie Müller, University of Bamberg, Germany
sophie.mueller@uni-bamberg.de
Thorsten Schneider, University of Leipzig, Germany

(Received June 2013 Revised August 2013) doi:10.14301/llcs.v4i3.251

Abstract

Extending access to higher education has led to a growing heterogeneity in the social origins and previous educational biographies of first-year students. They differ in their socialization, their preparedness for tertiary studies, and the salience of alternative options. How do these differences relate to social inequality in dropout from higher education? Drawing on theories and concepts of rational choice, differential learning environments, and selection, we argue that social origins and pre-tertiary educational pathways have at least an initial impact on dropout risks. We draw on retrospective life course data from the "Adult Education and Life-Long Learning" stage of the German National Educational Panel Study (NEPS) to use pre-tertiary pathways to reconstruct educational careers and perform an event history analysis on dropping out of higher education. Results suggest that these pathways substantially influence dropout rates in Germany. Students taking the direct pathway via the Gymnasium (i.e. the school type representing the highest school track) have significantly lower dropout rates than students with an upwardly mobile educational biography or students who obtained a vocational qualification before starting higher education. Whereas students from a higher social background are less prone to drop out than students from a lower social background at universities, social origins do not have any significant direct influence on dropout rates at universities of applied sciences and influence only the likelihood of entering these institutions.

Keywords: dropout, higher education, tertiary education, event history analysis, educational pathways, social origin

Introduction

Sociological studies exploring the mechanisms of dropout from tertiary education are rather scarce in Europe in general and Germany in particular. However, research in this field is important, because dropping out relates to issues of social inequality and influences how far graduation rates will increase in the population.

In times of accelerated technological change, modern societies respond to economic competition by increasing the proportion of higher education graduates. They pursue this goal by opening up access to higher education to a larger proportion of the population, and by fostering permeability from secondary and post-secondary education into tertiary education. However, due to their different qualifications and prerequisites, these newly-entitled groups might face higher obstacles to obtaining a higher education degree.

In this context, Germany is a particularly interesting case. Graduation rates at tertiary level

are far below the OECD average (OECD, 2012, p. 68) and have increased more slowly than in other countries due to institutional barriers (Powell & Solga, 2011). In addition, Germany has very high social inequality in access to tertiary education (Hillmert & Jacob, 2010). Therefore, providing new paths to obtain a higher education entrance certificate is an important step in raising tertiary completion rates. Some attempts have been made to provide new educational options which afford eligibility to higher education. However, the persistently high segregation of the German education system, particularly the "schism in education" (Baethge, 2010, p. 281) between vocational and general education, might render it particularly difficult for those who entered higher education via nonstandard pathways to complete it successfully.

In this article, we want to gain insights into the mechanisms of dropout from a life-course perspective using recent longitudinal nationwide data from Germany. The main research questions are: Do previous secondary or post-secondary educational pathways and social origins influence dropping out of higher education? Do these influences differ by type of higher education institution? And do these features become less important as the length of study time increases?

The German education system

The German education system is characterized by early tracking, high differentiation, and high stratification (Allmendinger, 1989). The lower secondary level differentiates between three tracks. In most federal states, each of these three tracks has been represented for many years by one specific school type: the lower track by the *Hauptschule*, the intermediate track by the *Realschule*, and the upper, academic track by the *Gymnasium*. The *Hauptschule* provides basic general education and usually covers Grades 5 to 9 (sometimes also including Grade 10). At the *Realschule*, students receive more extensive general education. This school type usually covers Grades 5 to 10. The *Gymnasium* aims to teach intensified general education and comprises Grades 5 to 12 or 13. The *Gymnasium* is the only one of the three school types offering students a final examination, known as *Abitur*, entitling them to enter higher education. Some federal states provide the lower and intermediate track within one school type; others also offer comprehensive schools containing all three tracks. However, comprehensive schools have not replaced the traditional *Gymnasium* in any single federal state.

High-achieving students from the *Haupt*- or *Realschule* who want to gain eligibility to higher education can transfer to a *Gymnasium* or to other schools offering senior classes (*gymnasiale Oberstufe*) at the upper secondary level. Some of these schools allow for part-time schooling, offer (either exclusively or additionally) restricted forms of the higher education entrance certificate, or also award vocational qualifications. A further possibility is to enter higher education without a higher education entrance certificate via a special examination (*Begabtenprüfung*), for talented or vocationally qualified persons. However, between 1995 and 2010, only about 1 per cent of students accessed higher education on this path (Autorengruppe Bildungsberichterstattung, 2012, p. 118, Tab. F1-4A).

There has been a considerable expansion of tertiary education in Germany in recent decades. For example, the proportion of 18- to 21-year-old students entering tertiary education increased from 17 per cent in 1980 to 44 per cent in 2009 (Statistisches Bundesamt, 2012). However, expansion in tertiary education has not been as extensive as that in secondary education (Lörz & Schindler, 2009, p. 98). Moreover, Germany has lower proportions of participation and graduation in tertiary education, and has experienced a smaller increase in tertiary enrolment and completion since the 1980s, than a number of other European countries (Arum, Gamoran, & Shavit, 2007; Powell & Solga, 2011; OECD, 2012, p. 68). Although inequality in completing secondary education decreased during the period of educational expansion, social inequality in post-secondary decisions increased (Mayer, Müller, & Pollak, 2007). Given the same level of achievement, students with *Abitur* from a higher social background more often access university directly, whereas students from the middle and lower classes more often opt for the less risky strategy of vocational training first and studies afterwards, or they even decide against higher education studies despite their eligibility (Becker & Hecken, 2009; Hillmert & Jacob, 2003; Mayer et al., 2007; Müller & Pollak, 2004; Reimer & Pollak, 2010). Those who have decided to take up higher education studies differ in their educational biography depending on their social origins. Higher

social classes follow the standard sequence from primary school to *Gymnasium* to university significantly more often (Jacob & Weiss, 2010).

The decision to enter higher education involves options in itself. The higher education system in Germany is binary (Mayer et al., 2007). On the one hand, there are the more practically oriented universities of applied sciences (*Fachhochschulen*) offering shorter and more structured study programmes of about 3.5 years. On the other hand, there are the more theory-oriented universities with longer study programmes of 5 to 6 years. Whereas students from a higher social background are more likely to enrol at a university, students from a lower social background are more likely to opt for a university of applied sciences. The degrees awarded by the different institutions are also connected to unequal chances. Graduates from the universities of applied sciences face more restrictions in accessing postgraduate studies, and their labour market outcomes are also less favourable (Müller, Brauns & Steinmann, 2002)[1].

To provide information on the prevalence and social selectivity of different educational pathways, we developed a basic classification of these pathways, which is also used later in our main analysis. This classification was based on first, whether students attended *Gymnasium* from the lower secondary level; second, whether they completed vocational education and training and third, whether they took up tertiary studies and, if so, whether they attended university (U) or university of applied sciences (UAS; see Table 1). If individuals completed vocational training and started tertiary studies, only vocational qualifications before tertiary studies are taken into account, while vocational qualifications acquired after tertiary studies are discounted.

Table 1. Types of educational pathways

Pathway	Gymnasium at lower secondary level	Completed vocational and educational training	Tertiary education at U = university, UAS = university of applied sciences
1a	x		U
1b	x		UAS
1c	x		-
2a	x	x	U
2b	x	x	UAS
2c	x	x	-
3a			U
3b			UAS
3c			-
4a		x	U
4b		x	UAS
4c		x	-

Table 2 gives an overview of the frequency of these pathways, both in total and by parental education, based on the weighted percentages. We distinguished four parental education groups according to the parents' highest educational qualification: (1) no qualification or lower school diploma with or without non-tertiary vocational qualification (CASMIN 1a, 1b, 1c), (2) intermediate school diploma with or without non-tertiary vocational qualification (CASMIN 2a, 2b), (3) tertiary education entrance diploma with or without vocational qualification (CASMIN 2c_voc, 2c_gen), and (4) degrees from universities or universities of applied sciences (CASMIN 3a, 3b). Pathway 1 comprises students who attend the *Gymnasium* at the lower secondary level (before Grade 10). Most of them enter university (10.2 per cent of all pathways); others enter a university of applied science (2.3 per cent) or do not enrol in tertiary education (2.0 per cent). Persons classified in

Pathway 2 also attend the *Gymnasium* at lower secondary level but additionally gain a vocational qualification. Some enter university afterwards (1.6 per cent); others, a university of applied sciences (1.4 per cent). However, the majority do not continue into tertiary education (9.0 per cent). Individuals classified in Pathway 3 do not attend the *Gymnasium* before Grade 10, but a lower or intermediate school. A minority switch to an upper secondary school offering a higher education entrance certificate and continue to university (1.9 per cent) or a university of applied sciences (1.1 per cent). The majority do not enter tertiary education (11.6 per cent). Pathway 4 comprises students who attend lower or intermediate schools at lower secondary level and obtain a vocational qualification. Some of them attain eligibility before, during, or after completing vocational education and training and enter university (1.0 per cent) or a university of applied sciences (2.9 per cent). Most of them do not enter tertiary studies, which is the most common of all pathways (55.1 per cent).

Table 2. Prevalence of educational pathways by education of parents

Pathway	No.[1]	%[1]	Total %[2]	CASMIN 1a, 1b, 1c %[2]	CASMIN 2a, 2b %[2]	CASMIN 2c_voc, 2c_gen %[2]	CASMIN 3a, 3b %[2]
1a	1,191	15.5	10.2	3.5	11.1	18.3	35.5
1b	284	3.7	2.3	1.2	2.6	3.4	6.6
1c	133	1.7	2	1.4	2.4	3.1	3.3
2a	180	2.3	1.6	0.6	2	2.3	4.9
2b	155	2	1.4	0.7	1.7	2.1	3.6
2c	783	10.2	9	7	12.5	12.4	11.6
3a	190	2.5	1.9	1.2	2.8	3	3.4
3b	115	1.5	1.1	0.9	1.3	1.2	1.6
3c	554	7.2	11.6	14.4	9.1	9.9	3.1
4a	126	1.6	1	0.9	1.1	1.8	1.1
4b	301	3.9	2.9	2.6	3	4.5	2.9
4c	3,657	47.7	55.1	65.6	50.3	38	22.3
Total (No.)	7,669	7,669	7,669	4,664	1,451	549	1,005
Total (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Notes. Source: NEPS: SC6: 1.0.0; own calculations.
[1]*Unweighted.*
[2]*Standardized weights based on Mikrozensus 2009.*

The distribution over pathways depends strongly on the parents' education: the higher the latter, the more likely are pathways in which students attend the *Gymnasium* at lower secondary level (1a–2c). The strongest differences by parental education prevail in Pathway 1a. Whereas 35.5 per cent of students whose parents have tertiary degrees attend the *Gymnasium* before Grade 10 and enrol in tertiary studies later without completing any vocational training before (Pathway 1a), only 3.5 per cent of those whose parents have no more than basic vocational training above and beyond compulsory schooling do so. By contrast, the share of students entering tertiary education without having attended the *Gymnasium* at lower secondary level (Pathways 3a, 3b, 4a, 4b) is distributed more equally with regard to parental education. The percentage of students not attending the *Gymnasium* and not entering tertiary education (Pathways 3c and 4c), however, is much higher if parents are less educated.

Theoretical considerations and hypotheses

This section considers explanations for why and when students drop out. First, we discuss mechanisms linking pre-tertiary pathways and dropout from higher education. Then, we explain how social origin might relate to both (pre-tertiary pathways and dropout). Finally, we reflect on the role of time.

Mechanisms linking pre-tertiary pathways and dropout from higher education

Previous research has shown that students at the same level of initial competencies develop differently according to which track they are assigned to (Hanushek & Wössmann, 2006; Maaz, Trautwein, Lüdtke, & Baumert, 2008; Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006). Due to institutional (e.g. curricula, timetables, education of teachers) and compositional differences (distribution of achievement and social background), progress in competencies is higher at the *Gymnasium* (Maaz et al., 2008). Although the majority of tertiary students, including those who have previously attended a lower or intermediate track, attained eligibility in senior classes of the highest track (*Gymnasiale Oberstufe*), there might be negative long-term effects. First, it is unlikely that three years of senior classes overwrite the effects of about five years of tracking at lower secondary level. Second, tertiary students who have attended lower or intermediate tracks at lower secondary level are more likely to join senior classes that are more heterogeneous in terms of achievement and social background (see Maaz, Chang, & Köller, 2004). Beyond effects on competencies, tracking also gives the opportunity to meet students with certain characteristics, and this may shape aspirations (Hallinan & Williams, 1990). Students on upwardly mobile paths are more likely to become acquainted with persons not entering tertiary education. Values shared in these friendships might be opposed to higher educational aspirations and might render integration into tertiary education more difficult (Tinto, 1975). The example given by friends might make educational or occupational alternatives appear more desirable than continuing tertiary education.

Tracking was found to have an effect on students' self-concepts. On the one hand, students in schools with lower average achievement develop higher academic self-concepts (big-fish-little-pond effect (BFLP), Marsh & Parker, 1984; for evidence on the German education system, see e.g. Trautwein et al., 2006). On the other hand, students in selective schools might develop higher academic self-concepts because they attribute the positive collective identity to themselves (basking-in-the-reflected-glory effect (BIRG), Cialdini et al., 1976; for evidence on the German education system, see e.g. Knigge & Hannover, 2011). However, the latter seems to be much weaker than the BFLP effect (Marsh, Kong, & Hau, 2000; Trautwein et al., 2006). Having formed a high level of academic self-concept at school, students who enter tertiary education via upwardly mobile pathways might experience substantial losses of academic self-concept when they arrive at an institution in which the average performance level is higher.

Drawing on differential learning environments and their effects on competencies, social relations, and self-concepts, we therefore expect that students who reach higher education from outside the standard academic track of the *Gymnasium,* will have higher dropout risks (Hypothesis 1).

In Germany, vocational education and tertiary education are particularly strongly divided within the general education system and have developed along separate paths for the last 200 years (Baethge, 2010, p. 281). Vocational and general education, therefore, differ in institutional characteristics. For example, the main goal in vocational education is to impart competencies related to an occupational field, whereas in general education, it is to develop the personality and to promote autonomy. Moreover, whereas vocational education is organized as a combination of practical work on the job and learning that is related to practical tasks at school, general education is mainly theoretical (Baethge, 2010, pp. 278–280).

This strong division between non-tertiary vocational education plus training (VET) and tertiary education in the German education system, renders the transition from VET to tertiary education particularly difficult (Powell & Solga, 2011, p. 54). Students with vocational qualifications who enrol in tertiary education might be frustrated due to the lack of recognition of prior qualifications, and they might find it difficult to adapt to the stronger theoretical orientation.

Following a reformulation of the rational-choice model on educational decisions, the key

determinants of the decision between tertiary and vocational education and the labour market can be considered to be the subjective probability of success, the expected income, and the time horizon (i.e. the time span taken into account for calculating the returns; see Hillmert & Jacob, 2003). If these determinants explain the decision to enter tertiary education after obtaining a vocational qualification, any changes occurring in them might lead to a revision of this decision and to a dropout from tertiary studies. Empirical evidence shows that students adjust their perceptions of the probability that they will succeed in tertiary education after receiving feedback through, for example, grades (Stinebrickner & Stinebrickner, 2012). Because students with vocational qualifications who drop out can anticipate the income of a skilled rather than an unskilled employee, their expected income gains through tertiary education will be lower. As a result, they might be more sensitive to reductions in their perceived probability of success. Due to the difficult transition from non-tertiary vocational education to tertiary education and the lower additional benefits of graduation, we therefore expect higher dropout risks for vocationally qualified students than for students without prior vocational qualification (Hypothesis 2).

Being the most direct pathway to tertiary education, the standard academic pathway via the Gymnasium is usually also the shortest in duration. Students who additionally obtain a vocational qualification before or after acquiring eligibility for tertiary studies, are therefore older when they enter tertiary studies. Older students might be under stronger time constraints because they are more likely to have children or to work longer hours in a job (DesJardins, Ahlburg, & McCall, 1999, p. 385; Smith & Naylor, 2000). We therefore expect older students to be more prone to drop out, and that age at enrolment will explain part of the higher dropout rates of pathways other than the *Gymnasium,* especially those including vocational qualifications (Hypothesis 3).

Universities have always been the central institutions of the German tertiary education system. Universities of applied sciences emerged from the non-tertiary vocational system and only later became recognised as tertiary institutions. In line with their origins, they are more practically oriented than academic universities (Mayer et al., 2007, p. 244). Thus, the transfer from more vocationally oriented schools or vocational training to a university of applied sciences, might be smoother than that to a university. Moreover, the prior pathways leading to universities of applied sciences are more heterogeneous (see Scheller, Isleib, & Sommer, 2013 and the descriptive section in this paper, esp. Table 3). Thus, students who arrive at a university of applied sciences via nonstandard pathways meet peers who share their school experience. This might ease integration. Consequently, we expect the effects of pre-tertiary pathways to be less pronounced at a university of applied sciences than at an academic university (Hypothesis 4).

The role of social origins

As mentioned above, educational pathways before higher education are class-specific (Table 2; see also Becker & Hecken, 2009; Jacob & Weiss, 2010; Reimer & Pollak, 2010): students from higher social classes have taken the standard path more often than students from lower classes. If pathways outside the academic track produce higher dropout rates, this should result in a social inequality of dropout (Hypothesis 5).

According to Tinto (1975), successful integration into the tertiary social and academic environment helps prevent students from dropout. Thus, students who engage more strongly in extracurricular activities and manage to establish a social network or those who find it easier to comply with the academic requirements based on the values and norms of the university system, are considered to be at less risk of quitting tertiary studies before graduation. Similar arguments can be found in studies following the tradition of Bourdieu (Reay, Davies, David & Ball, 2001; Thomas & Quinn, 2006) that highlight the rifts between the habitus of origin and the institutional habitus, and show how difficult they are to overcome for students from a lower social background. Hansen and Mastekaasa (2006), for example, found that after controlling for the secondary school grade-point average, students at Norwegian universities gain higher grades when they have higher levels of cultural capital.

According to the principle of relative risk aversion (Breen & Goldthorpe, 1997), students assign priority to the goal of avoiding downward social class mobility when taking educational decisions. Students whose parents have a higher education degree, risk a status loss by dropping out

of higher education. Hence, dropout involves higher costs for them compared to students whose parents have no tertiary degrees. Because students with more highly educated parents have access to a higher level of resources, find it easier to integrate into the college environment, and have higher costs of status demotion in case of dropping out, we expect them to have lower dropout rates than students with less well educated parents (Hypothesis 6a).

On the other hand, due to the highly socially selective German education system, a much smaller proportion of students from lower social origins than from higher social origins reach higher education. As a result, tertiary students from lower social origins might be a positive selection in terms of the qualities that help them to succeed in the educational system (Mare, 1980). These qualities might compensate for disadvantages and cancel out the effect of social origin (Hypothesis 6b).

The role of time

The literature on dropout emphasizes that it should be seen as a process in time (e.g. Chen, 2012; DesJardins, 2003; Tinto, 1988). The transition into tertiary studies can be considered as a succession of different stages (Tinto, 1988) that each have their own specific vulnerability to dropout. In addition, the determinants influencing dropout change over time and vary in the strength of their effects (DesJardins et al., 1999).

Differences in the starting conditions based on different educational histories should converge with the duration of a common history of studies under equal conditions. Students from upwardly mobile pathways who may well be less prepared for higher education might still catch up. Furthermore, newly acquired knowledge should become more important than that acquired before higher education. Consequently, the advantage of students who have taken the standard direct path should decrease over time. Finally, the lack of familiarity with the higher education culture might be overcome by making contact and exchanging ideas and information with other students. Indeed, Pascarella, Pierson, Wolniak, and Terenzini (2004) found that students with lower levels of social and cultural capital improve more strongly in educational outcomes through social interactions and academic activities, than students with higher levels of social and cultural capital. Thus, interactions with other students and academic effort seem to have compensatory potential. Regarding the timing of dropout among students who have already gained vocational qualifications, the driving force might well be cost–benefit calculations. According to human capital theory (e.g. Becker, 1964), the decision to invest in education depends not only on expected returns but also on opportunity costs. The latter are higher for students with vocational qualifications, because their qualification raises the costs of foregone earnings. This should result in an early dropout of students with prior vocational qualifications.

In summary, we expect educational pathways before higher education to lose their impact on dropout rates over the duration of higher education, because socialization and newly gained knowledge in the tertiary environment become more important than knowledge learned at school (Hypothesis 7). We also expect that the effect of a vocational qualification will diminish over study duration because the value of the prior qualification declines over time (Hypothesis 8).

Data and methods

We tested our hypotheses on data from the National Educational Panel Study (NEPS): Starting Cohort 6 – Adults (Adult Education and Lifelong Learning), doi:10.5157/NEPS:SC6:1.0.0. The NEPS data collection is part of the Framework Programme for the Promotion of Empirical Educational Research, funded by the German Federal Ministry of Education and Research and supported by the Federal States. The study population consists of 11,649 individuals born between 1944 and 1986 who were surveyed for either the first or second time in 2009/2010. Because the entire education and labour market history is surveyed on the basis of spells, i.e. the time spent in a specific state, we can identify starting points, durations, and successful completions of tertiary studies.

We doubt that our arguments apply to study episodes in East Germany before reunification, because access to higher education, study programmes, and the labour market situation differed strongly from conditions in West Germany (see Solga, 1997). Therefore, we dropped school and study episodes in the NEPS sample that were found to have taken place in the German Democratic Republic (GDR). This information is not available for those interviewed for the second time

in 2009/2010. Thus, respondents of this subsample were disregarded if they were born in East Germany and started secondary or tertiary education before the restructuring of the educational system in 1990.

Because our study is restricted to the German education system, we did not consider individuals who had migrated to Germany after the age of 7 years, had spent more than a year abroad during their school career, or had studied in a tertiary institution abroad.

As only 43 respondents (1.6 per cent of the analysed sample) had missing values on at least one variable of interest, we decided to exclude them listwise from the analysis. Additionally, we excluded 19 cases with extreme values (greater than 40) in the variable age at enrolment. For reasons of comparability over time, we did not take into account student entry cohorts before 1970. Although the institutional setting of higher education did not change dramatically afterwards, the recognition of universities of applied sciences as higher education institutions was a major change in the late 1960s.

We restricted our analyses to study episodes at universities and universities of applied sciences (both ISCED-level 5A). We excluded ISCED-level 5B institutions such as universities of cooperative education (*Berufsakademien*), business academies (*Wirtschaftsakademien*), or academies of public administration (*Verwaltungsakademien*) that are not yet recognized comprehensively as university-level institutions by federal law.

We analysed dropout from first tertiary studies as the dependent variable, defined as leaving first tertiary studies without a degree and not attending higher education again within the following 12 months. This definition ensures that we avoid considering changes in the field of studies as dropouts.

For the purpose of the present analysis, we distinguished four types of educational pathway after primary school and before tertiary entrance (see Pathways 1a, 1b, 2a, 2b, 3a, 3b, 4a, and 4b in Table 1, and the respective description in the text). We were not interested in Pathways 1c, 2c, 3c, and 4c, because students in these pathways do not enter tertiary education. Social origin was measured by the parents' highest educational qualification. As mentioned above, we distinguished four parental education groups: (1) no qualification or lower school diploma with or without non-tertiary vocational qualification (CASMIN 1a, 1b, 1c), (2) intermediate school diploma with or without non-tertiary vocational qualification (CASMIN 2a, 2b), (3) tertiary education entrance diploma with or without vocational qualification (CASMIN 2c_voc, 2c_gen), and (4) degrees from universities or universities of applied sciences (CASMIN 3a, 3b). Here again, we did not consider degrees from universities of cooperative education (*Berufsakademien*), business academies (*Wirtschaftsakademien*), or academies of public administration (*Verwaltungsakademien*) as higher education degrees.

We also took into account the respondent's age at entry into higher education derived from the respondent's birth year and the spell data on education.

Finally, all multivariate models include control variables. These are gender, first-year student cohorts, the type of higher education entrance certificate, changing between the two higher education institutions as a time-dependent variable, and the grade-point average of the highest school certificate. The latter variable was used in an attempt to control for primary effects, the relationship between socio-economic background and performance, as well as the relationship between school performance and educational pathway. In order to achieve better comparability over time, we standardized the grade-point average by the year of graduation. Additionally, we reversed the standardized values so that higher values indicate better grades. Although the type of higher education entrance certificate is related to the pathway, there is some variation within the pathways with regard to the type of the certificate. We also did not consider collinearity to be a problem. However, we were concerned that grading might differ between school types or, more precisely, that at the same level of competencies, students at other schools than the *Gymnasium* might obtain better grade-point averages. In this case, the predictive power of the grade-point average on dropout should be stronger for students who obtain their higher education entrance certificate at the *Gymnasium* than for students who obtain it at a vocational or other school. However, interaction effects between the school type awarding the entrance diploma and the grade-point average were not significant. This convinced us that the quality of the grade-point average as a proxy for primary effects does not differ strongly between school types. For a distribution of all covariates in both samples at different points in time, see Tables A1 and A2 in the Appendix.

We tested our hypotheses with event history analysis in order to take the time dependency of dropout into account. In a first step, we estimated the following functions indicating the proportion of students who have dropped out (H^a) or completed (H^b) tertiary education until a certain point in time (see Rohwer, 2006):

(1) $H^a(t) \coloneqq \sum_{k=0}^{t-1} r^a(k)G(k)$

(2) $H^b(t) \coloneqq \sum_{k=0}^{t-1} r^b(k)G(k)$

in which $r^a(k)$ represents the transition rate of dropout, $r^b(k)$ the transition rate of completion, and $G(k)$ the global survivor function of dropout and completion – that is, the probability of still persisting in tertiary studies at a certain point in time. We computed the survivor function with the Kaplan–Meier method.

For the multivariate analyses, we estimated transition rate models. The underlying dependent process is the monthly duration in first tertiary studies ending with an event for dropouts. The transition rate is the intensity of experiencing an event under the condition of not having experienced the event before (Blossfeld, Golsch & Rohwer, 2007). This is given in the following Equation 3:

(3) $r(t) = \lim_{t'-t} \frac{P(t \leq T < t' | T \geq t)}{t'-t}$

in which *T* is defined as a continuous random variable indicating the duration in tertiary education. Possible destination states are dropout, receiving a degree, or still being in tertiary education at the time of the interview. Whereas the last case is inherently right-censored, we also treated graduating as right-censored. We modelled the duration dependence with a piecewise constant exponential model that assumes the baseline hazard rate to be constant within time periods but allows it to vary between them (Blossfeld et al., 2007). A hazard rate for each predefined interval (l) was computed with the following equation, based on the coefficient for the baseline hazard of the interval (β_{0l}) and the regression coefficients of the covariates ($\beta_1 \ldots \beta_k$). Thus, it is assumed that covariates shift the rate proportionally up or down.

(4) $r(t_l) = \exp\{\beta_{0l} + \beta_1 x_1 + \ldots + \beta_k x_k\}$

In order to test whether the effects of covariates vary over time, we also estimated period-specific effects that can vary between the time intervals ($\beta_{1l} \ldots \beta_{kl}$):

(5) $r(t_l) = \exp\{\beta_{0l} + \beta_{1l} x_1 + \ldots + \beta_{kl} x_k\}$

At the time when large proportions of students have already graduated, the remaining population at risk might be highly selective, and this selectiveness might not be independent of studying successfully. We introduced right censoring into the multivariate analysis for cases that have no event during the first 8 years after starting tertiary studies. At this point in time, the risk set comprises 10 per cent of the original risk set at university and 2 per cent of the original risk set at universities of applied sciences. We ran the models separately for academic universities and universities of applied sciences because these two types of higher education institutions differ in important aspects: the composition of students regarding social origins, the composition of students regarding previously taken pathways, the length of study programmes, and the proportion of vocational versus academic orientation.

Results

In this section, we first present some descriptive results on the prevalence of pre-tertiary educational pathways and on the distribution of dropout and graduation over time, before moving on to the multivariate analyses of dropout.

Descriptive results

Table 3 shows the distribution of students at university and university of applied sciences over the four types of educational pathway. The weighted percentages give an impression of the prevalence of each pathway in the population. The distribution differs largely between the two types of higher education institution. The majority of first-year students (about 70 per cent) reach university via the standard pathway. At universities of applied sciences, however, the picture is different: here, less than every third student has taken the standard pathway. In comparison to academic university students, more students at the universities of applied sciences have obtained a vocational qualification before entering higher education. The largest group of these students reaches a university of applied sciences via an upwardly mobile pathway combined with non-tertiary vocational training (about 38 per cent). Also, the combination of the standard path-way with a vocational qualification is more frequent at universities of applied sciences (18 per cent) than at academic universities (about 11 per cent).

Table 3. Prevalence of pre-tertiary educational pathways among students at universities and universities of applied sciences

	Universities			Universities of applied sciences		
Types of pre-tertiary pathways	**Freq.[1]**	**%[1]**	**%[2]**	**Freq.[1]**	**%[1]**	**%[2]**
(1) Standard pathway	1,203	70.6	69.2	287	33.3	30.3
(2) Standard pathway + voc. qual.	183	10.7	10.8	156	18.1	18.0
(3) Upwardly mobile pathway	190	11.2	12.9	116	13.4	14.0
(4) Upw. mobile pathway + voc. qual.	128	7.5	7.1	304	35.2	37.7
Total	1,704	100	100	863	100	100

Notes. Source: NEPS: SC6: 1.0.0; own calculations.
[1]Unweighted.
[2]Standardized weights based on Mikrozensus 2009.

The functions in Figure 1 and 2 indicate the proportions of students who leave their university or university of applied sciences within 10 years after starting tertiary studies, because they either drop out (solid line) or graduate (dashed line). Because study programmes take longer at academic universities, students there graduate later than those at universities of applied sciences. In the first 3 years, graduations are extremely rare, but more than half of the dropouts have already taken place.

Figure 1. Proportion functions of dropout and graduation from higher education at universities, based on survivor functions (Kaplan–Meier method)

Notes. Source: NEPS SC6; own calculations.
Number of events: 320 dropouts; 1,204 graduations.

Figure 2. Proportion functions of dropout and graduation from higher education at universities of applied sciences, based on survivor functions (Kaplan–Meier method)

Notes. Source: NEPS SC6; own calculations.
Number of events: 116 dropouts; 705 graduations.

Multivariate results

We shall start with the findings from the multivariate analysis of university students. We estimated a piece-wise-constant exponential model. The hazard of dropping out does not change over time (see Table 4, Model 1). The risk seems to be time-constant. Regarding social origin, results suggest that students with at least one parent holding a tertiary degree drop out less often from university (see Table 4, Model 1). However, to some extent, this seems to be due to higher academic performance: if the grade-point average of the highest school certificate is controlled (as in Model 2), the effect of social origin decreases, although it is still statistically significant on a 10 per cent level. If, additionally, pre-tertiary pathways are taken into account (see Model 3), the effect of parents with tertiary education declines even more and ceases to be significant[2]. This is in line with the hypothesis that educational pathways explain social inequality in dropout rates to a notable degree (Hypothesis 5). As the remaining effects of parental education are not significant, Hypothesis 6a has to be rejected, whereas findings favour the conflicting Hypothesis 6b: students from a higher social background do not seem to be prevented from dropping out because they can rely on the resources of their parents, they integrate more easily at university, or they have higher dropout costs in terms of status demotion.

Table 4. Piecewise constant exponential models on dropout from universities within 8 years of starting tertiary studies

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Year 1	-10.53*	-13.56**	-11.46*	-11.64*	-11.55*	-11.72*
	[4.89]	[4.88]	[4.89]	[4.87]	[4.89]	[4.87]
Year 2	-10.23*	-13.25**	-11.15*	-11.33*	-11.18*	-11.36*
	[4.89]	[4.88]	[4.89]	[4.87]	[4.89]	[4.87]
Year 3–4	-10.99*	-13.99**	-11.88*	-12.06*	-11.94*	-12.10*
	[4.89]	[4.89]	[4.90]	[4.88]	[4.89]	[4.88]
Year 5–8	-10.76*	-13.77**	-11.65*	-11.82*	-11.67*	-11.83*
	[4.90]	[4.90]	[4.91]	[4.89]	[4.91]	[4.89]
Parents: CASMIN 1a/1b/1c	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
Parents: CASMIN 2a/2b	-0.22	-0.16	-0.09	-0.07	-0.09	-0.07
	[0.17]	[0.17]	[0.17]	[0.17]	[0.17]	[0.17]
Parents: CASMIN 2c	0.07	0.16	0.24	0.23	0.24	0.23
	[0.19]	[0.19]	[0.19]	[0.19]	[0.19]	[0.19]
Parents: CASMIN 3a/3b	-0.42**	-0.26^{+}	-0.14	-0.12	-0.13	-0.12
	[0.15]	[0.15]	[0.16]	[0.15]	[0.16]	[0.15]
GPA of highest school certificate		-0.67**	-0.66**	-0.64**		
		[0.08]	[0.08]	[0.08]		
in year 1					-0.70**	-0.68**
					[0.18]	[0.18]
in year 2					-0.49**	-0.47**
					[0.16]	[0.16]
in year 3–4					-0.81**	-0.78**
					[0.17]	[0.17]
in year 5–8					-0.70**	-0.66**
					[0.16]	[0.15]
Age at enrolment				0.09**		0.09**
				[0.02]		[0.02]
Pathway 1			*ref.*	*ref.*	*ref.*	*ref.*
Pathway 2			0.50**	0.15		
			[0.18]	[0.21]		
in year 1–2					0.67**	0.32
					[0.23]	[0.26]
in year 3–8					0.30	-0.05
					[0.27]	[0.29]
Pathway 3			0.45*	0.40*		
			[0.18]	[0.18]		
in year 1–2					0.64**	0.59*
					[0.24]	[0.24]
in year 3–8					0.23	0.19
					[0.27]	[0.27]
Pathway 4			0.64**	0.18		
			[0.21]	[0.25]		
in year 1–2					0.40	-0.06
					[0.31]	[0.34]
in year 3–8					0.85**	0.39
					[0.27]	[0.29]

*Notes. Source: NEPS SC6; own calculations. Standard errors in brackets. $^{+}$ p < 0.10, * p < 0.05, ** p < 0.01.*
Number of persons: 1,704; number of events (dropouts): 289. Controls (not reported in the table for the sake of clarity, see Table A3 in the Appendix for complete results): gender, entry year of tertiary studies (linear and squared), changing from university to a university of applied sciences, type of higher education entrance certificate. Due to low numbers of events per pathway in each single time period, Models 5 and 6 include constraints for educational pathways indicating that the effect for the first year equals the effect for the second year, and the effect for the third and fourth year equals the effect for the fifth to eighth year.

Model 3 indicates that pre-tertiary pathways seem to matter: students who came to university via the direct standard pathway have significantly lower dropout rates than students with a more complex educational history. This result favours the hypothesis that the learning environment of the *Gymnasium* eases integration and endurance at university (Hypothesis 1). Pathways with vocational training seem to induce the highest dropout rates. This is consistent with the hypothesis that these students are distracted from higher education more strongly, because of better alternative labour market opportunities and lower relative benefits from higher education, than students without additional vocational qualifications (Hypothesis 2).

In Model 4, we find that older students are more prone to drop out and that this explains higher dropout rates of students who obtained a vocational qualification before tertiary studies as expected in Hypothesis 3. However, pre-tertiary pathways and age at enrolment correlate highly. For example, 75 per cent of university students on Pathways 1 or 3 are respectively 21 years and 1 month and 21 years and 6 months old or younger, whereas 75 per cent of university students on Pathways 2 or 4 are respectively at least 24 years and 10 months and 27 years and 6 months old.

In Models 5 and 6, we introduced time-varying effects in order to test how the influence of the covariates on dropout develops over the duration of tertiary studies – in Model 5 without, and in Model 6 with, controlling for age at enrolment. We find results at the university to be consistent with the hypothesis that pre-tertiary pathways matter more strongly in the beginning and lose their impact later, which might be due to newly acquired knowledge and networks at university (Hypothesis 7). This seems to be the case for at least the standard pathway combined with vocational training and the upwardly mobile pathway without additional vocational qualifications. Students who have taken these pathways before entering university have higher dropout rates in the first 2 years than students who have followed the standard pathway, whereas these effects are less pronounced and not statistically significant afterwards. This is not the case, however, for students who have combined an upwardly mobile pathway and a vocational qualification. These students are more prone to a late dropout after the first 2 years. The effects on early dropout of students in Pathway 2 and the late dropout of students in Pathway 4 are no longer significant when age at enrolment is taken into account (Model 6).

Results for the universities of applied sciences differ from those from academic universities in four major points (see Table 5). First, social origins do not impact on dropout rates, so Hypothesis 5 has to be rejected. Second, vocational qualifications before tertiary education do not increase dropout rates significantly. This supports the hypothesis that the universities of applied sciences are more successful in retaining students with vocational qualifications because of their more practical orientation (Hypothesis 4). Third, older students do not differ significantly from younger students in their dropout risk. Fourth, effects of pre-tertiary pathways do not vary significantly over time. It has to be taken in account, however, that the sample and the total number of dropouts are smaller at universities of applied sciences, and thus the probability of significant effects is lower[3].

Table 5. Piecewise constant exponential models on dropout from universities of applied sciences within 8 years of starting tertiary studies

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Year 1	-8.51	-12.64^{+}	-14.53^{+}	-14.62^{+}	-13.77^{+}	-13.84^{+}
	[7.22]	[7.37]	[7.45]	[7.48]	[7.48]	[7.50]
Year 2	-8.74	-12.85^{+}	-14.73*	-14.82*	-13.98^{+}	-14.05^{+}
	[7.22]	[7.37]	[7.45]	[7.48]	[7.48]	[7.50]
Year 3–4	-9.11	-13.21^{+}	-15.07*	-15.17*	-14.52^{+}	-14.59^{+}
	[7.23]	[7.38]	[7.46]	[7.49]	[7.49]	[7.52]
Year 5–8	-8.40	-12.56^{+}	-14.45^{+}	-14.54^{+}	-13.78^{+}	-13.84^{+}
	[7.25]	[7.40]	[7.47]	[7.50]	[7.51]	[7.53]
Parents: CASMIN 1a/1b/1c	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
Parents: CASMIN 2a/2b	-0.23	-0.19	-0.17	-0.18	-0.18	-0.19
	[0.27]	[0.27]	[0.27]	[0.27]	[0.27]	[0.27]
Parents: CASMIN 2c	0.23	0.23	0.33	0.33	0.33	0.33
	[0.32]	[0.33]	[0.33]	[0.33]	[0.33]	[0.33]
Parents: CASMIN 3a/2b	0.08	0.01	0.07	0.07	0.07	0.07
	[0.25]	[0.24]	[0.25]	[0.25]	[0.25]	[0.25]
GPA of highest school certificate		-0.89**	-0.90**	-0.90**		
		[0.15]	[0.15]	[0.15]		
in year 1					-1.01**	-1.01**
					[0.27]	[0.27]
in year 2					-1.05**	-1.05**
					[0.31]	[0.31]
in year 3–4					-0.90**	-0.90**
					[0.30]	[0.30]
in year 5–8					-0.53	-0.53
					[0.33]	[0.34]
Age at enrolment				-0.01		-0.00
				[0.03]		[0.03]
Pathway 1			*ref.*	*ref.*	*ref.*	*ref.*
Pathway 2			0.14	0.15		
			[0.30]	[0.32]		
in year 1–2					-0.05	-0.03
					[0.39]	[0.40]
in year 3–8					0.36	0.37
					[0.46]	[0.47]
Pathway 3			0.73*	0.73*		
			[0.29]	[0.29]		
in year 1–2					0.72*	0.72*
					[0.36]	[0.36]
in year 3–8					0.69	0.69
					[0.46]	[0.46]
Pathway 4			0.01	0.02		
			[0.29]	[0.30]		
in year 1–2					-0.23	-0.22
					[0.36]	[0.37]
in year 3–8					0.30	0.31
					[0.40]	[0.41]

Notes. Source: NEPS SC6; own calculations. Standard errors in brackets. $^{+}$ *p < 0.10, * p < 0.05, ** p < 0.01.*
Number of persons: 863; number of events (dropouts): 113. Controls (not reported in the table for the sake of clarity, see Table A4 in the appendix for complete results): gender, entry year of tertiary studies (linear and squared), changing from a university to a university of applied sciences, type of higher education entrance certificate. We did not control for changing from a university of applied sciences to an academic university because of too few events in this group.Due to low numbers of events in each single time period, Models 5 and 6 include constraints for educational pathways indicating that the effect for the first year equals the effect for the second year, and the effect for the third and fourth year equals the effect for the fifth to eighth year.

Summary and conclusion

In this paper, we have analysed the impact of pre-tertiary educational pathways and social origins on dropout rates in higher education, and the role of time in this interplay. Our findings suggest that pre-tertiary pathways have an effect on dropout from higher education. At university, we find that pathways other than the academic track or including a vocational qualification have higher dropout rates than the standard path. This is in line with Hypothesis 1, that students from outside the standard pathway face adaptation difficulties, and that students with vocational qualifications might be less committed to graduating because they have lower additional benefits from a tertiary degree (Hypothesis 2). At the universities of applied sciences, pathways with vocational training do not show significantly higher dropout rates, which might be due to the more practical orientation of the institutions as predicted by Hypothesis 4. The upwardly mobile pathway without vocational qualification, however, is connected to higher dropout rates at the universities of applied sciences as well. The age of students at enrolment to tertiary education seems to explain to a great extent why students with vocational qualifications are more prone to drop out from tertiary education. If age is an approximate measure for the amount of external responsibilities and obligations due to children or a demanding job, this can be interpreted as support for Hypothesis 3. However, further variables would have to be included, such as the birth of a child or working time, in order to have more direct evidence. Moreover, this interpretation has to be viewed with caution, because age at enrolment correlates highly with pre-tertiary pathways, and might therefore serve as an approximate measure of whether students have obtained a vocational qualification or not. As far as social origins are concerned, there is some evidence that, at university, the class-specific choices of educational pathways influence social selectivity in dropout (Hypothesis 5): at university, students with more highly educated parents have lower dropout rates. If the educational pathways are taken into account, the effect of social origins is reduced and ceases to be statistically significant. Thus, we find no support for Hypothesis 6a, that students with more highly educated parents are less prone to drop out due to their access to higher levels of social, cultural, or economic capital or due to the motive of status maintenance, once prior educational decisions are taken into account. However, these prior educational decisions are socially selective. Thus, there might be indirect effects of the motive of status maintenance via the pre-tertiary career. As expected in Hypothesis 6b, tertiary students with lower social origins might be positively selected on the basis of unobserved characteristics, which help them to counterbalance the disadvantages of their social background.

As anticipated in Hypotheses 7 and 8, at the universities, the effect of pre-tertiary educational pathways on dropout rates seems to weaken over the duration of studies – at least with respect to the standard pathway with vocational qualification and the upwardly mobile pathway without vocational qualification. At the universities of applied sciences, we do not find this pattern of decreasing effects of educational pathways. Considering that adaptation difficulties might be a major cause for early dropout, this finding fits the hypothesis that at the universities of applied sciences, students from a lower social background and students who reach higher education from outside the standard pathway, experience lower cultural divisions (Hypothesis 4). However, due to the lack of direct measures of adaptation, further research is needed to carefully test this hypothesis in more detail.

Regarding the arguments based on theories of educational decisions, we find that the risk of dropping out of higher education is influenced by decisions taken years before. Because the decisions for educational pathways are class-specific, they partly explain social inequality in dropout. This result confirms the well-established finding that educational decisions are path-dependent (Breen & Jonsson, 2000; Pfeffer & Goldrick-Rab, 2011). At the same time, results point in the direction that the dropout decision – although depending on previous class-specific decisions for educational pathways – is not influenced directly by social background when previous educational pathways are controlled. This might be interpreted as support for the life-course hypothesis (Shavit & Blossfeld, 1993) stating that social background loses its impact on decisions as students become adults and more independent from their parents.

Having no means to measure the financial situation of the respondents during their studies, we do not focus on the impact of financial costs on dropout. However, at least we can say that direct

tuition fees cannot play an important role in our sample, because they were not introduced in Germany until 2005 and impact on only a small proportion of the sample.

Drawing on the concept of differential learning environments, we argued that students are differently socialized and prepared for higher education depending on the educational pathway they have taken beforehand. We find this a plausible explanation for our result that dropout rates differ depending on the pre-tertiary pathway. However, it is not the only possible explanation, and a more precise test of the theoretical concept would require more detailed and direct measures. Apart from path dependency, results also point out that time dependency plays an important role in dropout. Modelling dropout as a process with time-specific dropout intensities can help to reveal sensitive phases in which certain groups are especially vulnerable.

This study contributes to sociological research on higher education in the following ways: First, it uses nationwide longitudinal data to analyse the inclination towards dropout as a process over time. As far as we are aware, this has not been undertaken so far regarding dropout from higher education in Germany. Second, to the best of our knowledge, the study is the first to apply a multivariate approach to investigate the interplay between previous pathways, social origin, and dropout in Germany. This interplay is an important issue for highly stratified education systems with strong social inequality: Education systems such as Germany with its selective and stratified education system and low tertiary graduation rates have a high interest in encouraging permeability to tertiary education in order to increase graduation rates. According to the results of this study, about 30 per cent of students who enter universities and even about two thirds of students who enter universities of applied sciences in Germany have not taken the standard pathway. Instead, they either did not attend the *Gymnasium* until higher secondary level and/or they completed a vocational qualification before tertiary education. Increasing the non-standard pathways to higher education, which are used more often by students from a lower social background, seems to be one possible way to reduce social inequality in access to higher education. However, this study finds that tertiary students who have taken non-standard pathways are more vulnerable to drop out from higher education – at least at academic universities, which are the higher education institutions attended by the majority of tertiary students. Thus, on the one hand, non-standard pathways might lead persons who otherwise would not have entered tertiary education to gain a tertiary degree; on the other hand, they are less efficient than standard pathways in terms of leading students to a degree. Therefore, target-oriented support of students with a non-standard educational biography might be a very effective way to counter the loss of potential university graduates during tertiary studies. In our analysis, we suggested some mechanisms that – if further tested with more detailed measures – might give insights into which kind of support is needed. Some answers on how to counter the difficulties of students with different educational biographies might also be found at the more practically oriented universities of applied sciences, in which the dropout propensity does not differ between different pathways.

We find considerable differences between the institutions regarding the effects of covariates on dropout: Parental education, pre-tertiary vocational qualifications, and age, influence dropout at universities but not at universities of applied sciences. Thus, the latter seem to be more successful in attracting and retaining students from lower social backgrounds, with vocational qualifications, and/or with higher ages. Therefore, an expansion of universities of applied sciences might foster permeability between vocational education and tertiary education, and decrease social inequality in the attainment of tertiary degrees. From a critical perspective, however, this could be interpreted as “effectively maintained inequality” (Lucas, 2001), because quantitative inequalities (differences in the proportion of students who complete tertiary education) would be shifted to qualitative inequalities (differences in the distribution over institutions). Following this reasoning, universities would preserve their selectivity, whereas persons with a vocational background and lower social background would be diverted to universities of applied sciences that provide less prestigious degrees and lower labour market returns. However, this would still reduce inequality compared to the status quo in which the expansion of eligibility to tertiary education is leading to students from a lower social background

being diverted to vocational education rather than to the universities of applied sciences (Lörz, 2013; Mayer et al., 2007).

The present analysis has its limitations, because students select different educational pathways themselves, and this selection is dependent on further characteristics that might also play an important role for dropout from higher education. However, we hope to control for the most important of these characteristics by including the grade-point average of the highest educational certificate as a proxy variable of achievement. Another drawback is that period-specific effects might be caused by the changing composition: less motivated students might leave first, so that the remaining population becomes composed more strongly of students resisting any temptation to drop out. However, by controlling for the grade-point average of the school leaving certificate, we hope to rule out at least a part of any composition effects.

Acknowledgements

This work was supported by the German Federal Ministry of Education and Research within the Framework Programme for the Promotion of Empirical Educational Research (section: Support for Young Scientists and Researchers). We would like to thank the anonymous reviewers for their helpful comments and suggestions.

References

Allmendinger, J. (1989). Educational systems and labor market outcomes. *European Sociological Review, 5*, 231–250.

Arum, R., Gamoran, A., & Shavit, Y. (2007). More inclusion than diversion: Expansion, differentiation, and market structure in higher education. In Y. Shavit (Ed.), *Stratification in higher education: A comparative study* (pp. 1–35). Stanford, CA: Stanford University Press.

Autorengruppe Bildungsberichterstattung. (2012). Bildung in Deutschland 2012: Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf. Bielefeld: Bertelsmann. Retrieved from http://www.bildungsbericht.de/

Ayalon, H., & Yogev, A. (2005). Field of study and students' stratification in an expanded system of higher education: The case of Israel. *European Sociological Review, 21*, 227–241.

Baethge, M. (2010). Neue soziale Segmentationsmuster in der beruflichen Bildung. In H.-H. Krüger, U. Rabe-Kleberg, R.-T. Kramer, & J. Budde (Eds.), *Bildungsungleichheit revisited. Bildung und soziale Ungleichheit vom Kindergarten bis zur Hochschule* (1st ed., pp. 275–298). Wiesbaden: VS Verlag für Sozialwissenschaften.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. New York, NY: Columbia University Press.

Becker, R., & Hecken, A. E. (2009). Why are working-class children diverted from universities? An empirical assessment of the diversion thesis. *European Sociological Review, 25*, 233–250.

Blossfeld, H.-P., Golsch, K., & Rohwer, G. (2007). *Event history analysis with Stata.* Mahwah, N.J: Lawrence Erlbaum Associates.

Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials: Towards a formal rational action theory. *Rationality and Society, 9*, 275–305.

Breen, R., & Jonsson, J. O. (2000). Analyzing educational careers: A multinomial transition model. *American Sociological Review, 65*, 754–772.

Chen, R. (2012). Institutional characteristics and college student dropout risks: A multilevel event history analysis. *Research in Higher Education, 53*, 487–505.

Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology, 34*, 366–375.

DesJardins, S. L. (2003). Event history methods: Conceptual issues and an application to student departure from college. In J. C. Smart & W. G. Tierney (Eds.), *Higher education: Handbook of theory and research* (pp. 421–471). Dordrecht, Netherlands: Kluwer Academic.

DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (1999). An event model of student departure. *Economics of Education Review, 18*, 375–390.

Hallinan, M. T., & Williams, R. A. (1990). Students' characteristics and the peer-influence process. *Sociology of Education, 63*, 122–132.

Hansen, M. N., & Mastekaasa, A. (2006). Social origins and academic performance at university. *European Sociological Review, 22*, 277–291.

Hanushek, E. A., & Wössmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *The Economic Journal, 116*, C63–76.

Heublein, U., Hutzsch, C., & Schreiber, J. (2010*).* Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen. Ergebnisse einer bundesweiten Befragung von Exmatrikulierten des Studienjahres 2007/08. *HIS:Forum Hochschule* F02/2010*.* Retrieved from http://www.his.de/pdf/pub_fh/fh-201002.pdf

Hillmert, S., & Jacob, M. (2003). Social inequality in higher education. Is vocational training a pathway leading to or away from university? *European Sociological Review, 19*, 319–334.

Hillmert, S., & Jacob, M. (2010). Selections and social selectivity on the academic track: A life-course analysis of educational attainment in Germany. *Research in Social Stratification and Mobility, 28*, 59–76.

Jacob, M., & Weiss, F. (2010). Soziale Selektivität beim Hochschulzugang. Veränderungen der Zugangssequenzen zur Hochschule im Kohortenvergleich. In B. Becker & D. Reimer (Eds.), *Vom Kindergarten bis zur Hochschule. Die Generierung von ethnischen und sozialen Disparitäten in der Bildungsbiographie* (pp. 285–312). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Knigge, M., & Hannover, B. (2011). Collective school-type identity: Predicting students' motivation beyond academic self-concept. *International Journal of Psychology, 46*, 191–205.

Lörz, M. (2013). Differenzierung des Bildungssystems und soziale Ungleichheit: Haben sich mit dem Ausbau der beruflichen Bildungswege die Ungleichheitsmechanismen verändert? *Zeitschrift für Soziologie, 42,* 118–136.

Lörz, M., & Schindler, S. (2009). Educational expansion and effects on the transition to higher education: Has the effect of social background characteristics declined or just moved to the next stage? In A. Hadjar & R. Becker (Eds.), *Expected and unexpected consequences of the educational expansion in Europe and the US. Theoretical approaches and empirical findings in comparative perspective* (pp. 97–110). Bern, Switzerland: Haupt.

Lucas, S. (2001). Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects. *American Journal of Sociology, 106*, 1642–1690.

Maaz, K., Chang, P.-H., & Köller, O. (2004). Führt institutionelle Vielfalt zur Öffnung im Bildungssystem? Sozialer Hintergrund und kognitive Grundfähigkeit der Schülerschaft an allgemein bildenden und beruflichen Gymnasien. In O. Köller, R. Watermann, U. Trautwein, & O. Lüdtke (Eds.), *Wege zur Hochschulreife in Baden-Württemberg. TOSCA - eine Untersuchung an allgemein bildenden und beruflichen Gymnasien* (pp. 153–204). Opladen, Germany: Leske und Budrich.

Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives, 2*, 99–106.

Mare, R. D. (1980). Social background and school continuation decisions. *Journal of the American Statistical Association, 75*, 295–305.

Marsh, H. W., Kong, C.-K., & Hau, K.-T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of Personality and Social Psychology, 78*, 337–349.

Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology, 47*, 213–231.

Mayer, K. U., Müller, W., & Pollak, R. (2007). Germany: Institutional change and inequalities of access in higher education. In Y. Shavit (Ed.), *Stratification in higher education: A comparative study* (pp. 240–265). Stanford, CA: Stanford University Press.

Müller, W., Brauns, H., & Steinmann, S. (2002). Expansion und Erträge tertiärer Bildung in Deutschland, Frankreich und im Vereinigten Königreich. *Berliner Journal für Soziologie, 12*, 37–62.

Müller, W., & Pollak, R. (2004). Weshalb gibt es so wenige Arbeiterkinder in Deutschlands Universitäten? In R. Becker & W. Lauterbach (Eds.), *Bildung als Privileg? Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* (pp. 311–352)*.* Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

OECD (2012). *Education at a glance 2012: OECD indicators*. Paris, France: OECD Publishing. Retrieved from http://www.oecd.org/edu/EAG%202012_e-book_EN_200912.pdf

Pascarella, E. T., Pierson, C. T., Wolniak, G. C., & Terenzini, P. T. (2004). First-generation college students: Additional evidence on college experiences and outcomes. *Journal of Higher Education, 75*, 249–284.

Pfeffer, F. T., & Goldrick-Rab, S. (2011). Unequal pathways through American universities. *Institute for Research on Poverty: Discussion Paper no. 1391-11*. Retrieved from http://www.irp.wisc.edu/publications/dps/pdfs/dp139111.pdf

Powell, J. J., & Solga, H. (2011). Why are higher education participation rates in Germany so low? Institutional barriers to higher education expansion. *Journal of Education and Work, 24*, 49–68.

Reay, D., Davies, J., David, M., & Ball, S. J. (2001). Choices of degree or degrees of choice? Class, "race" and the higher education choice process. *Sociology, 35*, 855–874.

Reimer, D., & Pollak, R. (2010). Educational expansion and its consequences for vertical and horizontal inequalities in access to higher education in West Germany. *European Sociological Review, 26*, 415–430.

Rohwer, G. (2006). Verweildauern und Übergangsraten bei mehreren Folgezuständen. In A. Diekmann (Ed.), *Methoden der Sozialforschung* (pp. 348–367). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Scheller, P., Isleib, S., & Sommer, D. (2013). *Studienanfängerinnen und Studienanfänger im Wintersemester 2011/12: Tabellenband.* Forum Hochschule: 6/2013. Hannover: HIS, Hochschul-Informations-System. Retrieved from http://www.his.de/pdf/pub_fh/fh-201306.pdf

Shavit, Y., & Blossfeld, H.-P. (1993). Persisting barriers: Changes in educational opportunities in thirteen countries. In Y. Shavit & H.-P. Blossfeld (Eds.), *Persisting inequality: Changing educational attainment in thirteen countries* (pp. 1–23). Boulder, CO: Westview Press.

Smith, J. P., & Naylor, R. A. (2001). Dropping out of university: A statistical analysis of the probability of withdrawal for UK university students. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 164*, 389–405.

Solga, H. (1997). Bildungschancen in der DDR. In S. Häder, H.-E. Tenorth, & O. Anweiler (Eds.), *Bildungsgeschichte einer Diktatur. Bildung und Erziehung in SBZ und DDR im historisch-gesellschaftlichen Kontext* (pp. 257–274). Weinheim, Germany: Deutscher Studien Verlag.

Statistisches Bundesamt. (2012). *Bildung und Kultur: Nichtmonetäre hochschulstatistische Kennzahlen. 1980–2011.* Fachserie 11. Reihe 4.3.1. Wiesbaden. Retrieved from https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/KennzahlenNichtmonetaer2110431117004.pdf?__blob=publicationFile

Stinebrickner, T., & Stinebrickner, R. (2012). Learning about Academic Ability and the College Dropout Decision. *Journal of Labor Economics, 30*, 707–748.

Thomas, L., & Quinn, J. (2006). *First generation entry into higher education: An international study*. Buckingham, England: Open University Press.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*, 89–125.

Tinto, V. (1988). Stages of student departure: Reflections on the longitudinal character of student leaving. *The Journal of Higher Education, 59*, 438–455.

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology, 98*, 788–806.

Appendix

Table A1. Distribution of characteristics in the sample at certain points in time at universities

	Year 0		Year 2		Year 5	
	No.	%	No.	%	No.	%
Men	921	54	834	54.4	577	57.4
Women	783	46	700	45.6	429	42.6
Entry cohort						
1970–1979	380	22.3	356	23.2	224	22.3
1980–1989	566	33.2	518	33.8	407	40.5
1990–1999	343	20.1	311	20.3	212	21.1
2000–2010	415	24.4	349	22.8	163	16.2
HE entrance certificate						
Other	68	4	54	3.5	31	3.1
General	1,636	96	1,480	96.5	975	96.9
Change of HE institution	0	0	25	1.6	36	3.6
Field of study						
Medicine	168	9.9	160	10.4	120	11.9
Teacher training	360	21.1	334	21.8	173	17.2
Humanities/Arts	293	17.2	257	16.8	171	17
Natural sciences/Mathematics	283	16.6	251	16.4	187	18.6
Engineering	197	11.6	176	11.5	126	12.5
Law	104	6.1	94	6.1	60	6
Business/Economics	200	11.7	172	11.2	115	11.4
Other	99	5.8	90	5.9	54	5.4
Parent's education						
CASMIN 1a/1b/1c	478	28.1	418	27.2	290	28.8
CASMIN 2a/2b	366	21.5	333	21.7	198	19.7
CASMIN 2c	198	11.6	175	11.4	109	10.8
CASMIN 3a/3b	662	38.8	608	39.6	409	40.7
Pathway						
Pathway 1	1,203	70.6	1,115	72.7	747	74.3
Pathway 2	183	10.7	144	9.4	83	8.3
Pathway 3	190	11.2	165	10.8	109	10.8
Pathway 4	128	7.5	110	7.2	67	6.7
GPA of highest school certificate (mean)	0.087		0.119		0.156	
Age at HE entry	21.32		21.2		21.08	
Total	1,704	100	1,534	100	1,006	100

Source: NEPS: SC6: 1.0.0; own calculations.

Table A2. Distribution of characteristics in the sample at certain points in time at universities of applied sciences

	Year 0		Year 2		Year 5	
	No.	%	No.	%	No.	%
Men	555	64.3	472	62.8	101	61.6
Women	308	35.7	279	37.2	63	38.4
Entry cohort						
1970–1979	206	23.9	188	25	27	16.5
1980–1989	254	29.4	233	31	58	35.4
1990–1999	187	21.7	164	21.8	43	26.2
2000–2010	216	25	166	22.1	36	22
HE entrance certificate						
Other	411	47.6	355	47.3	76	46.3
General	452	52.4	396	52.7	88	53.7
Change of HE institution			13	1.7	9	5.5
Field of study						
Engineering	322	37.4	285	38	72	44.2
Business science	135	15.7	112	14.9	21	12.9
Social work/Pedagogics	146	17	132	17.6	24	14.7
Other	257	29.9	221	29.5	46	28.2
Parent's education						
CASMIN 1a/1b/1c	405	46.9	363	48.3	72	43.9
CASMIN 2a/2b	184	21.3	159	21.2	35	21.3
CASMIN 2c	77	8.9	63	8.4	11	6.7
CASMIN 3a/3b	197	22.8	166	22.1	46	28
Pathway						
Pathway 1	287	33.3	257	34.2	53	32.3
Pathway 2	156	18.1	136	18.1	28	17.1
Pathway 3	116	13.4	96	12.8	21	12.8
Pathway 4	304	35.2	262	34.9	62	37.8
GPA of highest school certificate (mean)	-0.029		0.006		-0.066	
Age at HE entry	22.33		22.20		22.26	
Total	863		751		164	

Source: NEPS: SC6: 1.0.0; own calculations.

Table A3. Piecewise constant exponential models on dropout from universities within 8 years of starting tertiary studies

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Year 1	-10.53*	-13.56**	-11.46*	-11.64*	-11.55*	-11.72*
	[4.89]	[4.88]	[4.89]	[4.87]	[4.89]	[4.87]
Year 2	-10.23*	-13.25**	-11.15*	-11.33*	-11.18*	-11.36*
	[4.89]	[4.88]	[4.89]	[4.87]	[4.89]	[4.87]
Year 3–4	-10.99*	-13.99**	-11.88*	-12.06*	-11.94*	-12.10*
	[4.89]	[4.89]	[4.90]	[4.88]	[4.89]	[4.88]

(Table A3 cont'd)	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Year 5–8	-10.76*	-13.77**	-11.65*	-11.82*	-11.67*	-11.83*
	[4.90]	[4.90]	[4.91]	[4.89]	[4.91]	[4.89]
Women	0.19	0.24*	0.23^{+}	0.28*	0.23^{+}	0.28*
	[0.12]	[0.12]	[0.12]	[0.12]	[0.12]	[0.12]
Entry year	0.12	0.17	0.12	0.08	0.12	0.08
	[0.11]	[0.11]	[0.11]	[0.11]	[0.11]	[0.11]
Entry year sq.	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Change of HE inst.	-0.15	-0.26	-0.21	-0.19	-0.24	-0.21
	[0.45]	[0.45]	[0.45]	[0.45]	[0.45]	[0.45]
General HE entrance certificate	-0.18	-0.05	0.26	0.42	0.29	0.45
	[0.29]	[0.29]	[0.31]	[0.32]	[0.31]	[0.32]
Parents: CASMIN 1a/1b/1c	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
Parents: CASMIN 2a/2b	-0.22	-0.16	-0.09	-0.07	-0.09	-0.07
	[0.17]	[0.17]	[0.17]	[0.17]	[0.17]	[0.17]
Parents: CASMIN 2c	0.07	0.16	0.24	0.23	0.24	0.23
	[0.19]	[0.19]	[0.19]	[0.19]	[0.19]	[0.19]
Parents: CASMIN 3a/3b	-0.42**	-0.26^{+}	-0.14	-0.12	-0.13	-0.12
	[0.15]	[0.15]	[0.16]	[0.15]	[0.16]	[0.15]
GPA of highest school certificate		-0.67**	-0.66**	-0.64**		
		[0.08]	[0.08]	[0.08]		
in year 1					-0.70**	-0.68**
					[0.18]	[0.18]
in year 2					-0.49**	-0.47**
					[0.16]	[0.16]
in year 3–4					-0.81**	-0.78**
					[0.17]	[0.17]
in year 5–8					-0.70**	-0.66**
					[0.16]	[0.15]
Age at enrolment				0.09**		0.09**
				[0.02]		[0.02]
Pathway 1				*ref.*	*ref.*	*ref.*
Pathway 2			0.50**	0.15		
			[0.18]	[0.21]		
in year 1–2					0.67**	0.32
					[0.23]	[0.26]
in year 3–8					0.30	-0.05
					[0.27]	[0.29]
Pathway 3			0.45*	0.40*		
			[0.18]	[0.18]		
in year 1–2					0.64**	0.59*
					[0.24]	[0.24]
in year 3–8					0.23	0.19
					[0.27]	[0.27]
Pathway 4			0.64**	0.18		
			[0.21]	[0.25]		
in year 1–2					0.40	-0.06
					[0.31]	[0.34]
in year 3–8					0.85**	0.39
					[0.27]	[0.29]

*Notes. Standard errors in brackets. $^{+}$ p < 0.10, * p < 0.05, ** p < 0.01.Source: NEPS SC6; own calculations.*
Number of persons: 1,704; number of events (dropouts): 289. Due to low numbers of events per pathway in each single time period, Models 5 and 6 include constraints for educational pathways indicating that the effect for the first year equals the effect for the second year and the effect for the third and fourth year equals the effect for the fifth to eighth year.

Table A4. Piecewise constant exponential models on dropout from universities of applied sciences within 8 years of starting tertiary studies

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Year 1	-8.51	-12.64^{+}	-14.53^{+}	-14.62^{+}	-13.77^{+}	-13.84^{+}
	[7.22]	[7.37]	[7.45]	[7.48]	[7.48]	[7.50]
Year 2	-8.74	-12.85^{+}	-14.73*	-14.82*	-13.98^{+}	-14.05^{+}
	[7.22]	[7.37]	[7.45]	[7.48]	[7.48]	[7.50]
Year 3–4	-9.11	-13.21^{+}	-15.07*	-15.17*	-14.52^{+}	-14.59^{+}
	[7.23]	[7.38]	[7.46]	[7.49]	[7.49]	[7.52]
Year 5–8	-8.40	-12.56^{+}	-14.45^{+}	-14.54^{+}	-13.78^{+}	-13.84^{+}
	[7.25]	[7.40]	[7.47]	[7.50]	[7.51]	[7.53]
Women	-0.29	-0.18	-0.18	-0.18	-0.18	-0.18
	[0.21]	[0.21]	[0.21]	[0.21]	[0.21]	[0.21]
Entry year	0.06	0.15	0.19	0.19	0.17	0.18
	[0.16]	[0.16]	[0.17]	[0.17]	[0.17]	[0.17]
Entry year sq.	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
General HE entrance certificate	-0.30	-0.48*	-0.52*	-0.52*	-0.50*	-0.50*
	[0.20]	[0.20]	[0.23]	[0.23]	[0.23]	[0.23]
Parents: CASMIN 1a/1b/1c	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
Parents: CASMIN 2a/2b	-0.23	-0.19	-0.17	-0.18	-0.18	-0.19
	[0.27]	[0.27]	[0.27]	[0.27]	[0.27]	[0.27]
Parents: CASMIN 2c	0.23	0.23	0.33	0.33	0.33	0.33
	[0.32]	[0.33]	[0.33]	[0.33]	[0.33]	[0.33]
Parents: CASMIN 3a/3b	0.08	0.01	0.07	0.07	0.07	0.07
	[0.25]	[0.24]	[0.25]	[0.25]	[0.25]	[0.25]
GPA of highest school certificate		-0.89**	-0.90**	-0.90**		
		[0.15]	[0.15]	[0.15]		
in year 1					-1.01**	-1.01**
					[0.27]	[0.27]
in year 2					-1.05**	-1.05**
					[0.31]	[0.31]
in year 3–4					-0.90**	-0.90**
					[0.30]	[0.30]
in year 5–8					-0.53	-0.53
					[0.33]	[0.34]
Age at enrolment				-0.01		-0.00
				[0.03]		[0.03]
Pathway 1			*ref.*	*ref.*	*ref.*	*ref.*
Pathway 2			0.14	0.15		
			[0.30]	[0.32]		
in year 1–2					-0.05	-0.03
					[0.39]	[0.40]
in year 3–8					0.36	0.37
					[0.46]	[0.47]
Pathway 3			0.73*	0.73*		
			[0.29]	[0.29]		
in year 1–2					0.72*	0.72*
					[0.36]	[0.36]
in year 3–8					0.69	0.69
					[0.46]	[0.46]
Pathway 4			0.01	0.02		
			[0.29]	[0.30]		
in year 1–2					-0.23	-0.22
					[0.36]	[0.37]
in year 3–8					0.30	0.31
					[0.40]	[0.41]

*Notes. Standard errors in brackets. $^{+}$ p < 0.10, * p < 0.05, ** p < 0.01. Source: NEPS SC6; own calculations.*
Number of persons: 863; number of events (dropouts): 113.
Due to low numbers of events in each single time period, Models 5 and 6 include constraints for educational pathways indicating that the effect for the first year equals the effect for the second year and the effect for the third and fourth year equals the effect for the fifth to eighth year.

Endnotes

[1] The "contrasting nature" (Mayer et al., 2007) between universities and universities of applied sciences might have converged to some degree due to the restructuring of study programmes decided in Bologna in 1997. In this article, however, we still concentrate mainly on study programmes heading towards the traditional degrees, because only a small portion of the present sample is already subject to this recent development.

[2] Even when we control for pre-tertiary pathways but not for the grade-point average of the highest school certificate, the effect of parental education is still significant on a 10 per cent level.

[3] The literature shows that the choice of fields of study relates to social origins (e.g. Ayalon & Yogev, 2005). There is also evidence that dropout depends on the field of study (Heublein, Hutzsch & Schreiber, 2010). Moreover, educational pathways might be connected to certain fields of study, making fields of study a potential mediator in the relationship between social origins, educational pathways, and dropout. Unfortunately, our data do not allow us to include fields of study as a time-varying covariate, because interviewers were asked to record only the last major in respondents who changed their field of study. In order to test for robustness, however, we include groups of study fields as time-constant variables. The effects of social origins and educational pathways do not change when the fields of study are included as covariates (see Tables A5 and A6 in the supplementary material).

Does a stepped approach using mixed-mode data collection reduce attrition problems in a longitudinal mental health study?

Adriaan W Hoogendoorn [1]
a.hoogendoorn@ggzingeest.nl
Femke Lamers [1]
Brenda W J H Penninx [1, 2, 3]
Johannes H Smit [1]

[1] Department of Psychiatry and the EMGO Institute for Health and Care Research, VU University Medical Center Amsterdam, The Netherlands
[2] Department of Psychiatry, Leiden University Medical Center, The Netherlands
[3] Department of Psychiatry, University Medical Center Groningen, University of Groningen, The Netherlands

(Received January 2013 Revised August 2013) doi:10.14301/llcs.v4i3.240

Abstract

According to epidemiological standards for longitudinal studies, any appropriate attempt should be made to reach a high response rate. This also concerns re-contacting respondents in follow-up waves. In the Netherlands Study of Depression and Anxiety (NESDA), respondents, who refused to visit the study site for a follow-up interview and would have been lost in a single step approach, were offered an interview at their home (in a second step) or an interview by telephone (in a third step). The stepped approach intended to increase response rates, but as a by-product, the switching of the interview response setting introduced effects due to the three different modes for obtaining data. This study aimed to evaluate the benefits and detriments of the introduction of a stepped approach using mixed-mode data collection. The attained attrition in the stepped approach was compared to the attrition rate that would have been achieved if the design would have remained in a single mode. Logistic regression models were used to study if the attrition was related to patient characteristics, i.e. the attrition was selective. Propensity score matching was used to study if similar respondents reacted differently to different modes. The stepped approach using mixed-mode collection reduced the attrition from 22% to 13%, while the selectivity of the attrition was reduced on almost all socio-demographic variables, although the selectivity increased in variables that reflect the respondents' mental health. Propensity score matching demonstrated no evidence of mode effects. Although the introduction of the stepped approach using mixed mode data collection did not completely solve the problem of attrition, we conclude that adding different modes was worthwhile.

Keywords: methodology, attrition, mixed-mode data collection, anxiety, depression

1. Background

In health sciences, longitudinal research designs are used to follow samples of persons over a period of time. Such designs permit researchers to study the progression of diseases or other changes in health, wellbeing and illness at the individual level. For example, in the mental health setting, an important research topic is the course of anxiety and depression. In the Netherlands, approximately one out of four persons will be faced with these disorders at some time during their lives (Bijl, Ravelli & Van Zessen, 1998). It is still not clear why some persons are more susceptible to these disorders than others, or why some persons recover quickly while others suffer for long periods of time. A longitudinal (panel) design will provide data that allow researchers to investigate these kinds of questions.

One of the major problems in the analysis of longitudinal data is attrition. Attrition refers to the phenomenon that respondents drop out of a longitudinal study. There are several reasons why respondents drop out of the study. Some respondents cannot be traced after they have moved, some respondents can be traced but are hard to contact and other respondents refuse to cooperate when contacted. The occurrence of attrition is problematic, since it not only reduces sample size, but may also lead to biased estimates where the drop-out is non-random (Goldstein, 2009). For example, in a longitudinal study on the effect of smoking on cognitive decline, it appeared that attrition was associated both with smoking and with cognitive functioning. A relatively high attrition of smokers was found (due to mortality) in combination with a relatively low attrition of respondents with high cognitive scores. In this case, ignoring attrition resulted in an underestimation of the negative effect: weighting to account for attrition yielded estimates that were up to 86% higher (Weuve et al., 2012).

Similar problems occur in longitudinal mental health studies or in cohort studies in psychiatric epidemiology, where researchers found that mental status is predictive for attrition. In an epidemiological follow-up study (NHANES I), the researchers found that depressed respondents had a higher attrition rate even after adjustment for socio-demographic variables, health status variables and smoking behaviour, and suggested that depressive symptoms be measured at the baseline interview in order to predict attrition, or to use the information to correct for potential bias (Farmer, Locke, Liu, & Moscicki, 1994). Similar results that mental status is predictive for attrition were found in other studies (de Graaf, Bijl, Smit, Ravelli, & Vollebergh, 2000; Fischer, Dornelas, & Goethe, 2001).

In order to deal with attrition, the standard practice is trying to keep it as low as possible. Low attrition can best be realized by putting effort into: reducing the different sources of panel attrition (Watson & Wooden, 2009), locating the sample members (Couper & Ofstedal, 2009): contacting the sample members (Groves & Couper, 1998), and getting cooperation from sample members by using incentives, stressing topic saliency and taking previous interview experiences into account (Groves & Couper, 1998; Lepkowski & Couper, 2002; Laurie & Lynn, 2009). In mental health settings, some researchers recommend intensive strategies to limit attrition by maintaining a continuous personal relationship between researcher and study participants (Susser, Schwartz, Morabia, & Bromet, 2012). In the general survey context, a modern tool to improve response rates is the use of mixed-mode surveys (de Leeuw, 2005; Dillman, Smyth, & Christian, 2009). Offering participants a certain mode initially, and - in case of non-response - offering participants a second or even third mode, can improve response rates and reduce non-response errors by getting responses from participants who refused the initial mode of data collection. This method, that we describe as a stepped approach using mixed-mode data collection, is also known as “offering multiple modes in a specific sequence” or a Type III of mixed mode survey (Dillman et al., 2009).

However, the use of different survey modes also introduces mode effects, since respondents may react differently within different interview modes (Groves, 1989). Several meta-studies discuss effects of data collection modes (de Leeuw, 1992; Bowling, 2005). Differences can be large if data are collected either by interviewers, or by self-interviewing, or if data collection is computer-assisted or not. Mode effects concern differences in both measurement errors (social desirability, acquiescence bias, interviewer bias) and non-measurement errors (coverage errors, non-response errors, item non-response). In fact, the differences in non-measurement errors make mixed-mode data

collection attractive for the purpose of attrition reduction, because respondents who were missed by a first mode can be reached by a second mode (Voogt & Saris, 2005). In the context of the present study, mode effects are delineated in the narrow sense of measurement errors, since differences in measurement errors are undesirable. Such effects on measurement errors can occur even if modes are comparable with respect to non-measurement error differences, e.g. if an interviewer is present or not, or if an interview is computer-assisted. From a meta-study on several studies that compared data collected in face-to-face and telephone interviews, the authors conclude that compared to face-to-face interviewing, telephone interviewing is associated with respondents who have a lower level of engagement in the interview, present themselves in more socially desirable ways and show more satisficing behaviour (Holbrook, Green, & Krosnick, 2003; Jackle, Roberts, & Lynn, 2006). An experimental study in the context of the European Social Survey confirmed the finding that telephone respondents were more likely to give socially desirable responses across a range of indicators, while no evidence was found for the hypothesis that telephone respondents were more likely to satisfice (Jackle et al., 2006).

A switch in survey mode in a longitudinal survey makes it problematic to determine if an observed change is a result from actual changes or from a change in survey mode (Dillman et al., 2009). Therefore, introducing a mixed-mode approach in a longitudinal survey may give us mixed results: on the one hand we expect a reduction of attrition as a favourable outcome, but on the other hand we may end up with mode effects that could complicate analyses.

This study aims to analyze the results of the introduction of a mixed-mode design, into a 2-year follow-up wave in a longitudinal naturalistic cohort study (n = 2,981), looking at the course of depressive and anxiety disorders. We will begin by describing the effect of introducing the mixed-mode survey design on the *attrition rate*. Recent literature points out that higher response rates (or lower attrition) do not guarantee lower non-response error. Therefore, we will also study the degree to which the stepped approach using mixed-mode data collection, prevents the attrition being dependent on respondent characteristics, i.e. to what extent we are faced with *selective attrition*. Next, we will give attention to disentangling measurement errors from non-measurement errors (Van der Laan & Van Nunspeet, 2009; Vannieuwenhuyze, Loosveldt, & Molenberghs, 2010). Finally, we will consider the positive and negative effects of the introduction of the mixed-mode effects, and will evaluate whether the introduction of the mixed-mode design was worthwhile in a mental health setting.

2. Method

2.1. NESDA Sample

The Netherlands Study of Depression and Anxiety (NESDA) is a longitudinal naturalistic cohort study, consisting of 2,981 persons that included individuals aged between 18 and 65 year, the majority of which (2,329; 78%) had anxiety or depressive disorders and a smaller group of healthy controls (n = 652; 22%). Specification of the sampling and recruitment procedures is discussed in detail elsewhere (Penninx et al., 2008). Participants were recruited from several settings: from primary health care (general practitioners), from specialized mental health care and from the community, i.e. outside any health care, in order to increase the generalizability of results and to cover the full range of psychopathology. The participants in the study were recruited from September 2004 to the end of February 2007 at three study sites (Amsterdam, Groningen and Leiden). Approval of the study protocol was granted by the ethical review boards of all participating centres, and all participants gave written informed consent.

2.2. Switching from a single mode to a mixed-mode design

The baseline measurement involved a four hour interview combined with two self-administered questionnaires, in which information was gathered on demographic variables, physical and social functioning, psychopathology, a medical examination, a cognitive computer task and collection of blood and saliva samples. All participants were interviewed in a face-to-face setting at the study site, and received travel expenses and a 15 euro gift certificate.

Two years after the baseline measurement, all participants were re-invited for the follow-up measurement. A letter announcing the follow-up measurement was followed a week later by attempts to make an appointment by telephone. At

least eight attempts were made, at several times and days. If no contact could be made, the contact information was checked with care providers or with the contact persons reported by the participants at the baseline measurement. Then five further attempts were made. If the contact attempts were not successful, the contacting procedure was terminated. When contacted, participants who were unable or unwilling to visit the study site were offered an interview at their home, or subsequently, an interview by telephone. This strategy resulted in a mixed-mode design at follow-up, with respondents interviewed in a face-to-face setting at the study site, respondents interviewed in a face-to-face setting at home and respondents interviewed by telephone. Outside these three groups, a remaining proportion of those contacted still refused to participate, as described in the next section.

2.3. Measurements

2.3.1. Attrition

The main outcome variable is the indicator that a follow-up measurement took place. Since the attained response was achieved by combining different modes, we will refer to this indicator as 'response in the mixed-mode design' and its complement as 'attrition in the mixed-mode design'. In case of attrition in the mixed-mode design, the given reason for attrition was coded as 'non-contact', 'unable to participate', 'refusal' or 'deceased'.

Next to 'response in the mixed-mode design' we studied 'response in the single mode design' which we call the 'on-site response' and which would have been the response in the hypothetical situation that no interviews were held at respondents' homes, and no telephone interviews were held, and only the on-site respondents responded. For reasons of clarity, we used the response behaviour at the follow-up wave to partition the total group of baseline responders into four groups: the 'on-site response', the 'at home response', the 'telephone response' and the 'non-response in the mixed-mode design'. The last three groups correspond to the 'non-response in the single mode design', while the first group corresponds to the 'response in the single mode design' and the first three groups correspond to the 'response in the mixed-mode design'.

2.3.2. Mode effects

As described above, in the follow-up measurement of the NESDA study, for some respondents the interview mode changed from a face-to-face interview at the study site to a face-to-face interview at the respondents' home or to a telephone interview (at the respondents' home). Introducing different survey modes in a survey design introduces different measurement errors if respondents provide different answers to the same questions depending on the mode being used to answer the question (Dillman et al., 2009). The amount of mode effect differs between questions due to differences in required cognitive efforts, or to differences in their susceptibleness to social desirability. Mode effects have been found in several studies comparing different interview modes. For example, differences in response behaviour have been assessed between telephone interviews and web-based self-interview, on questions related to experiences with environmental hindrance (Lugtig, Lensvelt-Mulders, Frerichs, & Greven, 2011), between telephone interviews and face to face interviews with respect to income related questions (Fessler, Kasy, & Lindner, 2012) and between face to face interview, mail (postal) mode and web mode, with respect to questions in a public health survey (Tipping, Hope, Pickering, Erens, Roth, & Mindell, 2010; Lorenc, 2010).

In this study, we examined if mode effects occurred with respect to two carefully selected variables: the number of depression symptoms measured by the Composite Interview Diagnostic Instrument (CIDI) – lifetime version 2.1 – to diagnose depressive and anxiety disorders according to Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition (DSM-IV) algorithms and the score on self-rated health. Reasons to select these specific variables to investigate mode effects included the fact that these variables came from the interview - not from the self-administered paper questionnaire that accompanied the interview - the relatively high measurement level compared to other (mostly dichotomous) variables in the interview, their susceptibility to social desirability, and the key role these variables play in the NESDA study.

To find out if mode effects occurred, we analyzed to what extent the responses to the two key variables differ between the three response

groups 'onsite', 'at home' and 'telephone'. In fact, as a result of respondents' preferences, we expected the three response groups to differ with respect to many variables. For example, we expected that the respondents who were interviewed in their homes would suffer more from psychiatric distress, but we also expected that they would prefer an at home interview because, at the baseline interview, they were older, and already suffering from psychiatric distress. Therefore we had to find out if the differences in the key variables were a result of different modes or a result of selection: the fact that the respondents differ between response groups.

In disentangling mode effects from selection effects, the choice of background variables is crucial, since these background variables determine the aspects in which the matched samples will be similar. We used background variables that were strongly related to the interview mode and strongly related to the two key variables. For the set of variables that related strongly to the interview mode, we used the set that Lamers et al. (2012) used to study determinants of attrition. For the set of variables that related strongly to the two key outcome variables, we used their equivalents measured at the baseline interview, since for these variables the scores on baseline and follow-up were highly correlated. The idea is that after controlling for both sets of variables, the observed differences in the key variables can be ascribed to mode effects.

2.3.3. Socio-demographic characteristics

Socio-demographic variables include age, gender, number of years of education, non-Northern European ancestry (yes/no), sampling site (Amsterdam, Groningen, Leiden), urbanization (high/low), distance to interview site (km), employment status (yes/no), partner status (yes/no) and number of chronic diseases. Urbanization was defined as high if the number of addresses was larger than 1500 per square kilometer. Because individuals who have previously participated in longitudinal research may be more motivated and therefore less likely to drop out of the NESDA study, previous participation in research (yes/no) was considered as a determinant of attrition. Distance to the interview site was determined by converting digit postal codes of respondents and interview sites to geographical positions (latitude and longitude) and by computing the geographical distance between these positions. The distance to the interview site was not incorporated in the Lamers et al. (2012) study, but was added to the analysis here, since this variable can be expected to play an important role in comparing attrition in the case where respondents have to visit the on-site location for their interview, with the case where interviews are carried out at the respondents' home or by telephone. The number of chronic diseases is not usually taken as a socio-demographic variable, but it was added as an indicator of general health status, being a count of self-reported presence of chronic diseases for which the respondent received treatment.

2.3.4. Psychiatric characteristics

Psychiatric characteristics included the 1-year diagnosis of depressive disorders (major depressive disorder (MDD) and dysthymia), anxiety disorders (social phobia, panic disorder with agoraphobia, panic disorder without agoraphobia, agoraphobia, general anxiety disorder (GAD)), and alcohol use disorders: alcohol dependence and alcohol abuse. All characteristics were obtained from the Composite International Diagnostic Interview (CIDI, lifetime version 2.1, (World Health Organization, 1997) to diagnose disorders according to the criteria of Diagnostics and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV). The number of DSM-IV depression symptoms (lifetime) was assessed using the depression section of the CIDI. Based on the items of the CIDI, the number of MDD symptoms ranged from 0 to 9 as stated by DSM-IV. Self-rated health was assessed by asking respondents to rate their own health on a scale from 0 to 100 after they rated the health of a person with a broken leg, a person with severe arthritis, a person with terminal cancer. Participants rated their own physical and mental health on the same scale, similar to the items in the '30-day function' section of the Collaborative Psychiatric Epidemiology Surveys (Alegria, Jackson, Kessler, & Takeuchi, 2008).

2.4. Statistical analyses

To study attrition selectivity, we started with a bivariate approach, comparing means of socio-demographic and psychopathological variables between the response group and the attrition group. Next, we took a multivariate approach, modeling attrition by use of multiple logistic regression analysis. For both modes – the single

mode and the mixed-mode – three logistic regression models will provide analyses of determinants of attrition, where all other potentially confounding variables are taken into account: one model using socio-demographic variables only, one using psychiatric variables only and a comprehensive model that combines these sets of variables. These logistic regression models allowed us to determine the *R*-indicator (Schouten, Cobben, & Bethlehem, 2009), a measure for representativeness which assesses the similarity of the response and the sample of the survey with respect to a certain set of auxiliary variables. Instead of comparing the response with the sample, we will apply the *R*-indicator to compare the response at the follow-up wave to the response at the baseline wave. The representativeness of the response at the follow-up wave provides us with a measure of the selectivity of the attrition. The *R*-indicator with respect to the auxiliary variable vector X is defined as $R(\rho_X) = 1 - 2S(\rho_X)$, where ρ_X is the propensity to respond given background X and where $S(\rho_X)$ is the standard deviation of these propensities. Ideally, X contains all relevant characteristics. In practice, one has to make a reasonable choice of available background variables and, since the response propensities ρ_X are unknown, one has to estimate the response propensities using a *logit* or *probit* model. The indicator takes values between 0, representing maximal selectivity, and 1, representing no selectivity at all, and can be translated into upper bounds for the non-response bias of the response mean. In the context of attrition, the *R*-indicator measures the similarity of the response at the follow-up wave and the response at baseline.

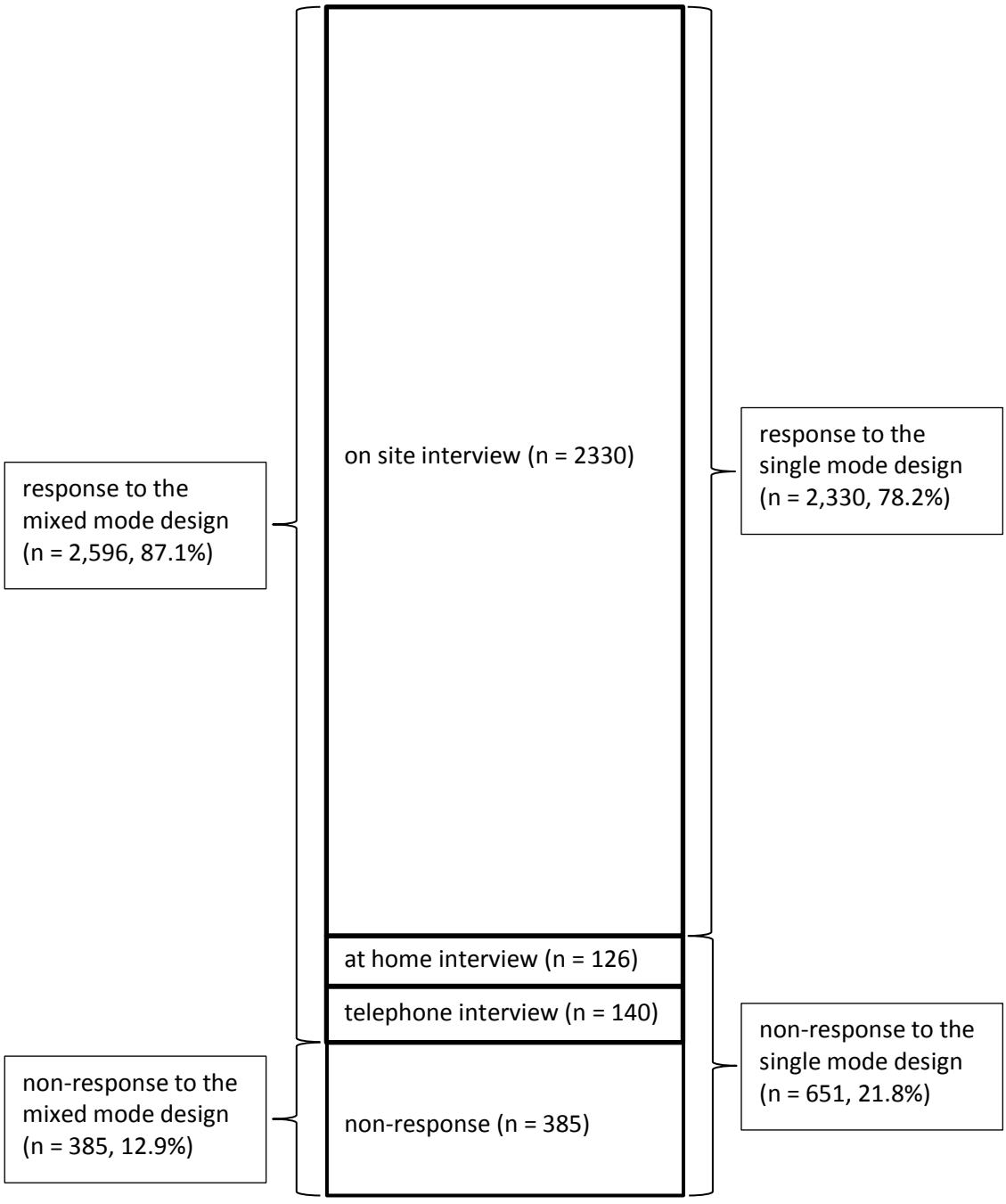
To study mode effects we considered that these were confounded by self-selection (de Leeuw, 2005). Propensity score methods have been used in several studies to investigate mode effects in mixed mode (Tipping et al., 2010; Lorenc, 2010; Lugtig et al., 2011; Fessler et al., 2012). Similar to (Lugtig et al., 2011) we used a method called 'coarsened exact matching' (CEM) (Iacus, King, & Porro, 2012) to eliminate differences in sample composition using a set of covariates. CEM is a form of propensity score matching that allows the researcher to "coarsen" data by groups for which the matches are made. The degree of coarseness is determined by the researcher. For example: for education, based on the number of years of education, we created three meaningful groups: a group with zero to seven years of education, corresponding to primary school level, a group with more than seven but less or equal to twelve years of education, corresponding to secondary school level, and a group with more than twelve years of education, corresponding to tertiary school level. Also, the matching is not one to one, but respondents are matched if they fall within the same stratum of "coarsening". After matching, the groups of respondents are similar with respect to all the covariates used in the matching procedure, so that the remaining differences can be attributed to mode effects, not to selection effects. Since there are two different response groups that use face-to-face interviewing – the onsite group and the at home group – we performed two matches. We statistically matched the onsite group to the telephone group, and the at home group to the telephone group, and compared these two pairs of groups to evaluate if there were mode effects. To compensate for the differential strata sizes, weights were computed and used in subsequent analyses.

3. Results

3.1. The amount of attrition

Figure 1 shows how the 2,981 respondents of the baseline sample were distributed across the different response categories for the follow up measurement. The attrition rate was 12.9%, and was mainly due to refusal (8.9%) and, to a lesser extent, to being unable to participate (2.3%) or to non-contact (1.5%). In evaluating the results of moving from the single mode design to the mixed-mode design, we kept in mind that the mixed-mode design is obtained sequentially, starting as a single mode design and then turned into a mixed-mode design by adding alternative 'modes'. As a result, the attrition rate can only decrease if we move from the single mode design towards the mixed-mode design. Figure 1 indicates that the attrition would have been 21.8% for the single mode design, and decreased to 12.9% in the mixed-mode design. The table reveals that the reduction in attrition by 8.9% was attributed both to offering the option of an interview at home – reducing the attrition by 4.2% – and to the option of a telephone interview at home – reducing the attrition by an additional 4.7%.

Figure 1. Response and non-response to the follow-up interview of the base line sample of 2,981 respondents in NESDA, according to the (realized) mixed mode design and to the (hypothetical) single mode design



Notes. Mixed mode design = on site interview, at home interview and telephone interview.
Single mode design = on site interview only.

Table 1. Baseline socio-demographic and psychopathological characteristics (means and percentages) of response on non-response groups for single and mixed-mode design

	Single mode			Mixed-mode		
	response *n*=2330	non-response *n*=651	*p-value*	response *n*=2596	non-response *n*=385	*p-value*
Socio-demographic variables						
Age – in years (SD)	42.2 (13.2)	40.7 (12.5)	.009	42.0 (13.1)	41.0 (12.9)	.144
Female vs Male	65.2	70.5	.012	66.1	68.3	.392
Education – in years (SD)	12.3 (3.2)	11.5 (3.3)	.000	12.3 (3.3)	11.2 (3.1)	.000
Non-Northern European descent vs Northern European descent	4.2	9.1	.000	4.5	10.1	.000
Sampling Site						
Amsterdam	36.7	56.8	.000	39.6	50.6	.000
Leiden	32.0	23.2	.000	30.2	28.8	.574
Groningen	31.4	20.0	.000	30.1	20.5	.000
High vs low urbanization	70.8	68.0	.179	69.6	73.8	.099
Distance in km to interview site (SD)	17.7 (37.6)	26.1 (45.0)	.000	20.4 (40.5)	13.9 (31.2)	.000
Previous participation in research	17.8	23.0	.002	20.1	10.6	.000
Not employed vs employed	35.4	37.5	.330	34.9	42.6	.003
No partner vs partner	30.9	29.8	.576	30.7	30.6	.984
Number of chronic diseases	0.9	1.0	.031	0.9	1.0	.058
Psychopathology						
Depressive disorders						
MDD	39.1	49.3	.000	38.8	58.4	.000
Dysthymia	9.5	15.8	.000	9.9	17.9	.000
Anxiety disorders						
Social phobia	22.1	28.0	.002	22.2	31.2	.000
Panic disorder with agoraphobia	13.4	19.0	.000	13.6	21.6	.000
Panic disorder without agoraphobia	8.6	11.2	.040	8.7	12.2	.026
Agoraphobia	6.5	6.3	.836	6.2	8.1	.178
General anxiety disorder	15.2	21.7	.000	15.3	25.2	.000
Alcohol use disorders						
Alcohol dependence	15.2	16.3	.479	14.9	18.4	.076
Alcohol abuse	11.7	12.3	.690	11.6	13.2	.361

Note. p-values are obtained from independent samples t-tests.

3.2. Attrition selectivity

Table 1 shows mean baseline characteristics for the response and non-response groups, both for the single and the mixed-mode design. For the *mixed-mode design* the comparison was published previously by our group (Lamers et al., 2012). We concluded that respondents who dropped out were less well educated, more often of non-Northern European descent, were more often (by comparison with Leiden) recruited in Amsterdam and less often in Groningen, lived in very urbanized areas, had not previously participated in a psychiatric study, were less often employed, and more often suffered, at baseline, from MDD, dysthmia, social phobia, panic disorders and general anxiety disorders. For the *single mode design*, the situation is roughly similar. Comparing the two designs, we observe that for gender, age, sampling site, number of chronic diseases the differences between the response group and the non-response group in the *mixed-mode design* were non-significant, while these differences in the *single mode design* were

significant at the p = 0.05 level. In other words: for these variables the amount of selectivity was reduced. On the other hand, for urbanization, employment status and for most of the psychopathological variables, moving from single to mixed-mode design made the selectivity worse.

We estimated multiple logistic regression models to explain the attrition in both designs (the single mode design and the mixed-mode design) from socio-demographic and psychopathology characteristics. Table 2 shows that in the *mixed-mode design*, younger age, fewer years of education, not being of Northern European descent, and having been recruited in Amsterdam were determinants of attrition. Of the psychiatric diagnoses, MDD and dysthymia were determinants of attrition. For the *single mode design*, Table 2 indicates roughly the same determinants, but with generally stronger effects (odds ratios that are generally more distant from 1). A remarkable exception is the selectivity with respect to the depressive disorder MDD: the odds ratio increased from 1.27 to 1.60. We will discuss the selectivity with respect to this variable in more detail later in this section. Note that the odds ratio from 'distance to site' drops from 1.58 in the single mode design to 0.73 in the mixed-mode design, indicating that adding the additional modes reduced distance from the study site as a risk factor for attrition.

Table 2 also shows values of *R*-indicators (representativity indicators) based on the logistic regression models. Note that the value of the *R*-indicator of the comprehensive model of the mixed-mode design (0.84) was higher than the corresponding value for the single mode design (0.75). The increase in *R*-indicator indicates that the selectivity – with respect to the socio-demographic and psychopathological variables in the model – was reduced. If we focus on the models using socio-demographic and psychopathology characteristics separately, we see that there was a gain in representativity in socio-demographic variables (from 0.77 to 0.87), but not in psychopathology (from 0.90 to 0.89), so that we may conclude that the reduction in selectivity is mainly obtained by the demographic variables.

That the mixed-mode design realized lower selectivity measured by the *R*-indicator, suggests that the at home interview and the telephone interview were able to reach specific groups of respondents who would have been lost if data collection had been restricted to the single mode design. Table 3 shows characteristics of the three response modes: on-site, at home and telephone interviews. Comparing the different modes in the mixed-mode survey, we found that respondents who did an at home interview were generally older, more often female, less urbanized, lived further away from the interview site and more often had a partner, while the respondents who did a telephone interview were generally younger, more often non-Northern European, lived in a more urban area, were more often employed, and more often single. Differences with respect to psychopathology were insignificant between the response groups. This suggests that with respect to psychopathology, the at home interview and the telephone interview reached rather similar individuals in terms of their psychiatric status.

Table 2. Odds ratios of attrition after 2 years according to baseline socio-demographic and psychopathology variables

	Single Mode			Mixed-mode		
	model 1	model 2	model 3	model 1	model 2	model 3
Socio-demographic variables						
Age (per y increase)	0.98***		0.99***	0.98***		0.99**
Female (versus male)	1.21+		1.23*	1.04		1.08
Education (per y increase)	0.90***		0.91***	0.90***		0.91***
Non-Northern European descent vs Northern European descent	1.91***		1.85***	1.82**		1.71**
Sampling Site						
Amsterdam						
Leiden	0.40***		0.39***	0.62***		0.61***
Groningen	0.29***		0.28***	0.53***		0.53***
High vs low urbanization	0.80*		0.84	0.93		1.00
Previous participation in Research	1.46**		1.45**	0.75		0.72
Distance to site (per 100 km)	1.25		1.58**	0.53**		0.73
Not employed vs employed	1.00		0.93	1.25+		1.14
No partner vs partner	0.88		0.83+	0.9		0.83
No. of chronic diseases	1.11*		1.08	1.08		1.04
Psychopathology						
Depressive disorders						
MDD		1.25*	1.27*		1.78***	1.60***
Dysthymia		1.42*	1.56**		1.29	1.31+
Anxiety disorders						
Social phobia		1.09	1.07		1.11	1.08
Panic disorder with agoraphobia		1.32*	1.27+		1.43*	1.23
Panic disorder without agoraphobia		1.27	1.26		1.31	1.22
Agoraphobia		0.96	0.88		1.31	1.16
General anxiety disorder		1.21	1.18		1.25	1.14
Alcohol use disorders						
Alcohol dependence		0.96	1.08		1.09	1.14
Alcohol abuse		1.06	1.11		1.23	1.19
n	2981	2981	2981	2981	2981	2981
Pseudo R^2	0.07	0.01	0.08	0.05	0.03	0.06
R-indicator	0.77	0.90	0.75	0.87	0.89	0.84

*Notes. ***: p<0.001, **: p<0.01, *: p<0.05, +: p<0.10*

Table 3. Baseline socio-demographic and psychopathological characteristics (means and percentages) of response on non-response groups for single and mixed-mode design

	on-site *n*=2330	at home *n*=126	telephone *n*=140	*p-value*
Socio-demographic variables				
Age – in years (SD)	42.2 (13.2)	43.7 (11.5)	37.2 (11.3)	.000
Female vs Male	65.2	73.8	73.6	.022
Education – in years (SD)	12.3 (3.2)	11.7 (3.3)	11.8 (3.6)	.024
Non-Northern European descent vs Northern European descent	4.2	4.0	10.7	.001
Sampling Site				
Amsterdam	36.7	69.8	62.1	.000
Leiden	32.0	13.5	16.4	.000
Groningen	31.4	16.7	21.4	.000
High vs low urbanization	70.8	40.5	77.1	.000
Distance in km to interview site (SD)	16.2 (31.8)	66.5 (53.7)	19.5 (36.9)	.000
Previous participation in research	17.8	69.0	15.7	.000
Not employed vs employed	35.4	34.9	25.7	.065
No partner vs partner	30.9	21.4	35.0	.041
Number of chronic diseases	0.9	1.0	0.9	.470
Psychopathology				
Depressive disorders				
MDD	39.1	30.2	41.4	.109
Dysthymia	9.5	10.3	15.0	.106
Anxiety disorders				
Social phobia	22.1	23.0	23.6	.899
Panic disorder with agoraphobia	13.4	14.3	16.4	.591
Panic disorder without agoraphobia	8.6	10.3	9.3	.773
Agoraphobia	6.5	4.8	2.9	.171
General anxiety disorder	15.2	12.7	20.0	.214
Alcohol use disorders				
Alcohol dependence	15.2	8.7	17.1	.109
Alcohol abuse	11.7	8.7	12.9	.535

Note. The p-values were obtained from ONEWAY ANOVAs.

We mentioned earlier that the change from single mode to mixed-mode increased the selectivity with respect to depressive disorder MDD, both in a bivariate sense (see Table 1) and in a multivariate sense (see Table 2). The percentage of MDD sufferers measured at baseline differed for the four response groups. We found that 58% of the non-responders (in the mixed-mode design) had MDD at baseline, while the equivalent percentages for the responding groups were 39%, 30% and 41% for the on-site, at home and telephone group respectively, with a combined mean of 39%. The contrast of 58% versus 39% was greater than the contrast in the single mode design, which was 49%, (obtained from adding the at home and the telephone group to the non-response group) versus 39%. This increase in contrast, explains the increase in selectivity shown in Table 1. The increased selectivity confirms that the at home interviews and the telephone interviews reached respondents who – with respect to psychiatric variables – were more similar to the on-site respondents than to the non-responders.

3.3. Survey mode

To find out if mode effects occurred, we studied differences between response groups (on-site interviewing, at home interviewing and telephone interviewing) with respect to two key variables: the number of DSM-IV depression symptoms, and self-rated health. As a result of selection – the fact that specific respondents preferred a certain mode – we expected that mean scores of these variables would differ for different groups. If there were no mode effects, we would expect that controlling for selection by using statistical matching would make these differences disappear. If on the other hand, controlling for selection did not (entirely) resolve these differences, we would conclude that mode effects existed.

Table 4 shows mean scores of the two variables for the three response groups (see the results in the bottom part of Table 4 in the columns for the unmatched samples, i.e. columns 2, 3 and 6; notice that columns 3 and 7 are identical). Since we expected selection effects, it was unexpected that the differences of the mean scores between the response groups hardly differed at all, both for the number of MDD symptoms and for self-rated health. As a result, the differences in means scores in the unmatched samples were statistically insignificant, as demonstrated by the test statistics at the bottom part of Table 4, for the comparison between unmatched samples for on-site face-to-face interviews and telephone interviews (columns 2 and 3), and for the comparison between unmatched samples for at home face-to-face interviews and telephone interviews (columns 6 and 7).

To separate selection effects from mode effects, we applied the method of Coarsened Exact Matching (CEM). As matching variables, we selected all variables that were significantly related to the response groups (see Table 3): age, gender, education, Northern European descent, sampling site, urbanization, distance to interview site, previous participation in research and partnership status. Next to these socio-demographic variables, we added the baseline versions of the two key variables: 'the number of MDD symptoms at baseline' and 'Self-rated health at baseline'. Creating matching cells for the binary variables (female, gender, Northern European descent, urbanization, previous participation, partnership) and for sampling site was straightforward: two and three cells respectively. For education, we created three cells: up to seven years, eight to twelve years and more than twelve years. For the remaining continuous variables (age, distance, number of MDD symptoms at baseline and self-rated health at baseline), we created two groups by specifying cut-off values: 45 for age, 40 for distance, 5 for number of MDD symptoms and 80 for self-rated health. The CEM procedure resulted in matching 713 on-site interviews to 112 telephone interviews, and in matching 25 at home interviews to 33 telephone interviews. Table 4 shows the results for unmatched and matched samples. Notice that the results for the matching variables in the matched subsamples of the on-site face-to-face sample and the telephone sample show that the matching was successful, as can be concluded from the identical columns 4 and 5 in Table 4 with respect to the matching variables. Similarly we conclude that the matching of the at home face-to-face sample and the telephone sample was successful (see columns 8 and 9). Computing the difference in means scores for the outcome variables of the follow-up measurement for the matched samples (see the test statistics at the bottom part of Table 4 in columns 4 and 5 and in columns 8 and 9), showed no significant differences. These results suggest that there is no mode effect between face-to-face interviewing and telephone interviewing for these two key outcome variables.

Table 4. Propensity score matching results, showing percentages of coarsened matching variables for unmatched and matched differences between face-to-face and telephone interviewing and showing unmatched and matched mean scores of the outcome variables 'number of MDD symptoms' and 'self-rated health' at follow-up wave and corresponding test statistics

matched or unmatched samples:	**unmatched**		**matched**		**unmatched**		**matched**	
interview setting:	**on-site f2f**	**telephone**	**on-site f2f**	**telephone**	**at home f2f**	**telephone**	**at home f2f**	**telephone**
column 1	*column 2*	*column 3*	*column 4*	*column 5*	*column 6*	*column 7*	*column 8*	*column 9*
Matching variables (baseline)								
Age (>=45)*	47.6%	27.9%	28.6%	28.6%	51.6%	27.9%	30.3%	30.3%
Gender: female*	65.2%	73.6%	72.3%	72.3%	73.8%	73.6%	90.9%	90.9%
Education: primary	5.7%	11.4%	5.4%	5.4%	8.7%	11.4%	9.1%	9.1%
Education: secondary	57.1%	57.1%	63.4%	63.4%	60.3%	57.1%	57.6%	57.6%
Education: tertiary	37.2%	31.4%	31.3%	31.3%	31.0%	31.4%	33.3%	33.3%
Non-Northern European descent*	4.2%	10.7%	6.3%	6.3%	4.0%	10.7%	0.0%	0.0%
Sampling site: Amsterdam	36.7%	62.1%	62.5%	62.5%	69.8%	62.1%	81.8%	81.8%
Sampling site: Leiden	32.0%	16.4%	17.9%	17.9%	13.5%	16.4%	3.0%	3.0%
Sampling site: Groningen	31.4%	21.4%	19.6%	19.6%	16.7%	21.4%	15.2%	15.2%
High urbanization *	70.8%	77.1%	78.6%	78.6%	40.5%	77.1%	72.7%	72.7%
Distance to interview site (>=40 km)*	13.7%	17.1%	11.6%	11.6%	61.1%	17.1%	21.2%	21.2%
Previous participation*	17.8%	15.7%	13.4%	13.4%	69.0%	15.7%	18.2%	18.2%
No partner*	30.9%	35.0%	32.1%	32.1%	21.4%	35.0%	18.2%	18.2%
Number of MDD symptoms (>=5)*	70.1%	73.6%	69.6%	71.4%	65.9%	73.6%	75.8%	81.8%
Self-rated health (>=80)*	53.5%	52.1%	53.6%	53.6%	51.6%	52.1%	54.5%	54.5%
Sample size	2330	140	713	112	126	140	25	33
Outcome variables (follow-up)								
Mean number of MDD symptoms (SD)	2.96 (3.42)	2.86 (3.44)	2.83 (3.53)	2.70 (3.37)	3.11 (3.40)	2.86 (3.44)	3.38 (3.50)	2.36 (3.10)
test statistics	t(2467) = -0.33, p = .743		t(823) = -0.36, p = .717		t(264) = -0.59, p = .558		t(56) = -1.17, p = .248	
Mean self-rated health (SD)	78.0 (16.2)	78.8 (15.8)	78.4 (15.9)	79.9 (15.3)	78.8 (17.4)	78.8 (15.8)	81.9 (16.7)	79.6 (15.1)
test statistics	t(2451) = 0.49, p = .621		t(816) = 0.81, p = .419		t(253) = -0.01, p = .991		t(51) = -0.53, p = .600	

*Notes. f2f refers to a face-to-face interview setting. SD = standard deviation * To save space, the complementary groups are not shown in the table.*

4. Discussion

This study evaluated the introduction of a mixed-mode survey design in the NESDA follow-up study in order to reduce the problem of attrition. Introducing a mixed-mode design adds channels to reach respondents, and can consequently only lower the attrition rate if we compare this new situation to a single mode design. We found that if no at home interviews and no telephone interviews had been held, the attrition rate would have been 22%, instead of 13% which we obtained in the mixed-mode design. From this observation, we concluded that the mixed-mode design reduced the attrition level by 9%. Studying the selectivity of attrition, we found that the introduction of the mixed-mode design decreased selectivity with respect to combined socio-demographic and psychiatric variables as measured by the *R*-indicator. However, the decrease in selectivity is not the same for all (sets of) variables. There appears to be a drastic decrease in selectivity with respect to socio-demographic variables, while a (small) increase in selectivity appears for psychiatric variables. This suggests that the introduction of the mixed-mode survey made the sample more representative with respect to demographic variables, and at the same time, slightly less representative with respect to psychiatric variables.

Another aspect that we considered was the fact that the introduction of mixed-mode design may also introduce problems related to mode effects. However, in our sample we fortunately did not find any evidence for mode effects. This finding is in line with those of Sobin et al. (1993), who found no differences between face-to-face methods for the diagnosis of lifetime psychiatric disorders. Our results are also in line with those of Voogt & Saris (2005), who introduced a mixed-mode design together with the use of a so-called central question, in a cross-sectional election survey. They concluded that the mixed-mode survey design led to larger response rates and less non-response bias. The positive effects appeared to be larger than the negative effects from the increase in response bias (or measurement error).

Our study dealt with the direct consequences of the introduction of a mixed-mode design only, and so far, we have not taken long term consequences into account. The reduction in attrition from 22% to 13% may not be considered large, but if the same attrition were maintained over the total of four waves, then the attrition would accumulate to $1 - (1 - 0.22)^4 = 63\%$ for the single mode design, while it would be $1 - (1 - 0.13)^4 = 43\%$ for the mixed-mode design. Keeping attrition low is important for longitudinal surveys, as sample members, once lost, are irreplaceable. Lynn (2011) discusses an experiment where mixed-mode survey design was introduced to reduce cost, and found that the lower attrition rate implied by the mixed-mode survey persisted in a subsequent wave. It is not yet clear what the long-term consequences of the introduction of the mixed-mode design would be for the level and selectivity of the attrition. This may be a topic for future research.

In our study the survey modes that were added to the single mode design were “interviews at the respondents’ home” and “telephone interviews”. The addition of other survey modes to reduce attrition should be studied: web-based methods for self-interviewing, interviews by on-line video and audio conversation methodology such as Skype or FaceTime (Loftis, Zelko, Jewett, & Meyer, 2012) or web-based methods using avatar interviewers (Bell, Castronova, & Wagner, 2009; Malakhoff & Jans, 2011).

In sum, we conclude that the mixed-mode design lowered the level of attrition, but did not necessarily reduce the selectivity of the attrition. This finding implies that the problem of attrition is not solved by the introduction of the mixed-mode design. Naïve analyses ignoring the attrition may well lead to misleading results. Fortunately, many variables at baseline are available to study the dependency of the attrition on key variables. Techniques to correct for the selective attrition, such as weighting or imputation (Goldstein, 2009) should be considered. Implementing these techniques is a challenge to all researchers analyzing data from longitudinal studies. Yet, the decrease in attrition by nine percentage points, the increase in the R-indicator for the combined set of socio-demographic and psychopathology variables, and the retention of sample members that would have otherwise been lost permanently from the longitudinal study, made the introduction of the mixed-mode survey design, very worthwhile.

Acknowledgements

The infrastructure for the NESDA study (www.nesda.nl) is funded through the Geestkracht program of the Netherlands Organisation for Health Research and Development (Zon-Mw, grant number 10-000-1002) and is supported by participating universities and mental health care organizations (VU University Medical Center, GGZ inGeest, Arkin, Leiden University Medical Center, GGZ Rivierduinen, University Medical Center Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Scientific Institute for Quality of Healthcare (IQ healthcare), Netherlands Institute for Health Services Research (NIVEL) and Netherlands Institute of Mental Health and Addiction (Trimbos Institute).

References

Alegria, M., Jackson, J. S., Kessler, R. C., & Takeuchi, D. (2008). *Collaborative Psychiatric Epidemiology Surveys (CPES), 2001-2003.* (Rep. No. ICPSR20240-v6). Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Bell, M. W., Castronova, E., & Wagner, G. G. (2009). *An expansion of survey data collection methods to the virtual worlds by means of VDCI.* (Rep. No. Research Notes of the German Council for Social and Economic Data 39). German Council for Social and Economic Data (RatSWD).

Bijl, R. V., Ravelli, A., & van Zessen, G. (1998). Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology*, 33, 587-595.

Bowling, A. (2005). Mode of Questionnaire Administration Can Have Serious Effects on Data Quality. *Journal of Public Health*, 27, 281-291.

Couper, M. P. & Ofstedal, M. B. (2009). Keeping in Contact with Mobile Sample Members. In P.Lynn (Ed.), *Methodology of Longitudinal Surveys* (pp. 183-203). Chichester: John Wiley & Sons, Inc.

de Graaf R., Bijl, R. V., Smit, F., Ravelli, A., & Vollebergh, W. A. (2000). Psychiatric and sociodemographic predictors of attrition in a longitudinal study: The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *American Journal of Epidemiology*, 152, 1039-1047.

de Leeuw, E. D. (1992). *Data Quality in Mail, Telephone, and Face-to-face Surveys*. Amsterdam: TT-Publicaties.

de Leeuw, E. D. (2005). To Mix or Not to Mix Data Collection Modes in Surveys. *Journal of Official Statistics*, 21, 233-255.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009*). Internet, mail, and mixed-mode surveys: The tailored design method, Third Edition.* Hoboken, NJ: John Wiley & Sons Inc.

Farmer, M. E., Locke, B. Z., Liu, I. Y., & Moscicki, E. K. (1994). Depressive symptoms and attrition: the NHANES I epidemiologic follow-up study. *International Journal of Methods in Psychiatric Research*, 4, 19-27.

Fessler, P., Kasy, M., & Lindner, P. (2012). Survey mode effects on income inequality measurement. http://scholar.harvard.edu/files/kasy/files/fesslerkasylindner-nov2012_0.pdf [On-line]. Available: http://scholar.harvard.edu/files/kasy/files/fesslerkasylindner-nov2012_0.pdf

Fischer, E. H., Dornelas, E. A., & Goethe, J. W. (2001). Characteristics of people lost to attrition in psychiatric follow-up studies. *Journal of Nervous and Mental Disease*, 189, 49-55.

Goldstein, H. (2009). Handling attrition and non-response in longitudinal data. *Longitudinal and Life Course Studies*, 1, 63-72.

Groves, R. M. (1989). *Survey Errors and Survey Costs*. New York: Wiley-Interscience.

Groves, R. M. & Couper, M. P. (1998). *Nonresponse in Household Interview Surveys*. New York: Wiley.

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone vs. Face-to-Face Interviewing of National Probability Samples With Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias. *Public Opinion Quarterly*, 67, 79-125.

Iacus, S. M., King, G., & Porro, G. (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis*, 20, 1-24.

Jackle, A., Roberts, C., & Lynn, P. (2006). *Telephone versus Face-to-Face Interviewing: Mode Effects on Data Quality and Likely Causes* (Rep. No. ISER Working Paper 2006-41). Colchester: Institute for Social and Economic Research, University of Essex.

Lamers, F., Hoogendoorn, A. W., Smit, J. H., van Dyck, R., Zitman, F. G., Nolen, W. A., & Penninx, B.W. (2012). Sociodemographic and psychiatric determinants of attrition in the Netherlands Study of Depression and Anxiety (NESDA). *Comprehensive Psychiatry*, 53, 63-70.

Loftis, C., Zelko, H., Jewett, C., & Meyer, M. (2012). A VOIP-based Telephone Interview and Recording System using Skype-TM. In Paper presented at the 46th Annual International Field Directors & Technology Conference, Orlando, Florida.

Lorenc, B. (2010). Item Nonresponse Analysis for a Mixed-Mode Survey. In M.Carlson, H. Nyquist, & M. Villani (Eds.), *Official Statistics - Methodology and Applications in Honour of Daniel Thorburn* (pp. 117-136). Stockholm: Department of Statistics, Stockholm University.

Lugtig, P., Lensvelt-Mulders, G. J. L. M., Frerichs, R., & Greven, A. (2011). Estimating nonresponse bias and mode effects in a mixed-mode survey. *International Journal of Market Research*, 53, 669-686.

Lynn, P. (2011). *The Effect of a Mixed-mode Wave on Subsequent Attrition in a Panel Survey: Evidence from the Understanding Society Innovation Panel* (Rep. No. Understanding Society Working Paper 2011-06). Colchester: ISER, University of Essex.

Malakhoff, L. & Jans, M. (2011). Towards usage of avatar interviewers in surveys. Survey Practice [On-line]. Available: http://surveypractice.wordpress.com/2011/06/20/usage-of-avatar/

Penninx, B. W., Beekman, A. T., Smit, J. H., Zitman, F. G., Nolen, W. A., Spinhoven, P., Cuijpers, P., De Jong, P.J., Van Marwijk, H.W.J., Assendelft, W.J.J., Van Der Meer, K., Verhaak, P., Wensing, M., De Graaf, R., Hoogendijk, W.J., Ormel, J., Van Dyck. R. (2008). The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *International Journal Methods Psychiatric Research*, 17, 121-140.

Schouten, B., Cobben, F., & Bethlehem, J. G. (2009). Indicators for the representativeness of survey response. *Survey Methodology*, 35, 101-113.

Sobin, C., Weissman, M. M., Goldstein, R. B., Adams, P. B., Wickramaratne, P. J., Warner, V. Lish, J. (1993). Comparing telephone and face-to-face interview methods for the diagnosis of lifetime psychiatric disorders in a family study. *Psychiatric Genetics*, 3, 227-233.

Susser, E., Schwartz, S., Morabia, A., & Bromet, E. J. (2006). *Psychiatric Epidemiology: Searching for the Causes of Mental Disorders.* Oxford: Oxford University Press.

Tipping, S., Hope, S., Pickering, K., Erens, B., Roth, M. A., & Mindell, J. S. (2010). The effect of mode and context on survey results: analysis of data from the Health Survey for England 2006 and the Boost Survey for London. BMC Medical Research Methodology, 10:84, 1-8.

Van der Laan, P., & Van Nunspeet, W. (2009). *Modernising household surveys in the Netherlands: Design, Efficiency Gains and Perspectives* (Rep. No. Discussion paper 09044). Statistics Netherlands.

Vannieuwenhuyze, J., Loosveldt, G., & Molenberghs, G. (2010). A Method For Evaluating Mode Effects In Mixed-Mode Surveys. *Public Opinion Quarterly*, 74, 1027-1045.

Voogt, R. J. J., & Saris, W. E. (2005). Mixed-mode Designs: Finding the Balance Between Nonresponse Bias and Mode Effects. *Journal of Official Statistics*, 21, 367-387.

Watson, N., & Wooden, M. (2009). *Identifying Factors Affecting Longitudinal Survey Response*. In P.Lynn (Ed.), Methodology of Longitudinal Surveys (Pp 157-181). Chichester: John Wiley & Sons, Inc.

Weuve, J., Tchetgen 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, 119-128.

World Health Organization (1997). *Composite International Diagnostic Interview (CIDI)*. Geneva: World Health Organization.

Identifying biases arising from combining census and administrative data – the fertility of migrants in the Office for National Statistics Longitudinal Study

James Robards [1,2]
james.robards@soton.ac.uk
Ann Berrington, [1,2,3]
Andrew Hinde [1,3]
[1]Division of Social Statistics and Demography, School of Social Sciences, University of Southampton
[2]ESRC Centre for Population Change, School of Social Sciences, University of Southampton
[3]Southampton Statistical Sciences Research Institute, University of Southampton

(Received January 2013 Revised June 2013) doi:10.14301/llcs.v4i3.244

Abstract

Demographic research is increasingly making use of longitudinal and life history data, given its strong analytical potential. Such data are frequently produced by linking and matching records from multiple sources. Where this is the case, there is the potential for a person's appearance in one source of data to be conditional on an event in another source of data. This can lead to bias in estimating occurrence/exposure rates concerning the event in question, unless the correct exposure can be identified. Achieving the latter requires understanding the reasons governing entry to the data. The Office for National Statistics (ONS) Longitudinal Study (LS) for England and Wales is a 1% sample of the population, constructed by combining data from the census, vital registrations (births and deaths) and the National Health Service Central Register (NHSCR). This paper examines the difficulties in obtaining the correct exposure for rates in complex data sets by studying the fertility of migrants using the ONS LS. Three tests in relation to the fertility of female migrants to England and Wales illustrate the possible association between exposure to risk and subsequent events. The first identifies the ability of the data set to record new migrants, the second is concerned with the mode of entry to the data set and subsequent fertility, and the third illustrates how the recorded fertility of migrants depends upon the way migration is measured.

Keywords

Office for National Statistics Longitudinal Study, migration, National Health Service Central Register, fertility, longitudinal data

1. Introduction

In demographic research there has been increasing interest in the exploration of associations and causation using longitudinal data sources, especially since the reporting of life course history and events in survey data can be incomplete (Murphy, 2009; Ní Bhrolcháin et al., 2011. Within longitudinal and life course research, it is becoming more common to combine data from different sources to produce complex data sets (Ford et al., 2009; Lyons et al., 2009). Indeed there is currently discussion as to whether linked administrative data sources could replace the decennial census for England and Wales (Ralphs and Staples, 2012).

Linked data sets can provide detailed information on dates of events and event sequencing. However, complications can occur as a result of the combination of data sources. Appearance in some sources can be related to the

events which the researcher wishes to measure using the combined data set, leading to potential bias in occurrence/exposure rates. In this paper we examine the potential biases that arise as a result of associations between life events and capture within a particular source. Using the Office for National Statistics (ONS) Longitudinal Study (LS), the issue is illustrated by studying female LS members' entry into the LS and the degree to which the timing of entry is related to subsequent fertility.

The ONS LS may be suitable for studying migrant fertility because of its large sample of migrants and accurate recording of births from registration data. Few studies have considered the timing of fertility among migrants to England and Wales (Waller, Berrington and Raymer, 2012); research has instead considered the absolute level of migrant fertility (Tromans, Natamba and Jefferies, 2009; Zumpe, Dormon and Jefferies, 2012). In other countries elevated fertility shortly after migration has been identified (Toulemon, 2004).

The ONS LS is composed of data from the 1971-2011 censuses, the National Health Service Central Register (NHSCR) and the vital registration system (births and deaths) (Adelstein, 1976; Hattersley and Creeser, 1995; Blackwell, Lynch, Smith and Goldblatt, 2003). The sample consists of persons born on one of four dates of the year, representing around 1% of the population of England and Wales. Individuals born in England and Wales on one of the four dates become new LS members at birth. For new migrants, entry to the ONS LS is made either by registering with a National Health Service (NHS) General Practitioner (GP) and reporting that the previous address was overseas, or by being recorded at the decennial census for the first time.

2. Composition of the ONS LS and implications for analysis

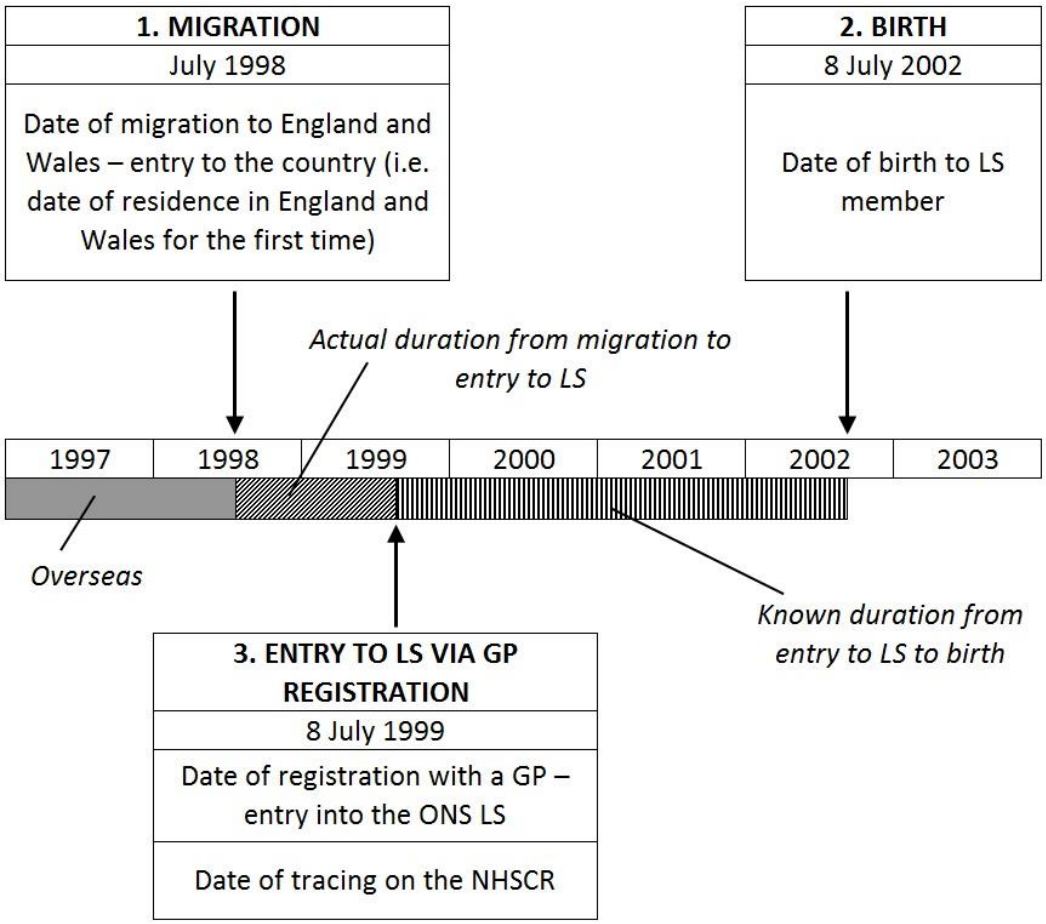
2.1 Data combined to produce life histories in the ONS LS

At each census, records for persons born on an LS birth date are selected. From the census, socio-economic and household information is recorded for the LS member. An attempt is then made to trace and match each census record to the NHSCR[1]. This enables the attachment of subsequent events recorded on the NHSCR for these LS members. Births on the four LS dates, and deaths of those born on the four dates, are extracted from vital registration data, as are all births to female LS members. Linkage rates have improved since the early 1990s (Blackwell et al., 2003). Immigrants with one of the four birth dates are captured either through censuses (persons enumerated at each census who were born on one of the four dates and who were not already LS members are added), or by newly registering with a NHS GP. In the latter context, the LS defines 'immigrants' as persons arriving from outside England and Wales, including Scotland, Northern Ireland and the Channel Islands, and '[t]he category of immigrant includes not only those individuals who describe themselves to their general practitioners as such, but also those who, having quoted a previous address abroad, cannot be matched to an existing NHS number' (Hattersley and Creeser, 1995 p.25).

Ideally, new immigrants will register with a GP soon after they arrive in England and Wales and be entered on the NHSCR, so that they will be captured by the LS close to the actual date they arrived. However, as there is no legal requirement for registration with a GP, many new arrivals in fact are not identified until they appear in the census.

Figure 1 shows a hypothetical entry to the data set for a female LS member and the way in which this is captured within the LS. Time point 1 (July 1998) is the date of migration. The date of birth of children born to migrants is known exactly (e.g. this woman gave birth on 8 July 2002 – time point 2). However, there was a time lag of one year before she registered with an NHS GP (on 8 July 1999 - point 3). This time lag is denoted by the period labelled 'Unknown duration from migration to entry to LS'. The duration is 'unknown' because, although the migration was in July 1998, the exact date of migration is not recorded. Upon registration with a GP, she is recorded in the NHSCR and enters the LS. The recorded exposure to the risk of an event (e.g. death or childbearing) will be the duration from the date of NHSCR registration, and will not include the exposure between the actual date of migration and the date of registration with a GP.

Figure 1. Terminology used to describe the entry of migrants to the ONS LS



2.2 Implications of the construction of the ONS LS for fertility research

Because of the way in which immigrants enter the LS through NHSCR registration it is important to consider the possible association between a woman's entering the exposed-to-risk and the risk of a subsequent birth. If migrant women who were pregnant, or intending to become pregnant, were more likely to register with a GP than other women, there would be an association between entry into the LS exposed-to-risk and the chance of a subsequent birth, and fertility rates of migrants computed using the LS would be inflated. Similarly, a birth to an immigrant mother who was born on one of the four LS dates but who was not in the NHSCR would probably trigger NHSCR registration, again leading to an association between the timing of birth and the timing of entry into the LS exposed-to-risk. The presence and magnitude of such a bias has not hitherto been studied.

3. Research questions and method

3.1 Research questions

We answer three related questions which attempt to quantify potential biases in using the LS for migrant fertility research:

1. How complete is the capture of new migrants to England and Wales in non-census years: what proportion is first identified by the census?
2. Is there evidence of an association between registering with a GP and the timing of a subsequent birth?
3. Is there is a relationship between the mode of entry to the ONS LS (between 1991 and 2001) and fertility after the 2001 census?

The questions are related as they study bias which could be arising from entry of migrants to the ONS LS (question 1), the fertility of migrants who register with a GP (question 2) and the fertility of migrants at the 2001 census (question 3).

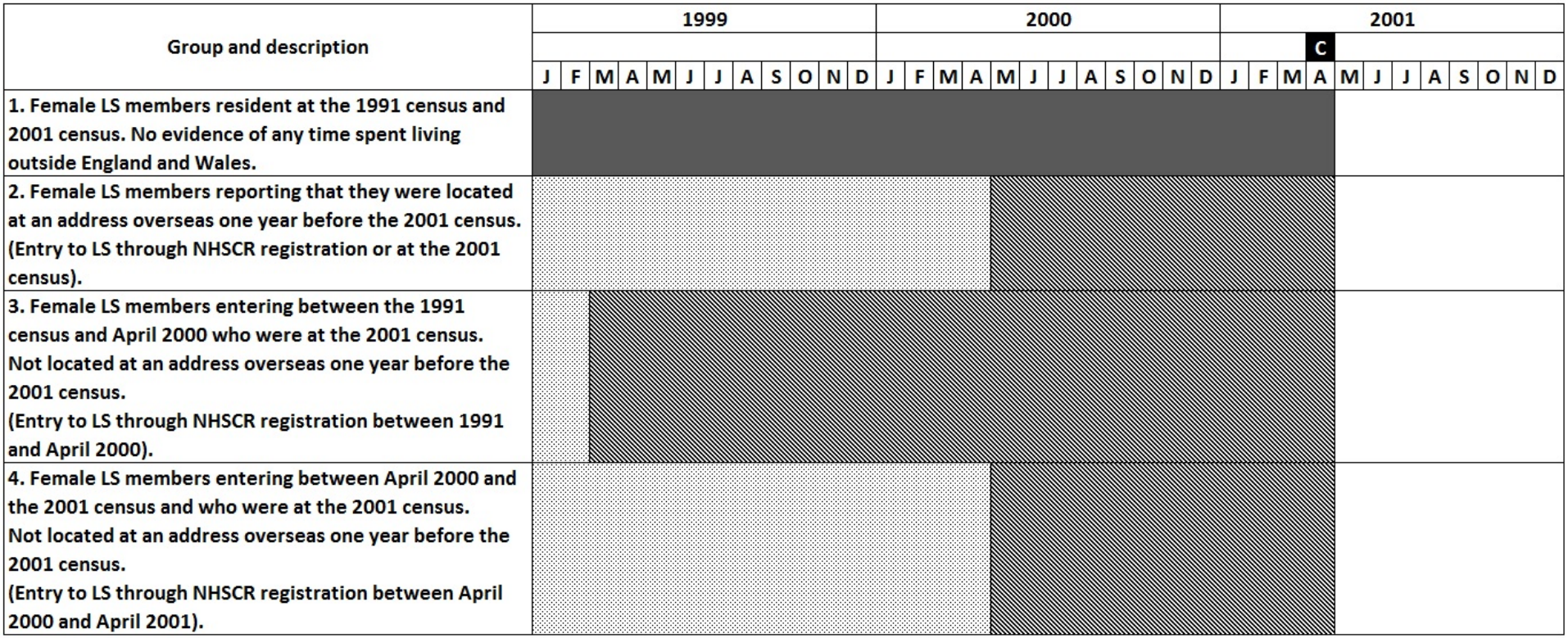
3.2 Method

To answer Question 1, the number of new LS female migrant members entering the data set in the five years prior to the census is divided by the number entering at the census for the first time. The sample is composed of LS members who entered via an NHSCR registration in the years 1996-2000 and were at the 2001 census, and those LS members who entered at the 2001 census for the first time. New entrants at the 2001 census are defined as those female migrant LS members who did not enter the data set at any point in the past and have not been resident at a past census. The analysis is by single year of age (based on age at the 2001 census) to allow the identification of age-group trends. The years 1996-2000 for new migrants will be used as they provide enough information to answer the question, while keeping to a minimum the risk that persons arriving in the period before the census may have left England and Wales before 2001. The biggest source of incomplete information in the LS for migrants between 1991 and 2001 was 'unrecorded embarkation' or persons leaving England and Wales without leaving a record of their departure.

To answer Question 2, the duration to the first birth after the date of GP registration (date of migration) is recorded in months for each female migrant into the ONS LS who was captured by the NHSCR. The number of first births per annum by duration from GP registration is calculated for the periods 1991-2000 and 2001-2006. The sample is composed of migrants who were identified by NHSCR registrations between 1991-2000 or 2001-2006 but who may have subsequently left again.

To answer Question 3, we identify four different types of female migrant entering the LS in the 1991-2001 period (Figure 2), and calculate age-specific and hence total fertility rates for each during the calendar years 2001, 2002, 2003 and 2004. The 2001 census asked all those living in England and Wales on the census night, for details of their place of residence one year before the census. Our four groups are defined as in Figure 2. The sample selection is described for each of the four types of migrant and, as with Question 1, we exclude migrants with a date of birth discrepancy[2].

Figure 2. Four types of female ONS LS migrant at the 2001 census

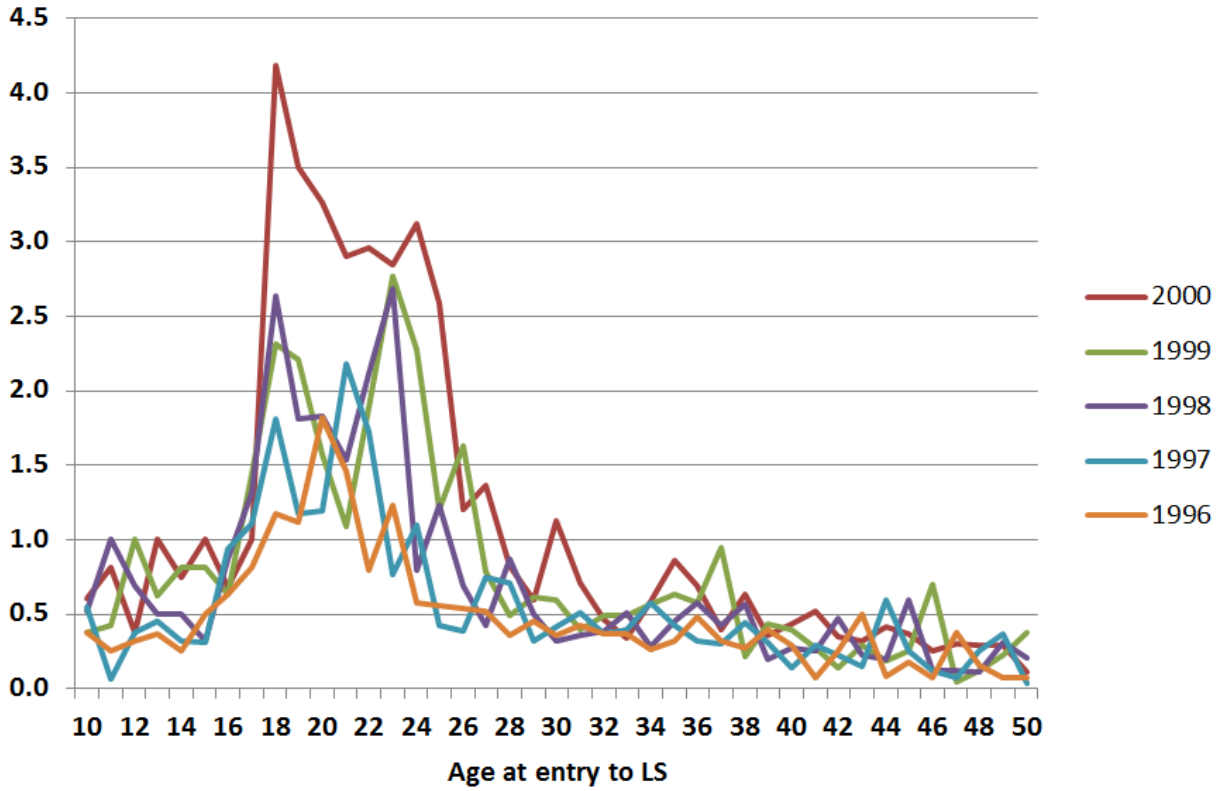


4. Results

4.1 How complete is the capture of new migrants to England and Wales in non-census years: what proportion is first identified by the census?

We compare the number of female LS members entering through an NHSCR GP registration in the five years before the 2001 census, with the number of female LS members entering at the 2001 census (Figure 3). A ratio below 1.0 indicates that more immigrants were first entered into the LS as a result of being present in the 2001 census, than registering with a GP in one of the years before the census. Among women aged 18-28 years in 2001, more entered the LS through registration with a GP in each of the years preceding the census, than did through being present at the census with an LS date of birth. For other ages, the ratios of entries are below 1.0, and LS members aged over 38 years at 2001 had the lowest ratios, typically below 0.5. We see high proportions of inter-censal capture among LS members entering around age 18 years. The increase in the ratio from age 17 years to age 18 / 19 years coincides with possible demand for reproductive health services. Women in the key reproductive age groups are more likely to register with a GP and enter the LS than women in older and younger age groups.

Figure 3. Ratio of ONS LS joiners in each of the years 1996-2000 to joiners at 2001 census in the same cohort by age at joining



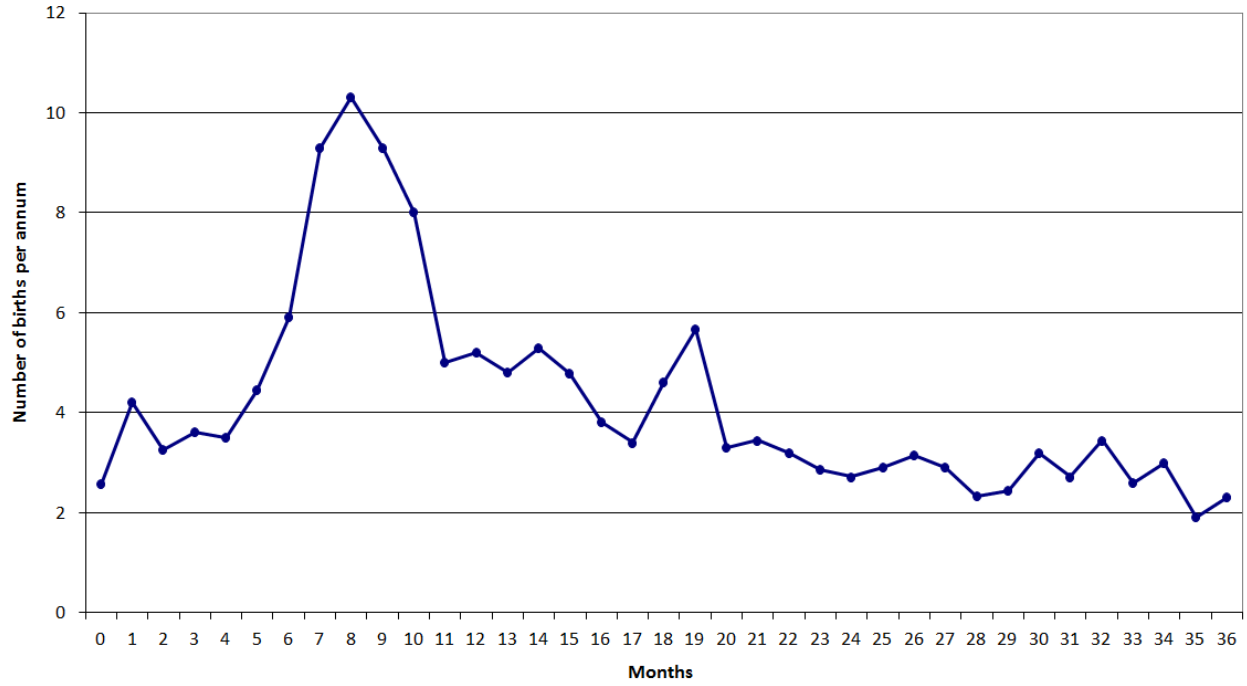
Source: Authors' analyses based on ONS LS.

4.2 Is there evidence of an association between registering with a GP and the timing of a subsequent birth?

We calculate the number of births which occurred on average at each duration since GP registration in two periods: 1991-2000 (Figure 4) and 2001-2006 (Figure 5). Figure 4 (1991-2000) reveals a peak in first births to new entrants eight months after entry to the LS. Following this, there is a decline in the number of first births to a point 11 months after the registration on NHSCR. A second rise in first births is observed around the 18 months period, after which there is a gradual decline up to a point 36 months after entry. Figure 5 (2001-2006) shows a more pronounced peak in first birth numbers in the eighth month after entry to the LS. The number of first births falls from this peak up to about 18 months after registration, and at longer durations remains roughly constant. The pattern is clear. Migrant women are especially likely to

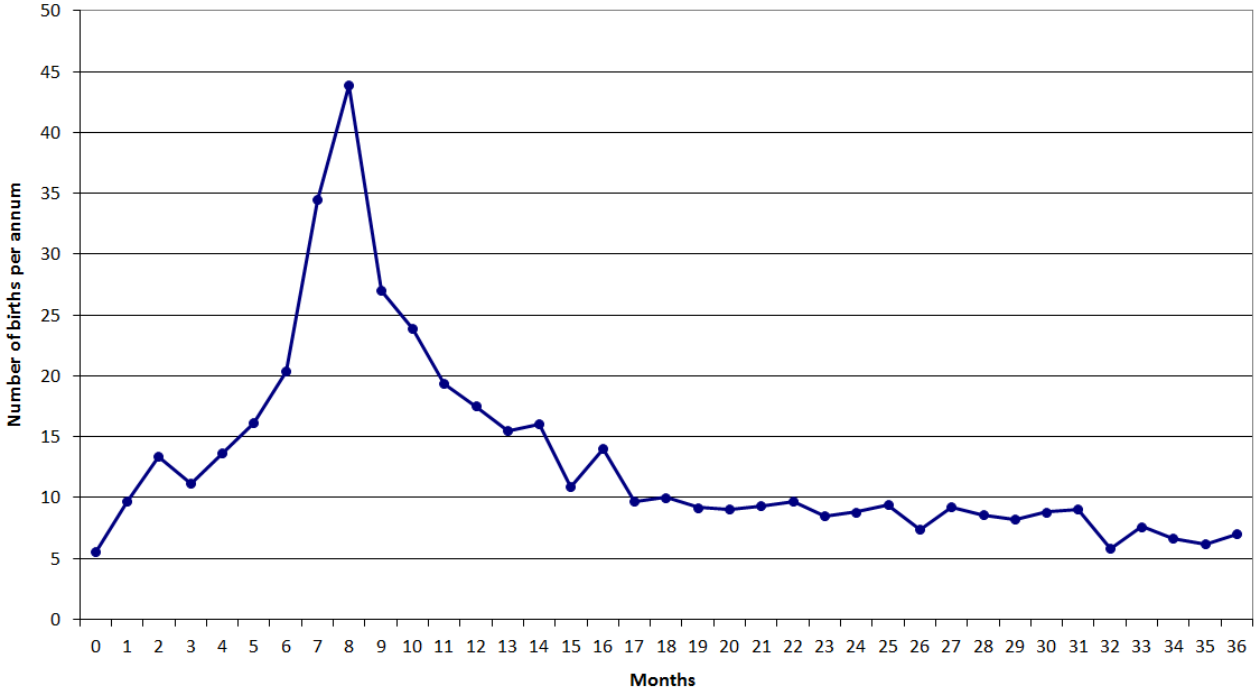
register with a GP around the time they become pregnant, leading to a strong association between registration with a GP (and hence entry into the LS) and a subsequent birth. If the date of registration with a GP is used as a proxy for the date of migration, the fertility of recent migrants to England and Wales calculated using LS data will be overestimated.

Figure 4. Average number of first births per month after registration on the NHSCR (and entry to ONS LS) for the period 1991-2000



Source: Authors analyses based on ONS LS.

Figure 5. Average number of first births per month after registration on the NHSCR (and entry to ONS LS) for the period 2001-2006



Source: Authors analyses based on ONS LS.

4.3. Is there is a relationship between the mode of entry to the ONS LS (between 1991 and 2001) and fertility after the 2001 census?

Table 1 presents total fertility rates (TFRs) for 2001-2004 for the four migrant groups listed in section 3.2. This time period was selected to enable the use of details on LS members recorded at the 2001 census; it allows a period of exposure long enough to provide reliable rates, while being short enough to minimise the impact of attrition.

Group 1, women continuously resident between 1991 and 2001, have a TFR of around 1.5 in all four years. Group 2 is a set of migrants between April 2000 and April 2001. There is a decrease in the TFR from 2001 – the rate is 2.1 in 2001 before dropping to 1.1 in the other years. For each year, the number of women resident remains roughly the same. Compared with the continuously resident women, these recent migrants have higher fertility in 2001, a period 8-21 months after the date of migration, but lower fertility thereafter. Information about these women's date of migration comes from the census, not from GP registration, and is therefore not associated with reproduction. Therefore the higher fertility in 2001 is probably genuine, and reflects the fact that fertility tends to be high among recent immigrants (Toulemon, 2004). The abrupt decline in fertility in 2002 and onwards is for two reasons: first, the high fertility among these women in 2001 will, for biological reasons, tend to be followed by a period of relatively few births; and, second, it is likely that attrition due to out-migration is more common among recent migrants than among those continuously resident (Group 1). As the LS is poor at capturing emigration in a timely fashion (Hattersley, 1999), the denominators of the fertility rates for years 2002, 2003 and 2004 will include an increasing number of women who have, in fact, departed from England and Wales.

Group 3 are LS members who migrated to England and Wales between 1991 and 2000 and reported at the 2001 census that they were not overseas 12 months before. The group has a high level of fertility compared to the continually resident sample and the recent migrants as of the 2001 census. Among this specific group, the fertility rate is high in 2001 and 2002 before dropping for 2003 and 2004. The difference between this group and Group 2 is that the members of Group 3 have, in 2001, been in England and Wales for some time, so that the rate of attrition between 2001 and 2004 is probably lower. The fact that they have higher fertility than Group 1 is probably associated with their status as recent migrants.

Group 4 has the highest fertility 2001-2004. This group was selected based on registration with a GP in the April 2000-April 2001 period, and reporting at the 2001 census that they were not living overseas 12 months before the 2001 census (April 2000). This indicates that there was a lag between entry to England and Wales (before April 2000) and registration with a GP (between April 2000 and April 2001). It is among these women that the association between registration with a GP and fertility is likely to be strongest, for we know that these women had been resident in England and Wales for some time before entering the LS through registration with a GP. In 2001, their TFR was 4.8. It seems from this that registration takes place in relation to intentions for subsequent fertility. Following the high TFR in 2001, there is a decline to 2.0 in 2002 and then 2.6 and 2.2 in 2003 and 2004 respectively. The timing of registration with a GP was associated with the date of conception of the first child; many of these women will have gone on to have a second child relatively soon after their first, which accounts for the continued high TFR of this group in 2002, 2003 and 2004. There is even some evidence of a two-year spacing between the first and the second child, in that the TFR is higher in 2003 than in either 2002 or 2004.

Table 1. Total fertility rates for 2001-2004 by group at the 2001 census

Sample	2001	2002	2003	2004
Group 1 – continually resident persons, 1991-2001	1.54	1.49	1.56	1.62
Group 2 – LS members at the 2001 census overseas 12 months before	2.12	1.05	1.07	1.07
Group 3 – LS members entering via NHSCR 1991-April 2000, not overseas 12 months before 2001 census	2.20	2.20	1.99	1.89
Group 4 – LS members entering via NHSCR April 2000-April 2001, not overseas 12 months before 2001 census	4.77	2.01	2.63	2.16

Source: Authors analyses based on ONS LS.

5. Conclusions

It is common for life history research to combine data from different sources to produce comprehensive life history data. Linkage of administrative data can often provide high quality, detailed and timely data, but exposure to risk and event likelihood can be related in combined datasets. The ONS LS is one such data set composed from multiple data sources. This research has sought to identify if there is an association between GP registration and subsequent fertility among migrants entering the ONS LS.

The three research questions covered the ability of migrants to be captured in NHS systems and enter the ONS LS, the biases in measuring duration between the entry of female migrants to the data set and a subsequent birth, and the fertility of recent migrants in the period after the 2001 census. It was shown that, during the five years before the 2001 census, the ONS LS generally collected more female migrants through registration with a GP than at the census. However, this does not necessarily mean that the female migrants who enter the LS through NHSCR registration enter at the same time as their migration event, and the analysis strongly suggested an association between the date of registration and the initiation of reproduction. Indeed, it is clear that becoming pregnant leads to a surge in GP registrations and hence in entries to the LS (shown by the peak in the number of births after 7-9 months). This is confirmed by the third piece of analysis presented in this paper, in which fertility during the period 2001-2004 was estimated for four groups of women classified according to their recorded migration history. The group of women who were known to have registered with a GP shortly before 2001 but had migrated earlier (Group 4) exhibited very high fertility in the calendar year 2001. For these women, it seems clear that their registration with a GP was triggered by their intention to become pregnant shortly, or by the fact that they had become pregnant.

Despite this, our analysis does suggest that the fertility of migrants in the years following their arrival in England and Wales is higher than that of non-migrants. A comparison of non-migrants (Group 1) with migrants (Group 3) suggested that those who migrated between 1991 and 2000 had fertility in the period 2001-2004 up to about 40 per cent higher than women who had been continuously resident in England and Wales between 1991 and 2001. Because Group 3 includes only those migrating between 1991 and April 2000 and registering with a GP, bias arising from the relationship between GP registration and subsequent fertility is excluded from these results. Results for Group 2 are key as this sample has been selected based on the response at the 2001 census indicating that the LS members were overseas 12 months before, and for this group we see an immediately higher TFR for 2001 compared to the subsequent years. Therefore, preliminary evidence suggests a high level of fertility among recent migrants, supporting the findings of Toulemon (2004).

At the 2011 census, for the first time since the 1971 census, a question was asked on the date of migration to the United Kingdom, which should provide a more precise date of migration. To further improve information in the ONS LS on the date of migration, a question asking for the date of migration or first permanent residence at the GP registration stage could be asked and included in the dataset. Such a question would also be beneficial for further linkage of administrative data and improving estimates of international migration to England and Wales.

The findings of this research highlight the potential bias which can be introduced into analyses of the risk of life course events, using data sets assembled from several sources. Combining data sources to produce life history data requires an understanding of the source of each data, and associations between the events in one source and appearance in another source.

Acknowledgements

This research has been facilitated by funding from the Economic and Social Research Council (ESRC) (studentship number ES/G018766/1).

The authors wish to acknowledge user support from Chris Marshall and staff at the Centre for Longitudinal Study Information and User Support (CeLSIUS) and assistance from the Office for National Statistics Longitudinal Study Development Team. The authors would like to thank the two anonymous reviewers for their helpful comments.

The permission of the Office for National Statistics to use the Longitudinal Study is gratefully acknowledged, as is the help provided by staff of the Centre for Longitudinal Study Information & User Support (CeLSIUS). CeLSIUS is supported by the ESRC Census of Population Programme (Award Ref: RES-348-25-0004). The authors alone are responsible for the interpretation of the data. Census output is Crown copyright and is reproduced with the permission of the Controller of Her Majesty's Stationery Office and the Queen's Printer for Scotland. Clearance number LS30106.

References

Adelstein, A. M. (1976) Policies of Office of Population Censuses and Surveys – Philosophy and Constraints, *British Journal of Preventive & Social Medicine*, *30*, 1-10. Retrieved from http://www.jstor.org/stable/25565877

Blackwell. L., Lynch, K., Smith. J. & Goldblatt, P. (2003) *Longitudinal Study 1971-2001: Completeness of Census Linkage, Series LS no. 10.*, Office for National Statistics, London. Retrieved from http://celsius.lshtm.ac.uk/documents/LS10.pdf

Ford, D. V., Jones, K. H., Verplancke, J. P., Lyons, R. A., John, G., Brown, G., Brooks, C. J., Thompson, S., Bodger, O., Couch, T., & Leake, K. (2009) The SAIL Databank: building a national architecture for e-health research and evaluation, *BMC Health Services Research, 9*: 157. doi: 10.1186/1472-6963-9-157

Hattersley, L. (1999) *LS User Guide 18 – International migration data in the Longitudinal Study*, Office for National Statistics, LS Unit, London. Retrieved from http://www.celsius.lshtm.ac.uk/documents/userguide18.pdf

Hattersley, L. & Creeser, R. (1995) *Longitudinal Study 1971-1991 - History, organisation and quality of data*, HMSO, London. Retrieved from http://celsius.lshtm.ac.uk/documents/LS%20No.7%20Hattersley%20&%20Creeser%201995.pdf

Lyons, R. A., Jones, K. H., John, G., Brooks, C. J., Verplancke, J. P., Ford, D. V., Brown, G., & Leake, K. (2009) The SAIL databank: linking multiple health and social care datasets, *BMC Health Services Research, 9*: 3. doi: 10.1186/1472-6947-9-3.

Murphy, M. (2009) Where have all the children gone? Women's reports of more childlessness at older ages than when they were younger in a large-scale continuous household survey in Britain, *Population Studies*, *63*, 115–133. doi: 10.1080/00324720902917238

Ní Bhrolcháin, M., Beaujouan, E. & Murphy, M. (2011) Sources of error in reported childlessness in a continuous British household survey. *Population Studies*, 65, 305-318. doi: 10.1080/00324728.2011.607901

Rahman, N. & Goldring, S. (2006) Factors Associated with household non-response in the Census 2001, *Survey Methodology Bulletin*, *59*, 11-24. Retrieved from http://www.s3ri.soton.ac.uk/isi2007/papers/Paper13.pdf

Ralphs, M. & Staples, V. (2012) *Exploring the Challenges of Using Administrative Data*. Office for National Statistics. Retrieved from http://www.ons.gov.uk/ons/about-ons/what-we-do/programmes---projects/beyond-2011/news/reports-and-publications/index.html

Toulemon, L. (2004) Fertility among immigrant women: new data, a new approach, *Population & Societies 400, 1-4.* Retrieved from http://www.ined.fr/en/resources_documentation/publications/pop_soc/bdd/publication/540/

Tromans, N., Natamba, E., & Jefferies, J. (2009) Have women born outside the UK driven the rise in UK births since 2001? *Population Trends, 136*, 28-42. Retrieved from http://www.ons.gov.uk/ons/rel/population-trends-rd/population-trends/no--136--summer-2009/have-women-born-outside-the-uk-driven-the-rise-in-uk-births-since-2001-.pdf

Waller, L., Berrington, A., & Raymer, J. (2012) *Understanding recent migrant fertility in the United Kingdom*, Centre for Population Change Working Paper Number 27. Retrieved from http://cpc.geodata.soton.ac.uk/publications/2012_Understanding_recent_migrant_fertility_WP27_Waller_et_al.pdf

Zumpe, J., Dormon, O., & Jefferies, J. (2012) *Childbearing among UK born and non-UK born women living in the UK*, 25 October 2012. Office for National Statistics. Retrieved from http://www.ons.gov.uk/ons/dcp171766_283876.pdf

Endnotes

[1] Change in address 12 months before the census (internal and international migration) has been identified as being related to higher census non-response (Rahman and Goldring, 2006).

[2] A small number of individuals gave LS dates of birth at the 2001 census and were subsequently found to have been registered before the census date on the NHSCR but with a different date of birth. These people have been excluded from analyses reported for Questions 1 and 3 as, because of their 'date of birth discrepancy', they could not have entered the LS at a date before the census.

KEYNOTE LECTURE

Socio-economic inequality in childhood and beyond: an overview of challenges and findings from comparative analyses of cohort studies*

Jane Waldfogel
Columbia University, USA; London School of Economics, UK
jw205@columbia.edu

(Received February 2013 Revised August 2013) doi:10.14301/llcs.v4i3.263

Abstract

The growing number of countries with large child cohort studies offers an unprecedented opportunity for comparative research. A topic of central interest in my research is, to what extent the sizable gaps in development that exist between children from different socio-economic status (SES) groups in the US are also present in other countries, and to what extent the mechanisms explaining these gaps are similar or different across countries. This overview draws on the results of comparative analyses of birth cohort study data collected in Australia, Canada, the United Kingdom, and the United States, to illustrate the challenges that arise in carrying out this kind of research and the way these challenges were met – in particular those having to do with data access and comparability, and those having to do with causal inference. I conclude that this type of research also offers great promise as shown by findings on SES gaps in child development in the four countries.

*This paper is the edited text of a keynote lecture given at the third annual conference of the Society for Longitudinal and Life Course Studies, Paris, 2012

Introduction

The growing number of countries with large child cohort studies now offers an unprecedented opportunity for comparative research. There are, of course, numerous challenges in carrying out this kind of work, in particular, challenges related to data access and comparability, and challenges related to causal inference. But this type of research also offers great promise. In this talk I will illustrate these points by drawing on comparative analyses of inequality in child development from a programme of research, using birth cohort data from Australia, Canada, UK, and the US, that I have carried out in collaboration with Bruce Bradbury, Miles Corak, and Elizabeth Washbrook. I conclude with a summary of key findings to date and next steps for further research.

The cohort data

Most of the work reported here, relies primarily on four main cohort datasets. (Other datasets used in the analyses are introduced briefly in subsequent sections). The sources are: Australia, *LSAC* (Longitudinal Study of Australian Children - Birth Cohort), representing children born in 2003-2004 (N=5,107); Canada, *NLSCY* (National Longitudinal Study of Canadian Youth) using a cohort of children born in 2000-2002 to match the other datasets (N=8,522); the UK, *MCS* (Millennium Cohort Study)

(N=19,517), for children born 2000-2002 (N=19,517); and the US, *ECLS-B* (Early Childhood Longitudinal Study-Birth Cohort), representing children born in 2001 (N=10,700; all sample sizes for ECLS-B are rounded to the nearest 50, in accordance with the requirements for the restricted data).

Challenges

Data access

The first challenge that arises in this kind of research is data access. Several of the cohort studies are restricted. For example, the Canadian NLSCY may be used only at approved research data centers in Canada, and the US ECLS-B may be used only by license holders in US.

These kinds of restrictions make comparative work more difficult. Analyzing another country's data is not always permitted, and datasets cannot be pooled. So, working in international teams becomes essential. International teams – with researchers from each country -- increase costs and coordination problems, but also confer benefits in terms of ensuring country expertise. This is especially useful in studies such as ours in which the country contexts, and in particular the policy contexts, are relevant.

Data comparability

A second challenge is that data must be "harmonized." It is rare that datasets from different countries use the same measures, so similar measures must be constructed to ensure that the constructs we are examining are equivalent. This challenge is a major one and, as Kohn (1987) pointed out in his seminal Presidential address to the American Sociological Association, poses major costs in terms of time, money, and effort.

Particularly important for this programme of work, measuring socio-economic status (SES) requires constructing similar measures of income or education, and measuring child development requires identifying similar measures of reading, math, behavior, or health. In doing so, there is a tension between using the best measures and the most comparable ones. So it is important to be clear in this kind of work about how data are harmonized and also how sensitive the results are to those decisions. We have also found it helpful to address this tension by presenting main results drawing on the most comparable measures, but then augmenting those with supplemental estimates using the additional measures available in specific countries.

Causal inference

A further challenge in this type of analysis is the problem of causal inference. With just a single cohort from each country, it is difficult to estimate causal effects of policies, since usually all cohort members are exposed to the same policy regime. Repeated cohorts (if available) can be used to estimate effects of *changes* in policies. In addition, geographic or other forms of variation *within* a country can be used (Washbrook, Ruhm, Waldfogel & Han, 2011; Esping-Andersen et al., 2012). But even with single cohorts, comparative analyses can still be informative. At a minimum, they can point to the existence of different patterns or associations, even when explanations for those differences cannot be firmly established.

Socio-economic inequality in childhood

Recent decades have seen an increase in income inequality across most OECD countries (OECD, 2011). This is of particular concern because of evidence that countries with high levels of inequality also tend to have low levels of inter-generational mobility. This relationship is clearest with respect to the US where high levels of income inequality are associated with low levels of mobility between generations (Corak, 2013; Björklund & Jäntti, 2008).

Why is there this link between inequality and immobility and at what stages of life does it develop? Increasingly, attention is turning to the early years – which are seen by many psychologists and economists as particularly important for the development of social and economic capabilities (e.g. Heckman & Lochner, 2000). Moreover, strong social inequalities in learning skills at school entry are well documented in the US (Burkam & Lee, 2002; Magnuson & Waldfogel, 2005, Magnuson & Duncan, 2006, Murnane et al, 2006). Is this pattern repeated in other countries, and which features of national environments are associated with these indications of early socio-economic immobility?

The research agenda of our four country study therefore focused on using cohort data from across similar countries to address two key questions:

1) Is inequality in school readiness, and the factors that explain it, different across countries?
2) Does inequality widen or narrow after school entry, and does this vary across countries?

Using cohort data to examine socio-economic status gaps in school readiness in US and UK

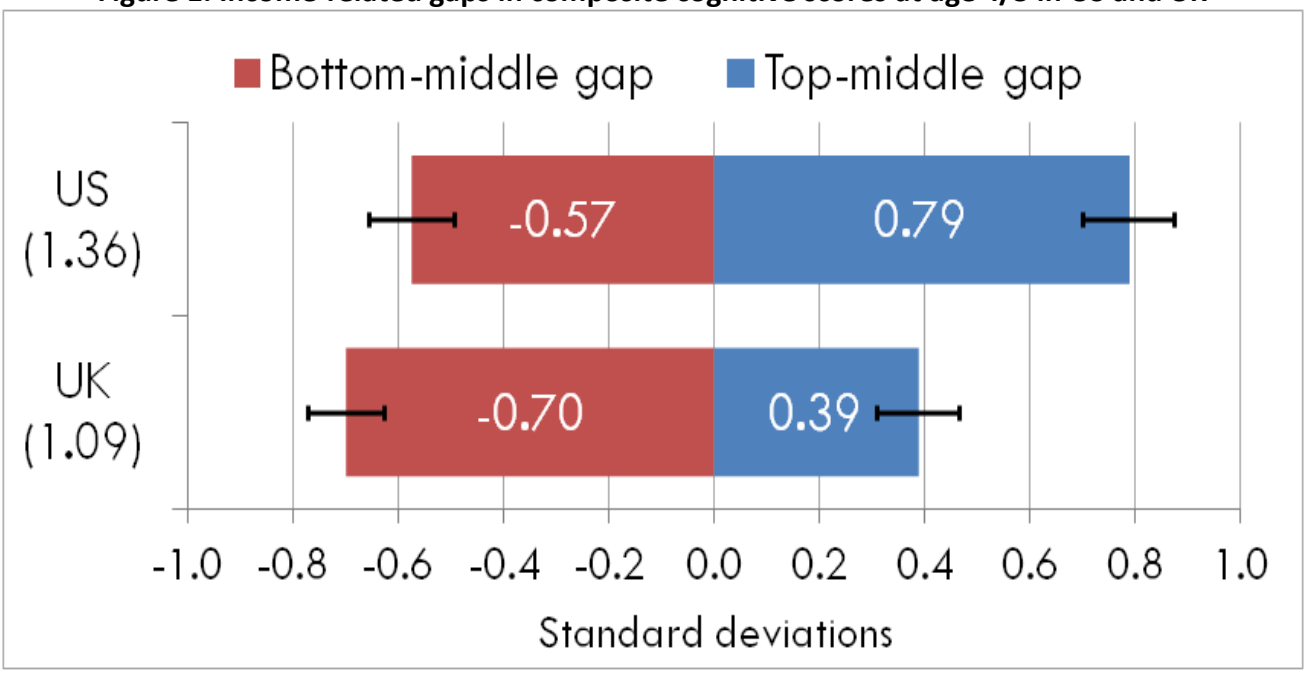
Using ECLS-B (US) and MCS (UK), Waldfogel and Washbrook (2011a, 2011b) examined SES gaps in school readiness across the two countries. To create comparable SES groups, we divided families in each country into income quintiles (with Q1 representing the lowest income quintile, and Q5 the highest).[i]

To construct comparable cognitive outcome measures, given that the underlying cognitive assessments differed across countries, we created cognitive composite scores by taking the first principal component from a battery of test scores.[ii]

Our analysis also explored differences in behavioral development, drawing on 25 items from the parent-reported Strength and Difficulties Questionnaire (SDQ) in the MCS, and 21 highly similar items (collectively drawn from a number of survey instruments) in the ECLS-B. Scores were adjusted for gender and a cubic trend in age in months, and principal components analysis (PCA) was then conducted separately to extract the cognitive and behavioral indices.[iii]

Our analysis focused on bottom-middle (Q1-Q3) and top-middle (Q5-Q3) gaps in the cognitive and behavioral outcomes. We found sizable SES gaps in school readiness in both countries, but disparities were proportionately larger in the cognitive than the behavioral domain in both countries. Figure 1 shows the key patterns for the cognitive outcome, with clear differences across countries not just in overall magnitude of the gap but in the parts of the income distribution in which it is concentrated. For example, the difference in ability between middle- and high-income children is larger in the US than the UK, whereas the reverse is true for the disparity between the middle and the poorest.

Figure 1. Income-related gaps in composite cognitive scores at age 4/5 in US and UK



Source: Waldfogel & Washbrook (2011b)

The next step was to use the rich data from the US and UK cohort studies to identify factors that might explain these gaps. To play a role in explaining gaps, a factor must:

- differ between income groups
- have an effect on the outcome

We considered a wide range of factors that met these criteria, and found in both countries that the single most important factor explaining the cognitive gaps was parenting. In this context, *parenting* is a construct that includes the sensitivity and responsiveness of parent-child interactions, as well as the quality of the home learning environment that parents provide. The importance of parenting was paralleled in the results for behavior gaps, although the analysis revealed some differences in the relative importance of different factors for the two dimensions of development.

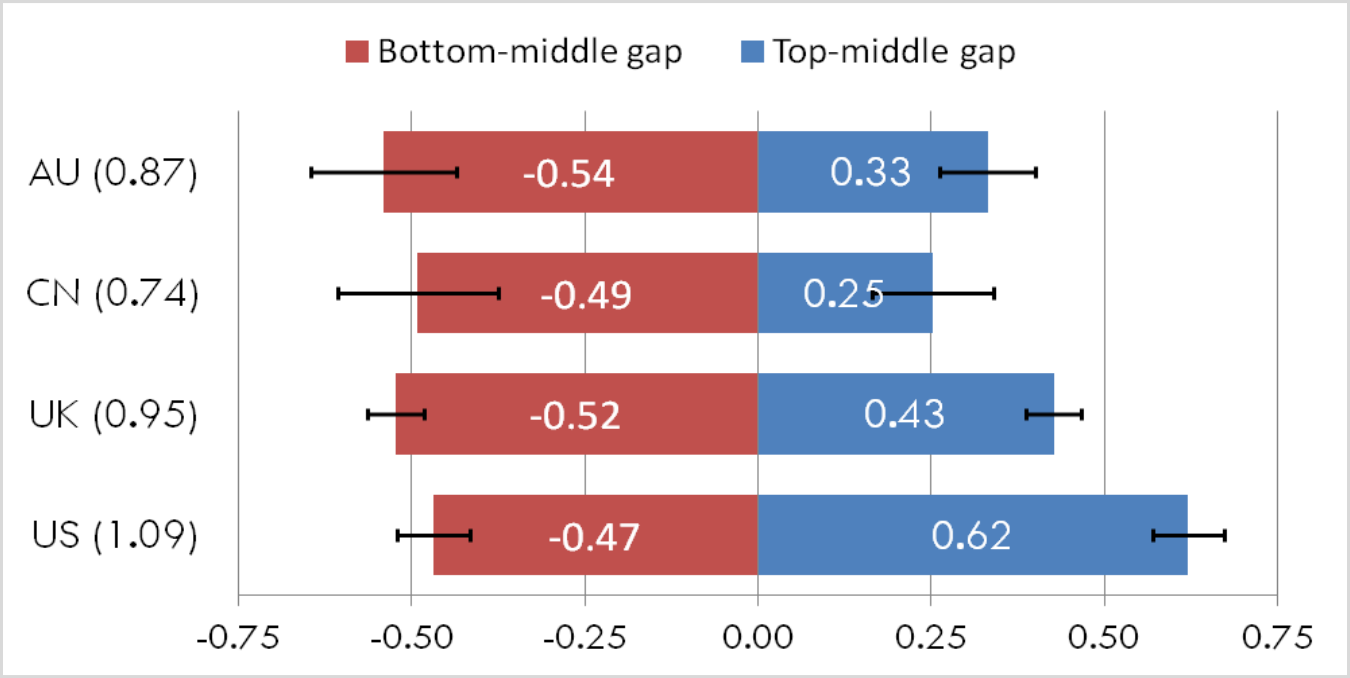
Are similar gaps in school readiness seen in other countries?

Bradbury, Corak, Waldfogel, and Washbrook (2012) extended this analysis to four countries, using cohort data from LSAC (AU), NLSCY (CA), MCS (UK), and ECLS-B US). To create comparable SES groups, we again divided families into income quintiles in the manner described above, but also extended the analyses to include comparisons by parental education level. While income quintiles are a relative measure (one that may result in families with very different average levels of resources being in the bottom group in different countries), parental education can arguably be seen as an absolute measure that serves to compare "like with like". Parental qualifications in all countries were coded using the International Standard Classification of Education (ISCED) measure, to ensure this was the case as far as possible.

Ensuring comparability of outcome measures becomes more difficult as the number of countries increases. An insufficient number of similar cognitive test scores were available in the NLSCY and LSAC, to enable the extraction of a principal component that measured the same latent construct in all four countries. Only two aspects of development were assessed directly and in a comparable way in all four countries. Vocabulary was measured using PPVT scales in three countries, and the BAS Naming Vocabulary scale in the MCS tapped a very similar ability. Externalizing behavior (a composite of hyperactivity/inattention and conduct problems) was assessed by two SDQ sub-scales in both the MCS and LSAC, and similar items could be selected from the behavior questions in the other two surveys. A number of other abilities were assessed comparably in at least two countries with, for example, a copying test of fine motor ability conducted in all surveys except the MCS. Our strategy was to focus our analysis on the vocabulary and externalizing behavior scores, since those measures existed for all four countries, but to calculate gaps for all available outcomes to see if the patterns would be replicated, if only partially, with other measures.

Focusing again on bottom-middle (Q1-Q3) and top-middle (Q5-Q3) income-related gaps in cognitive development, we found that SES gaps in school readiness exist in all four countries, but the size of the gaps differs, with generally smaller gaps at the top in AU and CN than in the UK and US (see Figure 2). The finding of smaller gaps in the Australian and Canadian cohorts was consistent for all the measures we were able to analyze, and differed little, whether income or education was used to capture SES. With regard to the measure of behavior problems, the results for all four countries echoed our earlier finding that socio-emotional disparities are consistently smaller than cognitive ones at the age of school entry.

Figure 2. Income gaps in vocabulary at age 4/5: four countries



Source: Bradbury, Corak, Waldfogel & Washbrook (2012)

Are similar gaps seen for children of immigrants?

In many countries, children of immigrants make up a growing share of the child population, and inequality between children of immigrants and children of native-born parents is a growing concern. Thus, although our main focus to date has been on inequality associated with socio-economic status, we have also used the same cohort datasets to examine gaps in school readiness between children of immigrants and children of native-born parents (Washbrook, Waldfogel, Bradbury, Corak, & Ghangro, 2012).

To do so, we divided children in each country into four groups, based on parents' immigration status and language: native-born parents, with the official language of the country spoken at home; at least one foreign-born parent, with the official language spoken at home; at least one foreign-born parent, with a foreign language spoken at home; and all others (this latter category includes families whose immigrant or language status could not be determined). The 'official language' was either English or French in Canada, and English in the other countries.

For the reasons discussed above, our plan was to focus our analyses on the two outcomes that are comparable across all four countries: the vocabulary and externalizing behavior scores. However, initial within-country analyses showed that, unlike SES disparities, gaps related to parental immigration looked very different between the verbal and non-verbal cognitive measures. It became clear that in this analysis, vocabulary tests administered in the country language could not be considered a single reliable indicator of overall cognitive ability and would provide only a partial picture. Specifically, large disparities between certain groups in verbal ability co-existed with markedly smaller differences in non-verbal skills that were sometimes even of the opposite sign. This was particularly marked in the case of copying skills (available in three countries) but also in the tests of non-verbal reasoning ability conducted in the MCS (the BAS Pattern Construction and Picture Similarities scales) and in several other tests of numerical and writing abilities.

Overall, we found that generally, children of immigrants enter school as, or more, ready to learn, as the children of native-born parents. The large

vocabulary deficits of the children of immigrants who speak a foreign language in the home, are the exception rather than the rule. These findings suggest that studies which rely upon measures of vocabulary may present a distorted picture of the capabilities of children from non-official language backgrounds. Indicators that are not reliant upon language will generally show children of immigrants to be much more capable. This suggests a need for future data collections to place more emphasis on assessing children in their home language (even for children who know enough of the official language to communicate with the interviewer).

Using cohort data to examine the evolution of inequality after school entry

We would also like to know what happens to inequality once children are in school. Do gaps widen or narrow?

Magnuson, Waldfogel, and Washbrook (2012) examined this for the US and UK, using data from ECLS-K (Early Childhood Longitudinal Study-Kindergarten Cohort, N=21,409, born 1993) and ALSPAC (Avon Longitudinal Study of Parents and Children, N=13,988, born 1991-1992). These cohorts are older than the ones discussed so far, enabling us to follow children from school entry up to the age of 14, in a way not possible for those born more recently. The ECLS-K is a nationally representative sample of children, clustered in schools and recruited in the fall of Kindergarten around the age of 5. Parental and teacher interviews have so far been conducted at six waves, the most recent in 2007 when the children were in 8th grade. In contrast to the other cohorts discussed here, ALSPAC is not a nationally representative sample, but an entire birth cohort from one region of England. Mothers were recruited in pregnancy, and they and their children have been surveyed with high frequency ever since. While ALSPAC lacks the advantage of national representativeness, its data is by far the richest of all the cohorts we use, and includes items not available elsewhere such as biological markers and a battery of psychological assessments. That said, validation studies have shown that the demographic make-up of the Avon region is very similar to that of England as a whole.

The outcome measures examined in our study, are assessments of reading and math ability at several points between school entry and the age of 14. The measures are unusual in that they are measured on the same scale at all ages, allowing exploration of the evolution of absolute as well as relative disparities in achievement (i.e. with and without standardizing outcome variance at different ages). The ECLS-K measures are direct assessments of ability administered by the survey; the ALSPAC measures are National Curriculum levels from the Key Stage assessments administered in all English state schools at age 7, 11 and 14.[iv]

Again, disparities in achievement according to two measures of SES were examined, contrasting income quintiles and ISCED levels of parental education. A full description of the results is beyond the scope of this lecture but one key finding was that the characterization of the evolution of disparities was different depending on whether raw or standardized scores were used, although it was relatively insensitive to the choice of measure of outcome (reading or math) or of SES (income or education). Another key finding was that the pattern of the gaps differed across the two countries, with more marked narrowing of gaps in the first few years of primary school in the UK, but then possibly more widening of the gaps thereafter, as compared to the patterns for the US.

Evolution of inequality after school entry in AU, CA, UK, and US

In a new project funded by Russell Sage Foundation (RSF) and the Australian Research Council (ARC), Bradbury, Corak, Washbrook, and I are extending the analysis of inequality after school entry to the four countries, using data from LSAC (AU), NLSCY (CA), MCS (UK), and ECLS-K (US). Specifically, we are following children from school entry to age 11-13, analyzing whether SES gaps widen, hold constant, or narrow, and relating differences in inequality across countries to differences in family resources, labor markets, and policy contexts.

Conclusions and next steps

Our programme of work to date has demonstrated the rewards of using cohort data for comparative analyses. Although arranging data access and ensuring comparability are challenging and have been costly, they have also yielded considerable insights that would not have been possible in single-country analyses.

Results to date indicate that achievement gaps at school entry differ across countries, and interestingly, these differences mirror differences in adult social mobility, with larger gaps in the US and UK than in AU and CN. These results suggest that future research probing more deeply into the role of country contexts, and in particular policy contexts, would be fruitful.

We have also found results that are robust across countries. Our US-UK comparative analyses show that parenting is a key factor explaining SES gaps in child outcomes in both countries. This is an important finding, and one that has implications for policy in both countries. Moreover, our four-country analyses provide robust evidence that language is a key factor in explaining gaps for children of immigrants. In particular, on language-free tasks, children of immigrants do as well/better than children of native-born parents. This too is a policy-relevant finding, as countries seek to understand the implications of the influx of children of immigrants into their school systems.

Finally, the evolution of gaps after school entry also appears to differ across countries. However, this conclusion is, thus far, based only on data from the US and UK, and it will be important to extend it to other countries. Accordingly, the next step in this programme of research is extending the study of gaps after school entry to four countries -- AU, CN, UK, US (the subject of our current RSF and ARC studies).

Looking further ahead, it would also be useful to examine gaps at school entry, and the evolution of gaps after school entry, in additional countries, to take advantage of additional variation in country contexts and, in particular, policy contexts. It would also be informative to examine changes in gaps for succeeding cohorts, both within and across countries, for those countries (such as the UK) that have repeated cohorts. Both such extensions would strengthen our ability to draw conclusions about the sources of and potential remedies for the types of inequality that the rich, cross-national cohort data have so amply illustrated.

More generally, our work with the cohort data illustrates the potential of comparative cohort analyses to shed light on the processes underlying child and youth health and development, and the potential role of policies both in improving average outcomes and in reducing disparities in outcomes. As more cohort datasets become available, with rich data on health and developmental outcomes, we hope to see more analyses of this kind. My experience working on this programme of research suggests that although this kind of comparative research is costly, it is also immensely rewarding. In my view, the time and effort involved in addressing the challenges of accessing data and ensuring comparability, while substantial, are well worth it.

Acknowledgements

The work reported here was conducted as part of an ongoing research programme in collaboration with Bruce Bradbury, Miles Corak, and Elizabeth Washbrook and with funding support from Sutton Trust, Russell Sage Foundation, Carnegie, NICHD and Australian Research Council.

References

Björklund, A., & Jäntti, M. (2008). Intergenerational Income Mobility and the Role of Family Background. In W. Salverda, B. Nolan, & T. Smeeding (Eds), *The Handbook of Economic Inequality*. Oxford: Oxford University Press.

Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2012). Inequality during the Early Years: Child Outcomes and Readiness to Learn in Australia, Canada, United Kingdom, and United States. In J. Ermisch, M. Jäntti, & T. Smeeding (Eds), *From Parents to Children: The Intergenerational Transmission of Advantage*. New York: Russell Sage Foundation.

Burkam, D. & Lee, V. (2002). *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School*. Washington, DC: Economic Policy Institute.
Corak, M. (2013). Income Inequality, Equality of Opportunity, and Intergenerational Mobility. *Journal of Economic Perspectives* (forthcoming).
Esping-Andersen, G., Garfinkel, I., Han, W.-J., Magnuson, K., Wagner, S., & Waldfogel, J. (2012). Child Care and School Performance in Denmark and the United States. *Children and Youth Services Review 34,* 576-589.
Heckman, J.,& Lochner, L. (2000). Rethinking Education and Training Policy: Understanding the Sources of Skill Formation in a Modern Economy. In S. Danziger and J. Waldfogel (Eds.), *Securing the Future: Investing in Children from Birth to College*. New York: Russell Sage Foundation.
Kohn, M. (1987). Cross-National Research as an Analytic Strategy. *American Sociological Review* 52, 713-731.
Magnuson, K., & Duncan, G. (2006). The Role of Family Socioeconomic Resources in the Black-White Test Score Gap among Young Children. *Developmental Review* 26, 365-399.
Magnuson, K., & Waldfogel, J. (2005). Child Care, Early Education, and Racial/Ethnic Test Score Gaps at the Beginning of School. *The Future of Children 15*, 169-196.
Magnuson, K., Waldfogel, J., & Washbrook, E. (2012). The Development of SES Gradients in Skills during the School Years: Evidence from the United States and England. In J. Ermisch, M. Jäntti, & T. Smeeding (Eds), *From Parents to Children: The Intergenerational Transmission of Advantage.* New York: Russell Sage Foundation.
Murnane, R., Willett, J., Bub, K., & McCartney, K. (2006). Understanding Trends in the Black-White Achievement Gaps during the First Years of School. *Brookings-Wharton Papers on Urban Affairs* (pp.97-135). Washington DC: Brookings Institution Press.
Organization for Economic Cooperation and Development (OECD) (2011). Divided We Stand: *Why Inequality Keeps Rising.* Paris: OECD Publishing.
Waldfogel, J., & Washbrook, E. (2011a). Early Years Policy. *Child Development Research* 2011, 1-12.
Waldfogel, J., & Washbrook, E. (2011b). Income-Related Gaps in School Readiness in the United States and United Kingdom. In T. Smeeding, R. Erikson, & M. Jäntti (Eds), *Persistence, Privilege, and Parenting: The Comparative Study of Intergenerational Mobility.* New York: Russell Sage Foundation.
Washbrook, E., Ruhm, C., Waldfogel, J., & Han, W.-J. (2011). Public Policies, Women's Employment After Childbirth, and Child Well-Being. *B.E. Journal of Economic Analysis and Policy* 11, Article 43. Available at: http://www.bepress.com/bejeap/vol11/iss1/art43.
Washbrook, E., Waldfogel, J., Bradbury, B., Corak, M., & Ghanghro, A.A. (2012). The Development of Young Children of Immigrants in Australia, Canada, the United Kingdom and the United States. *Child Development* 83, 1591-1607.

Endnotes

[i] Incomes were equivalized for household size, expressed in constant prices and averaged over three waves of data to reduce measurement error. (We thus did not examine income dynamics – for example the influence of changes in income or income at different stages of the child's life). Weights provided with the surveys were used when calculating the quintiles so that, as far as possible, the groups were representative of the underlying national populations.

[ii] To construct our composite, we used all 5 test scores available in the MCS - Bracken School Readiness (age 3); BAS Naming Vocabulary (age 3, 5), BAS Picture Similarities and Pattern Construction (age 5) – and all 6 available in the ECLS-B - literacy, math, receptive vocabulary, expressive language, color knowledge and fine motor copying ability (age 4).

[iii] PCA has the advantage that it extracts the maximum possible "signal" from a number of potentially noisy measures and the first component can be used as a single focal index in complex analyses between groups and across countries. When conducting international comparisons, however, careful inspection of the results of the PCA is necessary to ensure that the loadings and other diagnostic statistics are similar across countries.

[iv] An Entry Assessment score measured at age 4 can also be used for the ALSPAC cohort in the standardized analyses only.

Longitudinal and Life Course Studies 2013 Volume 4 issue 3 Pp 276-287 *ISSN 1757-9597*

RESEARCH NOTE

Siblings and child development

Elise de La Rochebrochard, Ined, Paris, France
roche@ined.fr
Heather Joshi, Institute of Education, London, United Kingdom

(Received March 2013 Revised August 2013) doi:10.14301/llcs.v4i3.248

Abstract

Having many siblings, or none, may impair, or improve, a child's development compared to being part of a two-child family. Any effect may vary for different aspects of development. This note describes, cross-sectionally, the observed association between child development at ages 3 to 7 years and the number of co-resident siblings, at three sweeps of the UK Millennium Cohort Study. Indicators of cognitive development (verbal and non-verbal), are taken from surveys at ages 3, 5 and 7 years. Behavioural problems are reported at the three surveys on the Strengths and Difficulties Questionnaire. We analyze its five sub-scales separately, and also the Total Difficulties score. For each of 26 outcomes, we estimate the risk of falling into the most problematic 10% of the child population, depending on the number of siblings at each survey, controlling for the child's gender and the level of the mother's education. In this descriptive exercise, maternal education stands in for a host of possible social covariates; and allows for the least educated mothers having larger families. Children with ≥2 siblings generally showed increased odds of adverse outcomes, especially in cognition at age 3. This is in line with the hypothesis of resource dilution, but only-children tend, for some outcomes, to score worse than those in two-child families. The odds ratios (ORs) for boys were, except one, unfavourable, of similar magnitude to estimates for larger families. The disadvantages associated with low maternal education were generally greater. These explorations lay the ground for longitudinal modelling of causal pathways.

Keywords: child development; school readiness; Strengths and Difficulties Questionnaire; siblings; resource dilution; Millennium Cohort Study

Abbreviations: BAS, British Ability Scales; MCS, Millennium Cohort Study; OR, Odds Ratio; PiM, The Progress in Mathematics; SDQ, Strengths and Difficulties Questionnaire; UK, United Kingdom.

Introduction

Quantitative studies investigating influences on child development often control for the number of siblings, but seldom focus attention on it. There are reasons to expect both positive and negative effects of the number of brothers and sisters.

Children may benefit directly from the learning, company and affection that siblings can provide each other. Siblings may be social, emotional, language and cognitive assets from a child's point of view (Azmitia & Hesser, 1993; Bowes, Maughan, Caspi, Moffitt & Arseneault, 2010; Brody, 2004; Gass, Jenkins & Dunn, 2007). Research on kindergarten children showed that children having at least one sibling display higher social skills (Downey & Condron, 2004). Children with no siblings may have low behavioural and social skills (self-centred, less cooperative), perhaps because doting parents bring them up to be "little emperors" as suggested in the case of China's One-Child policy (Cameron, Erkal, Gangadharan & Meng, 2013). Younger siblings could benefit from their older siblings as better agents of cognitive development than immediate age-peers (Azmitia & Hesser, 1993). In return, older siblings could benefit from these teaching interactions, which are considered as an important promoter of verbal intellectual development for the "instructor" child (Smith, 1993).

On the other hand, siblings could be a liability. First, siblings can be bullies, harming a child's emotional and behavioural development (Wolke & Skew, 2012a). The risk of being bullied at home is reported to rise with the number of siblings, from 46% with 1 sibling, 53% with 2 siblings, to 58% with 3 or more siblings (Wolke & Skew, 2012b). Secondly, siblings could create competition for parental attention, and a dilution of the resources that may be devoted to any one child. Downey posited that parents have finite levels of time, energy, and money, and that these resources are diluted among children as sibship size increases (Downey, 1995; Downey, 2001). This pattern is observed in research based on the Avon Longitudinal Study of Parents and Children (ALSPAC), finding family size to be the strongest influence on parental investment per child, even in wealthy or well educated families (Lawson & Mace, 2009). If parental investment is important for the acquisition of social or human capital, this could be key to a negative link between larger family size and educational achievement (Becker & Tomes, 1976; Blake, 1981; Coleman, 1988). Not only could siblings "dilute" parental resources, they could also "dilute" the intellectual home milieu, according to the "confluence theory" (Steelman, Powell, Werum & Carter, 2002; Zajonc & Markus, 1975). On this hypothesis, the average intellectual level of the home decreases with the arrival of a new sibling.

While the direction and the potential causal impact of siblings on the development of a given child remains ambiguous (Guo & VanWey, 1999), the number of children in a family may reflect other circumstances which also impinge on child development. This descriptive note takes only preliminary steps towards developing research on the impact of siblings on each other. It presents some basic correlations between the number of siblings and child development in the UK at ages 3, 5 and 7. It paves the way for the elaboration suggested in our conclusions, and offers guidance on how information on siblings might handled in research focussed on other aspects of child development.

Data

We used the UK Millennium Cohort Study (MCS), a representative, but disproportionately stratified, and clustered, sample of around nineteen thousand families having a child born in 2000-2001. Data were collected at 9 months and then at 3, 5 and 7 years through home visits. From age 3, the cognitive assessments of the cohort child were administered by trained interviewers and behavioural problems were parent-reported. Detailed information on the survey is available online (www.cls.ioe.ac.uk/mcs).

Number of siblings

Siblings were defined as co-resident children, whether they were fully related or not to the cohort child. Full siblings, half, step, adopted and foster siblings were all counted. In cases of multiple births in the cohort, one child was randomly defined as the index child for this analysis and his/her twin or triplets were counted among siblings. The number of siblings was recorded four times: at 9 months and then at 3, 5 and 7 years. To allow for possible non-linearity in the relationship between number of siblings and the cohort child's development, the variable "number of siblings" was coded in

categories: 0 sibling (the cohort child is an only-child), 1 sibling (the cohort child is in a two-child family), and in our main analysis 2 or more siblings (i.e. a 'large' family of a least three children). In supplementary work, we also distinguished 3- and 4+ child families among the latter. As two children is the norm for British families (Office for National Statistics, 2012), having one sibling was set as our reference category.

Measures of child development

We investigated the cohort child's development based on scores for cognitive and behaviour development at ages 3, 5 and 7 years, adjusted, in the case of the cognitive scores, for age at assessment. We focus exclusively on "low scores" defined as the bottom 10% of each observed distribution. We abandoned our initial intention to look at "high scores" (top 10%) as, for some of the behavioural sub-scales, the top 10% were not defined (up to 30% of the sample obtained the highest score of the distribution).

Measures of the cohort child's cognitive development

At age 3, cognitive development was measured by the Revised Bracken Basic Concept Scale (Bracken, 1998). The subtests administered were colours, letters, numbers and counting, sizes, comparisons, shapes. Several measures were based on the British Ability Scales (BAS) (Elliott, Smith & McCulloch, 1997). At ages 3 and 5, the BAS "naming vocabulary" instrument measured expressive (but not receptive) language ability. At age 7, BAS "word reading" measured recognition and oral reading of single words (but not reading comprehension). At age 5, BAS "picture similarities" measured non-verbal reasoning ability. At ages 5 and 7, the BAS "pattern construction" scale was used to measure non-verbal ability, especially spatial. Finally, at age 7, the "Progress in Mathematics" (PiM) test from NFER (www.gl-assessment.co.uk/products/progress-maths) was used. This included tasks based on numbers, shape, space, measures and data handling. Thus we had eight cognitive scores in total between 3 and 7 years. These cognitive scores measured different dimensions of cognitive development with a strong verbal component for the Bracken and the BAS Naming Vocabulary.

Measures of the cohort child's behaviour

Behaviour was measured using the Strengths and Difficulties Questionnaire (SDQ) at ages 3, 5 and 7 years (Goodman, 1997). The SDQ (www.sdqinfo.com), is recognised as a psychometrically sound indicator of overall child mental health problems (Goodman, Lamping & Ploubidis, 2010; Stone, Otten, Engels, Vermulst & Janssens, 2010; Vostanis, 2006). It consists of 25 items covering five subscales: emotional problems, peer problems, behavioural problems, hyperactivity and prosocial behaviour. The SDQ questionnaire was completed by the main respondent in the MCS, virtually always the mother (at least 97% of the respondents at these surveys). Usually, the SDQ is analysed in terms of its total difficulties score (which sums the four first sub-scales). However, as the presence of siblings could have different effects on the different SDQ sub-scales, we explored not only the total difficulties score but also the five sub-scales separately. At each of the three ages observed, we thus considered 6 SDQ scores with a total of 18 SDQ scores between 3 and 7 years.

Analyses

Basic analyses were carried out on cross-sections at each age, without attempting to trace the experience of individual children across sweeps. This meant that the fullest cross-sectional sample was considered at each age. For example, if a child participated in the age 3 and 7 data collections but not at age 5, s/he was included in the analyses of outcomes at 3 and 7, even if s/he could not be included in analyses of age 5 outcomes. For each different analysis done, the sample size is indicated in Tables 1 and 2 and varies from 13,338 to 14,924.

Logistic regressions were performed on the risk of having a low score (bottom 10% of the distribution) on the 26 child development outcomes.

The models focussed on the number of siblings at each survey. They controlled for just two other factors known to be related to child development: the gender of the cohort child and the level of the mother's education. The latter, for preliminary purposes, stands in for a host of possible social confounders, from poverty to parenting, including family building intentions (La Rochebrochard (de) & Joshi, 2013). Mother's level of education was divided in three groups. The highest level was 'college educated', i.e. those with degrees, bachelors or higher, or equivalent vocational qual-

ifications (National Vocational Qualification (NVQ) Level 4 or 5). Middle education covers those who did not proceed to college but gained academic qualifications at secondary school, or equivalent vocational qualifications, at NVQ levels 2 and 3. The low education group had minimal (NVQ Level 1, or other overseas) or no qualifications. In this preliminary research, we did not consider the cohort child's birth order, characteristics of the siblings, or other factors about the family. All percentages and odds ratios were weighted to reflect the original sampling probabilities. Variances were estimated with the correction for a finite population. As this is only a preliminary exploration, we treated a 0.05 *p-value* as a test of significance, even where multiple tests were carried out. A more conservative approach would be to apply a Bonferroni correction (Bland & Altman, 1995) which would set the threshold for significance at a *p-value* of 0.002. This stricter criterion is also indicated in tables of results. Data analyses were carried out using STATA/SE 11.1 (Stata Press, College Station, TX, US).

Results

Figure 1. Distribution of the number of siblings when the cohort child is 9 months, 3, 5 and 7 years (UK Millennium Cohort)

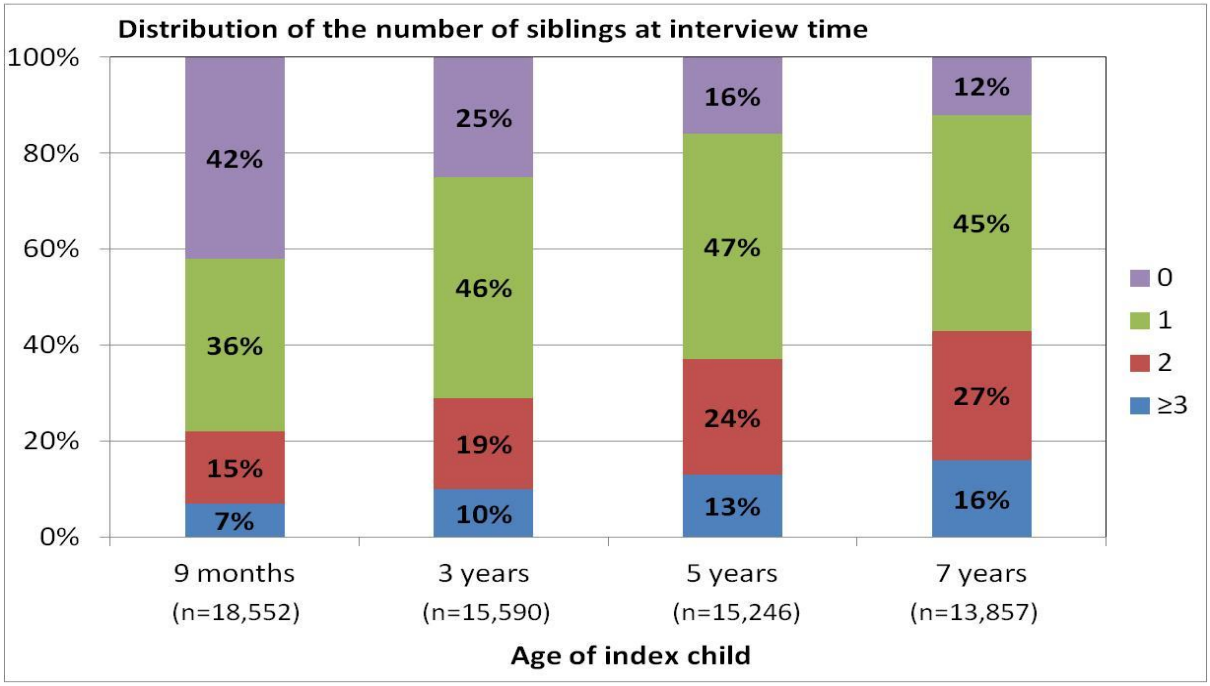


Figure 1 shows that that the proportion of cohort children with at least one sibling rose between surveys from 58% to 88%, reflecting particularly the arrival of new babies during the early years of the cohort child's life. Indeed, at ages 5 and 7, only-children belong to a diminishing minority (16-12%). Meanwhile, the proportion of children with three or more siblings (i.e. in a family of at least four children) rose from 7% to 16%. The two-child family was the most frequent at age 3 and after. The sample size (see figure 1, horizontal axis labels) declined over time as the survey suffered attrition. In order to check that this survey loss was neutral with respect to family size, the distributions of siblings at 9 months were compared across the samples observed at 9 months, 3, 5 and 7 years. The four distributions (not shown) were very similar and do not indicate any attrition bias associated with number of siblings.

Figure 2. Distribution of the number of siblings when the cohort child is 7 years by mother's level of education (UK Millennium Cohort)

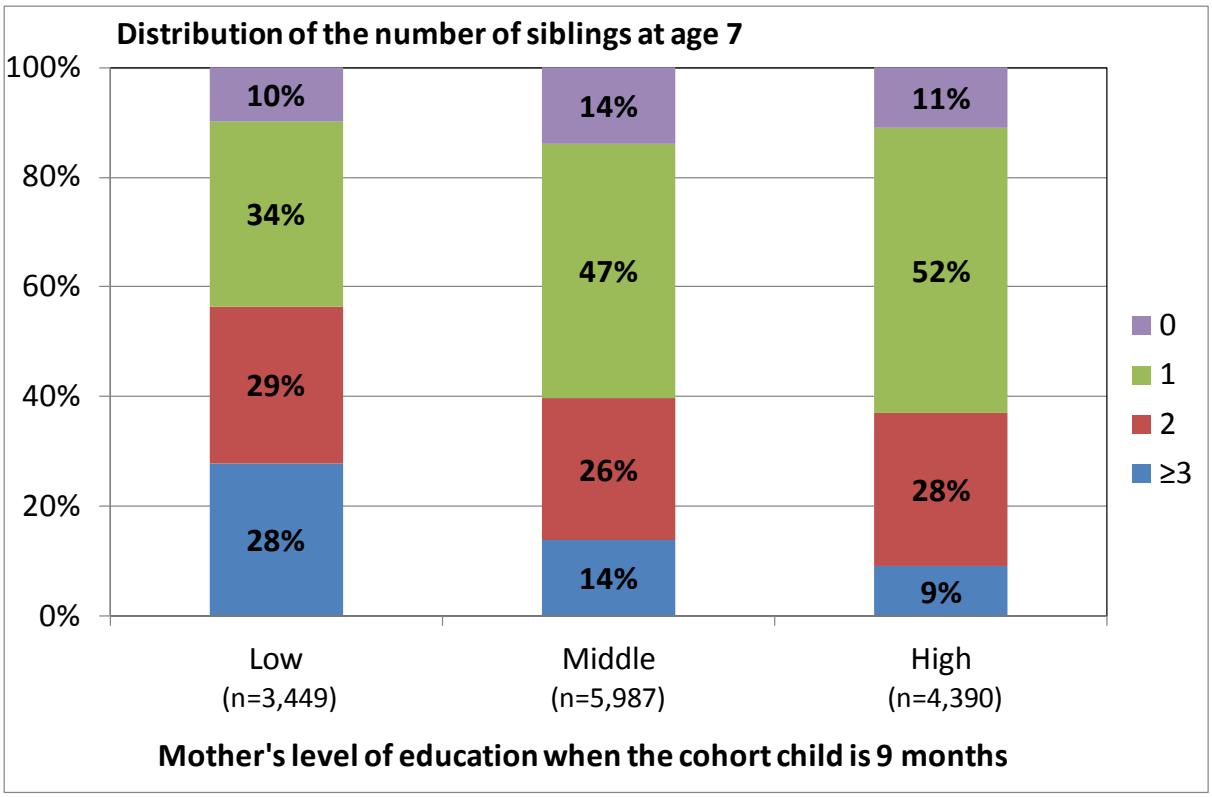


Figure 2 shows the distribution of number of siblings by mother's education when the cohort child was 7. For both high and middle educated mothers, the proportions of only-child families and of large families (at least four children) were low (below 15%). Among mothers with no or minimal qualifications, larger families (four plus) occurred most often (28%) and one child families least often (10%).

Table 1 and Table 2 report the estimates from the 26 logistic multivariate regressions on the risk of having a score in the bottom 10% of the observed distribution, adjusted for number of siblings, gender of child and mother's education. Children with no sibling (ie only-children) are compared to the reference group having 1 sibling in the column headed "only-children". An adjusted odds ratio (OR) over unity indicates that the only-child has a greater risk of being in the bottom 10% of the distribution, or vice-versa if the OR is less than one. A child with no siblings tends to do better than a child with one sibling on the Bracken and Naming Vocabulary at age 3 and 5, but s/he tends to do worse than a child with one sibling on the other cognitive outcomes at 5 and 7 (Picture Similarities, Pattern Construction, Word Reading, Number Skills). On the behaviour outcomes, being an only child tends to be an advantage for conduct problems and prosocial behaviour. However, this advantage tends to decrease and even to disappear for only-children aged 5 and 7 years. Being an only child tends to be a disadvantage for peer problems and hyperactivity which tends to increase for the group of children who remain without any sibling at age 5 and 7.

Table 1. Adjusted odds ratios for being in the bottom 10% of the observed scores distributions for cognitive outcomes by (i) number of siblings, (ii) gender and (iii) mother's education

	Sample size	Number of siblings (reference: 1 sibling)						Gender (reference: girl)			Mother's education (reference: High educated)					
		Only- children			≥2						Middle			Low		
		OR[1]	CI95%	*p*	OR[1]	CI95%	*p*	OR[2]	CI95%	*p*	OR[3]	CI95%	*p*	OR[3]	CI95%	*p*
Bracken age 3	13,820	0.91	0.75-1.10		1.91	1.65-2.22	***	1.81	1.58-2.08	***	2.33	1.87-2.90	***	6.31	5.15-7.73	***
BAS Naming Vocabulary age 3	14,542	0.81	0.67-0.97	‡	1.65	1.44-1.90	***	1.53	1.34-1.75	***	1.85	1.50-2.29	***	6.03	4.84-7.50	***
BAS Naming Vocabulary age 5	14,924	0.93	0.76-1.15		1.57	1.39-1.77	***	1.35	1.18-1.54	***	2.42	2.00-2.93	***	6.78	5.63-8.15	***
BAS Picture Similarity age 5	14,916	1.15	0.96-1.36	*	1.13	0.97-1.32	*	1.42	1.24-1.64	***	1.80	1.54-2.11	***	3.10	2.57-3.75	***
BAS Pattern Construction age 5	14,866	1.24	1.04-1.47	**	1.45	1.27-1.66	***	1.96	1.70-2.26	***	1.95	1.61-2.37	***	3.87	3.17-4.74	***
BAS Pattern Construction age 7	13,492	1.27	1.02-1.58	**	1.36	1.14-1.61	***	1.34	1.16-1.56	***	1.76	1.45-2.13	***	3.66	3.01-4.46	***
BAS Word Reading age 7	13,381	1.09	0.84-1.42		1.56	1.32-1.84	***	2.29	1.98-2.64	***	2.81	2.25-3.52	***	5.67	4.58-7.02	***
Number skills age 7	13,543	1.31	1.03-1.66	**	1.28	1.09-1.52	**	1.27	1.11-1.46	***	2.50	2.05-3.05	***	5.01	4.00-6.27	***

Notes. *Increased risk at p≤0.002* *** *Increased risk at 0.002<p≤0.05* ** *Increased risk at 0.05<p≤0.15* *
Decreased risk at p≤0.002 ‡‡ *Decreased risk at 0.002<p≤0.05* ‡ *Decreased risk at 0.05<p≤0.15* †

[1] *Adjusted for gender and maternal education*
[2] *Adjusted for number of siblings and maternal education*
[3] *Adjusted for number of siblings and gender*

Table 2. Adjusted odds ratios for being in the bottom 10% of the observed scores distributions for behavioural outcomes from the Strengths and Difficulties Questionnaire (SDQ) by (i) number of siblings, (ii) gender and (iii) mother's education

	Sample size	Number of siblings (reference: 1 sibling)						Gender (reference: girl)			Mother's education (reference: High educated)					
		Only-children			≥2						Middle			Low		
		OR[1]	CI95%	*p*	OR[1]	CI95%	*p*	OR[2]	CI95%	*p*	OR[3]	CI95%	*p*	OR[3]	CI95%	*p*
Total difficulties age 3	14,395	1.10	0.93-1.29		1.20	1.04-1.38	**	1.58	1.40-1.78	***	2.15	1.77-2.62	***	5.47	4.44-6.73	***
Total difficulties age 5	14,559	1.34	1.11-1.61	***	1.39	1.20-1.61	***	1.60	1.41-1.81	***	2.46	2.04-2.97	***	5.65	4.61-6.92	***
Total difficulties age 7	13,338	1.67	1.35-2.07	***	1.36	1.17-1.59	***	1.84	1.60-2.12	***	2.12	1.76-2.56	***	4.26	3.46-5.24	***
Peer problems age 3	14,633	1.42	1.23-1.64	***	1.21	1.06-1.38	**	1.30	1.15-1.46	***	1.47	1.28-1.70	***	2.59	2.24-3.00	***
Peer problems age 5	14,684	1.49	1.22-1.81	***	1.16	0.97-1.40	*	1.48	1.29-1.69	***	1.45	1.19-1.77	***	2.98	2.46-3.61	***
Peer problems age 7	13,427	1.95	1.58-2.39	***	1.21	1.03-1.41	**	1.54	1.33-1.79	***	1.64	1.37-1.96	***	3.24	2.69-3.90	***
Hyperactivity age 3	14,616	1.08	0.91-1.27		0.95	0.80-1.12		1.82	1.57-2.10	***	1.76	1.47-2.11	***	2.98	2.42-3.67	***
Hyperactivity age 5	14,627	1.25	1.07-1.46	**	1.07	0.92-1.24		1.86	1.64-2.12	***	2.06	1.72-2.46	***	3.51	2.90-4.24	***
Hyperactivity age 7	13,397	1.31	1.05-1.64	**	1.20	1.03-1.41	**	2.17	1.86-2.53	***	1.85	1.51-2.26	***	2.57	2.09-3.17	***
Emotional Symptoms age 3	14,727	1.01	0.84-1.23		1.20	1.01-1.41	**	1.05	0.91-1.20		1.56	1.31-1.85	***	3.49	2.90-4.20	***
Emotional Symptoms age 5	14,694	1.20	0.99-1.44	*	1.21	1.05-1.38	**	0.89	0.79-0.99	‡	1.41	1.21-1.64	***	2.72	2.32-3.18	***
Emotional Symptoms age 7	13,418	1.14	0.87-1.49		1.27	1.07-1.50	**	1.14	0.99-1.32	*	1.70	1.43-2.04	***	2.98	2.45-3.64	***
Conduct Problems age 3	14,754	0.80	0.68-0.94	‡	1.14	1.01-1.30	**	1.26	1.13-1.41	***	1.64	1.41-1.92	***	3.89	3.31-4.58	***
Conduct Problems age 5	14,713	0.89	0.73-1.09		1.31	1.14-1.51	***	1.57	1.38-1.78	***	1.74	1.48-2.03	***	3.73	3.16-4.40	***
Conduct Problems age 7	13,447	1.06	0.85-1.33		1.34	1.17-1.55	***	1.72	1.52-1.94	***	1.86	1.56-2.22	***	3.82	3.14-4.64	***
Prosocial behaviour age 3	14,651	0.67	0.53-0.84	‡‡	1.31	1.12-1.54	***	1.60	1.38-1.87	***	0.93	0.77-1.12		1.65	1.33-2.05	***
Prosocial behaviour age 5	14,712	0.76	0.59-0.97	‡	1.37	1.18-1.60	***	2.29	1.97-2.66	***	1.17	0.99-1.37	*	1.79	1.51-2.12	***
Prosocial behaviour age7	13,451	1.02	0.77-1.37		1.31	1.10-1.56	***	2.73	2.31-3.24	***	1.14	0.92-1.41		2.01	1.62-2.50	***

Notes. *Increased risk at p≤0.002* *** *Increased risk at 0.002<p≤0.05* ** *Increased risk at 0.05<p≤0.15* *
Decreased risk at p≤0.002 ‡‡ *Decreased risk at 0.002<p≤0.05* ‡ *Decreased risk at 0.05<p≤0.15* †

[1] *Adjusted for gender and maternal education*
[2] *Adjusted for number of siblings and maternal education*
[3] *Adjusted for number of siblings and gender*

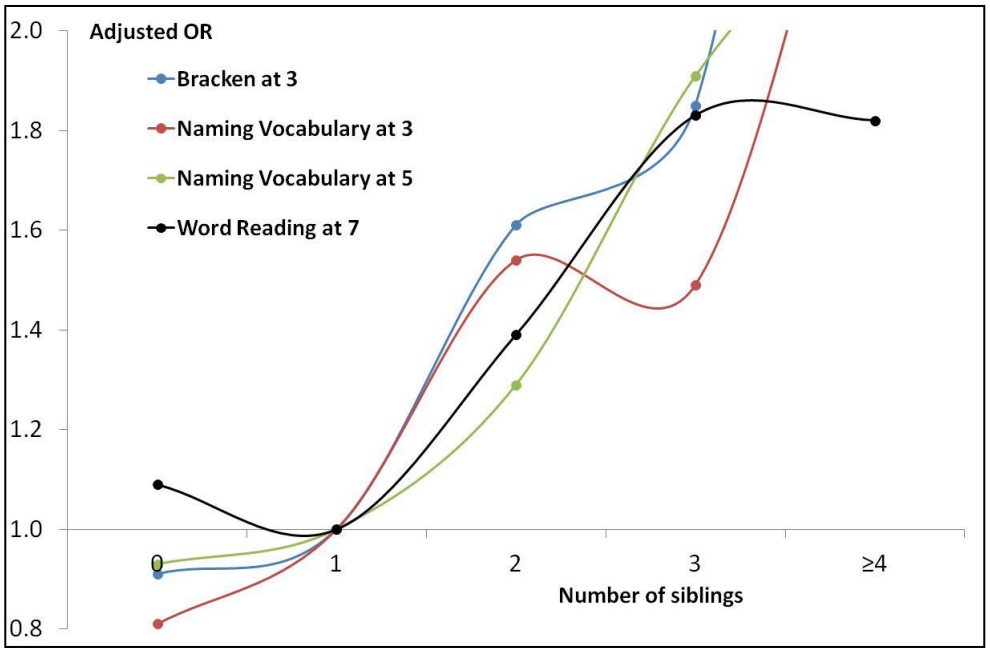
We turn to multi-child families in Tables 1 and 2 (column "number of siblings, ≥2). Compared to cohort children having just one sibling, children with two or more siblings nearly always exhibited a significantly raised risk of low scores. The adjusted odds ratios are never above 2, but they come closest to it for the Bracken and BAS Naming Vocabulary (OR between 1.6 and 1.9). The other cognitive outcomes and behaviour problems nearly all had adjusted odds ratios in the range 1.2 to 1.4. There was no significant difference in the risk of hyperactive problems between children with one or more than one sibling, at ages 3 and 5. Peer problems and lower levels of pro-social skills, might have been supposed to be less prevalent where there are plenty of siblings from whom to learn social skills, but we do not find such a pattern. On the whole, having several siblings was associated with behavioural and cognitive problems.

We also undertook a supplementary analysis by splitting those with two and more siblings into three sub-groups: 2 siblings, 3 siblings, and 4 and more siblings (Figures 3 and 4). There were too few families with 5 or more siblings to treat separately, and even between 3 and 4 siblings the confidence limits on odds ratios tend to overlap. For all cognitive outcomes and the total difficulties score, there is a tendency to an increase in the adjusted odds ratios as the number of siblings rises from one to 4 or more. This pattern was also observed to some degree for conduct and peer problems elements, but the relationship with sib size is less clear for the other sub-scales of the SDQ.

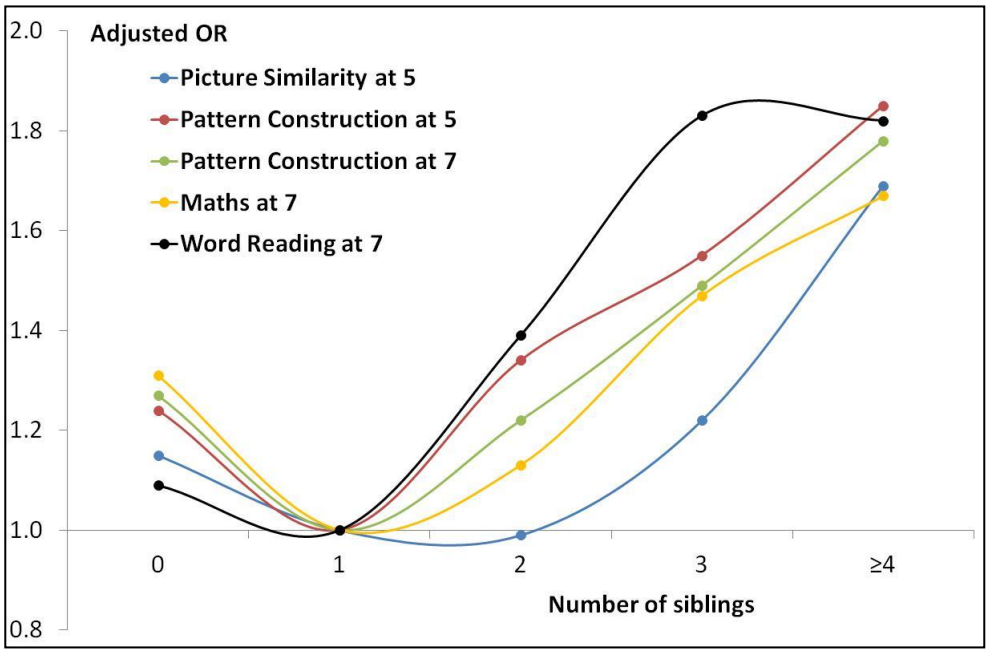
To provide some sense of the salience of these differences, Tables 1 and 2 also show the mutually adjusted ORs for the other two variables in the models. For the cohort chid's gender, girls were set as the reference category, thus an adjusted OR above 1 indicated that boys score worse than girls (all else being equal). In 23 out of 26 outcomes, boys are significantly more at risk of having a low score on cognitive and behavioural outcomes (Tables 1 and 2). The only exception is the emotional symptoms sub-scale of the SDQ at age 5, where girls have a higher risk. These odds ratios are mostly in the same range as the ratios comparing ≥2 siblings with one sibling. However, the estimates for gender and number of siblings are dwarfed by those estimated for mother's education. Compared to mothers with high education, the adjusted ORs for mothers with the low level of education were around 6 for the Bracken and the BAS Naming Vocabulary, around 4 for other cognitive outcomes and around 3 for the different sub-scales of the SDQ. We tested for interactions of the number of sibs with child sex and with mother's education, but found nearly nothing significant (results not shown). The relationships between child's outcomes and number of siblings do not appear to vary for families where there might well be different levels of resources to spread between siblings. The information on social background captured in mother's education is a considerably stronger correlate of child outcomes than either sibling number or gender.

Figure 3. Odds ratios for being in the bottom 10% of the observed score distributions for cognitive outcomes, for four categories of sibling numbers relative to 1 sibling, adjusted for index child's gender and mother's education (UK Millennium Cohort)

a) Bracken at age 3, Naming Vocabulary at age 3 and 5, Word Reading at age 7[(1)]



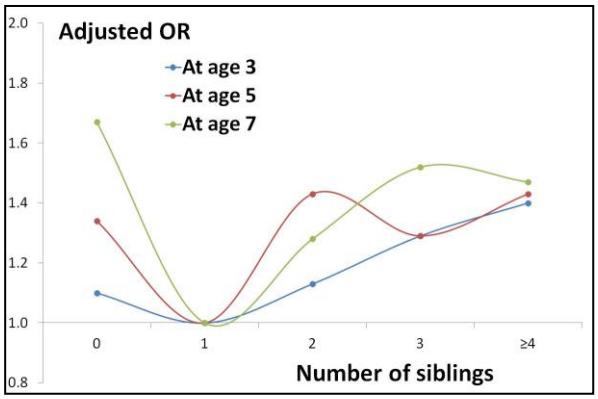
b) Picture Similarity at age 5, Pattern Construction at age 5 and 7, Maths at age 7, Word Reading at age 7



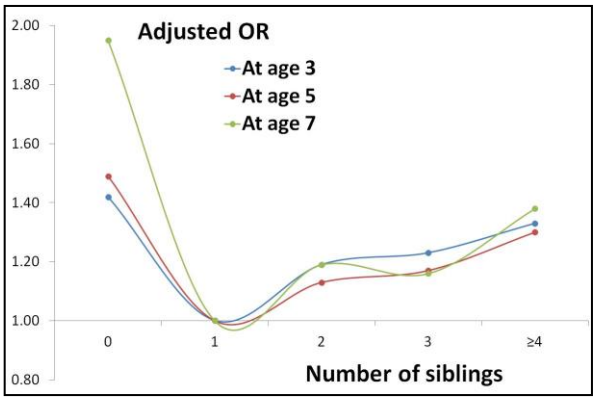
Notes.
(1) When the number of siblings is ≥4, the adjusted OR for the Bracken at age 3 was of 3.92 [CI95%: 3.08, 4.97], the adjusted OR for the BAS Naming Vocabulary was of 2.63 [CI95%: 2.09, 3.31] at age 3 and 2.31 [CI95%: 1.83, 2.91] at age 5.

Figure 4. Odds ratios for being in the bottom 10% of the observed score distributions for behavioural outcomes, for four categories of sibling numbers relative to 1 sibling, adjusted for index child's gender and mother's education (UK Millennium Cohort)

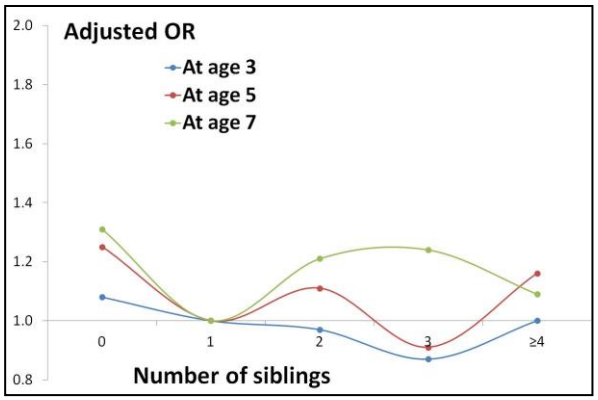
a) Total difficulties



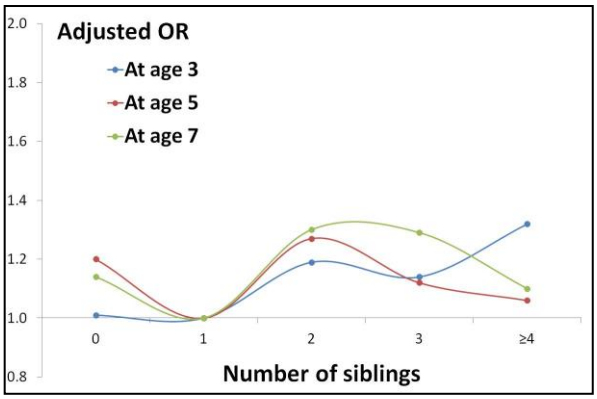
b) Peer problems



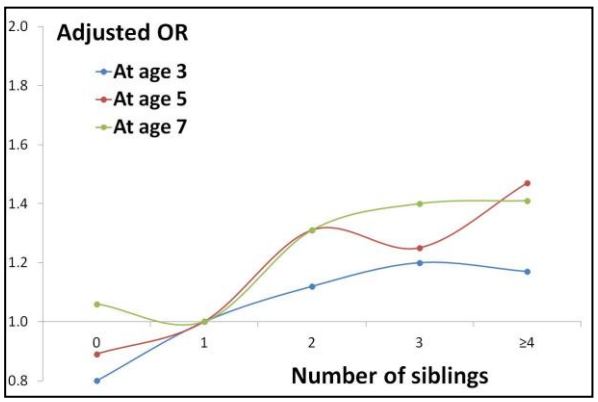
c) Hyperactivity



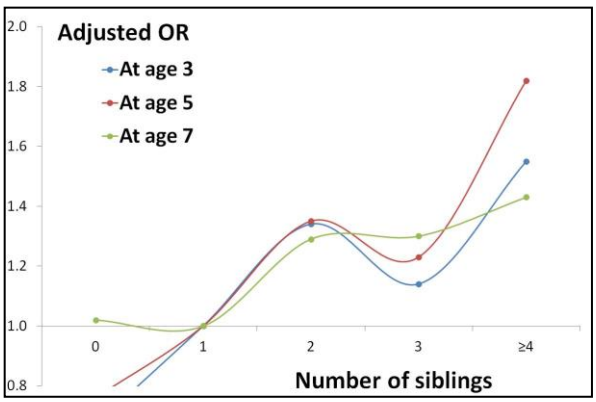
d) Emotional symptoms



e) Conduct Problems



f) Prosocial Behaviour [(1)]



Notes.
(1) When the number of siblings is 0, the adjusted OR for the SDQ Prosocial Behaviour was of 0.67 [CI95%: 0.53, 0.84] at age 3 and was of 0.76 at age 5 [CI95%: 0.59, 0.97]. These low values reflect high levels of prosocial behaviour.

Conclusion

Looking across a broad range of indicators of developmental difficulties, we observed that for all outcomes but one (hyperactivity), children with two or more siblings appear at disadvantage compared to children with only one sibling. The differentials associated with having several siblings are more pronounced for outcomes with a strong verbal component than other cognitive and behaviour outcomes. Such a result is in accordance with the resource dilution hypothesis, even though these cross-sectional descriptions are not an adequate test of a causal relationship of harm or gain from having siblings. The hypothesis of social gains from multiple siblings does not seem to fit with our results on behavioural problems. In particular, children in large families do not appear to be protected from peer problems or problems on the prosocial scale and, somewhat contrary to expectations, only-children, at least at 5 and under, have fewer prosocial problems than children with a sibling.

Comparing only-children with those having one sibling presents a more complex pattern. Being an only-child appears an advantage for outcomes with a strong verbal component at young age, but a disadvantage for other cognitive outcomes. On behavioural outcomes, being an only-child could be either an advantage or a disadvantage depending on the sub-scale, however the advantage seems to disappear for older lone children. Bear in mind that the only-child at 5 and 7 years is in a diminishing minority (16%-12%) whereas there were more of them at 3 years (25%). Thus, the fact that the contrast between no sibling and one sibling changes with the child's age, reminds us that family building is a dynamic and possibly a selective process. The apparent 'disadvantages' of being an only-child could reflect factors which select their parents out of further childbearing. There may be something particular about the families who do not proceed to have further children, which helps account for the modest extra developmental difficulties we have recorded here. Another hypothesis would be that developmental problems with one child may even inhibit progression to a next birth.

Although this is only a first step in the investigation of whether siblings help or hinder a child's development, we can report to other researchers who may wish to control for siblings in their investigations of other topics, that it would be better to control for the number of siblings as grouped variables than just a linear term. In this research note, several outcomes have been investigated in order to draw a general picture of differentials by number of siblings. However, each outcome would need to be investigated and discussed more thoughtfully. Moreover, before we can reach any conclusions about the mechanisms whereby siblings may impact on child development, it will be necessary to make use of evidence about the top end of the distributions as well as those in most difficulty, and of the longitudinal nature of the data, to see how children's scores change over time when the number of siblings changes and to model other factors. The most obvious one is birth order (Black, Devereux & Salvanes, 2005; Lawson & Mace, 2010). Is it an advantage to be the first-born regardless of the number of younger sibs who come along? Is it particularly important to be shielded from resource dilution in the earliest years of life? One could also consider, the interval between births, the age of the cohort child, the siblings' gender, whether the cohort birth or the next one had been intended, the quality of relationship between siblings, the arrival and departure of non-natural siblings, other predictors of parental investment and child development, parental partnership stability, the home learning environment, parenting practice and other circumstances.

References

Azmitia, M., & Hesser, J. (1993). Why Siblings Are Important Agents of Cognitive Development: A Comparison of Siblings and Peers. *Child Development 64*, 430-444.

Becker, G.S., & Tomes, N. (1976). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy 84*, S143-S162.

Black, S.E., Devereux, P.J., & Salvanes, K.G. (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics 120*, 669-700.

Blake, J. (1981). Family Size and the Quality of Children. *Demography 18*, 421-442.

Bland, J.M., & Altman, D.G. (1995). Multiple significance tests: the Bonferroni method. *British Medical Journal 310*, 170.

Bowes, L., Maughan, B., Caspi, A., Moffitt, T.E., & Arseneault, L. (2010). Families promote emotional and behavioural resilience to bullying: evidence of an environmental effect. *Journal of Child Psychology and Psychiatry and allied disciplines 51*, 809-817.

Bracken, B.A. (1998). *Bracken basic concept scale - revised (BBCS-R)*. San Antonio: Psychological Corporation.

Brody, G.H. (2004). Siblings' Direct and Indirect Contributions to Child Development. *Current Directions in Psychological Science 13*, 124-126.

Cameron, L., Erkal, N., Gangadharan, L., & Meng, X. (2013). Little emperors: behavioral impacts of China's One-Child Policy. *Science 339*, 953-957.

Coleman, J.S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology 94*, S95-S120.

Downey, D.B. (1995). When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance. *American Sociological Review 60*, 746-761.

Downey, D.B. (2001). Number of siblings and intellectual development. The resource dilution explanation. *American Psychologist 56*, 497-504.

Downey, D.B., & Condron, D.J. (2004). Playing Well with Others in Kindergarten: The Benefit of Siblings at Home. *Journal of Marriage and Family 66*, 333-350.

Elliott, C.D., Smith, P., & McCulloch, K. (1997). *British Ability Scales Second Edition (BAS II): Technical Manual*. London: nferNelson.

Gass, K., Jenkins, J., & Dunn, J. (2007). Are sibling relationships protective? A longitudinal study. *Journal of Child Psychology and Psychiatry and allied disciplines 48*, 167-175.

Goodman, A., Lamping, D.L., & Ploubidis, G.B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. *Journal of Abnormal Child Psychology 38*, 1179-1191.

Guo, G., & VanWey, L.K. (1999). Sibship Size and Intellectual Development: Is the Relationship Causal? *American Sociological Review 64*, 169-187.

La Rochebrochard (de), E., & Joshi, H. (2013). Children born after unplanned pregnancies and cognitive development at 3 years: social differentials in the United Kingdom Millennium Cohort. *American Journal of Epidemiology*, advance access published July 25, 2013.

Lawson, D.W., & Mace, R. (2009). Trade-offs in modern parenting: a longitudinal study of sibling competition for parental care. *Evolution and Human Behavior 30*, 170-183.

Lawson, D.W., & Mace, R. (2010). Siblings and childhood mental health: evidence for a later-born advantage. *Social Science & Medicine 70*, 2061-2069.

Office for National Statistics. (2012). Chapter 3: Fertility. In *National Population Projections, 2010-based reference volume* (pp. 1-16). UK: ONS.

Smith, T.E. (1993). Growth in Academic Achievement and Teaching Younger Siblings. *Social Psychology Quarterly 56*, 77-85.

Steelman, L.C., Powell, B., Werum, R., & Carter, S. (2002). Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges. *Annual Review of Sociology 28*, 243-269.

Stone, L.L., Otten, R., Engels, R.C., Vermulst, A.A., & Janssens, J.M. (2010). Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: a review. *Clinical Child and Family Psychology Review 13*, 254-274.

Vostanis, P. (2006). Strengths and Difficulties Questionnaire: research and clinical applications. *Current Opinion in Psychiatry 19*, 367-372.

Wolke, D., & Skew, A.J. (2012a). Bullying among siblings. *International Journal of Adolescent Medicine and Health 24*, 17-25.

Wolke, D., & Skew, A.J. (2012b). Family factors, bullying victimisation and wellbeing in adolescents. *Longitudinal and Life Course Studies 3*, 101-119.

Zajonc, R.B., & Markus, G.B. (1975). Birth order and intellectual development. *Psychological Review 82*, 74-88.

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 the Journal Manager **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 5000 words**, with longer papers also accepted up to a maximum of 7000 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.

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 Harvard Style 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 the Journal Manager. 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. Particular care should be taken by the author to check references.

Proofs. The corresponding author will be asked to view a layout proof of the article on the website and respond with final amendments within 3 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.

CONTENTS

PAPERS

KEYNOTE LECTURE

RESEARCH NOTE

This journal can be accessed online at: www.llcsjournal.org

Published by
The Society for Longitudinal and Life Course Studies
Institute of Education, 20 Bedford Way, London, WC1H 0AL
info@slls.org.uk http://www.slls.org.uk