VOLUME 6 | ISSUE 2 | **APRIL 2015** ISSN 1757-9597

Longitudinal and Life Course Studies: International Journal

Published by

SLLS Society for Longitudinal and Life Course Studies

Special Issue

Generation X enters middle age

Inside this Special Issue

- British 1970 Birth Cohort Study age 42
- Origins and contexts
- Missing data
- Obesity
- Vocabulary
- Social mobility
- Religion

Co-sponsors:

VU university medical center

ISER INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

SRC SURVEY RESEARCH CENTER UNIVERSITY OF MICHIGAN

LLCS EDITORIAL BOARD

SUBSCRIPTIONS

For immediate access to current issue and preceding 2014 Issues (Volume 5):

Annual Individual rate: **£20 (discounted rate £55 for 3 years)**.

Annual Library **rate: £200** to register any number of library readers.

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

SLLS members have automatic free access to the journal among many other benefits. For further information about SLLS and details of how to become a member, go to http://www.slls.org.uk/#!services/ch6q

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

Print on demand

An attractive printed version of each Issue of the journal, including all back numbers, is available at a price of £10 (plus postage and packaging). If you would like to purchase a full set or individual copies, visit the SLLS Bookshop at http://www.slls.org.uk/#!journal-bookshop/czkc Depending on distance, you will receive the order within two weeks of submitting your order.

Please note

The reselling of personal registration subscriptions and hard copies of the journal is prohibited.

CONTENTS

INTRODUCTION

PAPERS

LLCS 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

SLLS Society for Longitudinal and Life Course Studies

SLLS International Conference

'Life Courses in Cross-National Comparison: Similarities and Differences'

Dublin, Ireland

18 - 21 October 2015

(Pre-conference workshop on 18.10.15, with main conference starting on 19.10.15)

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

Confirmed keynote speakers:
Prof. Elizabeth Cooksey, Ohio State University, USA
'Intergenerational Linkages, Interdisciplinary Collaborations and Interesting Findings: Lessons Learned from 50 Years of the National Longitudinal Surveys';
Prof. Richard Layte, Economic and Social Research Institute (ESRI), Ireland
'National Crises and Personal Troubles: The Great Recession in Ireland and Family Processes';
Prof. Rainer Silbereisen, Center for Applied Developmental Science (CADS), Germany
'Dealing with Perceived Uncertainties in Work and Family and their Psychosocial Consequences: Comparative Studies in Germany and Poland in Times of Social Change'.

The conference theme will focus on life courses in cross-national comparison this year, but will cover other areas of longitudinal and life course study: physical, psychological, social developmental and ageing processes and functioning within and across life course stages from infancy to old age; methods and findings of cohort studies; other sources of longitudinal data such as panel studies and record linkage; international comparisons; household, and income dynamics; intergenerational transfers and returns to learning; gene-environment interactions; 'mixed', and comparative methods; innovative methodology in design, measurement, data management, analysis and research practice (quantitative and qualitative).

The fee for conference registration will be discounted for SLLS members so please consider joining if you have not already done so! Please visit the membership page of the website for full details of how to join online: http://www.slls.org.uk/#!services/ch6q

More conference information will be posted on the SLLS website as it becomes available. Registration will open in July 2015.

www.slls.org.uk

GUEST EDITORIAL

Generation X enters middle age

Alice Sullivan Institute of Education, University College London, UK
a.sullivan@ioe.ac.uk
Matt Brown Institute of Education, University College London, UK
David Bann Institute of Education, University College London, UK

(Received April 2015 Revised April 2015) http://dx.doi.org/10.14301/llcs.v6i2.351

Introduction

The 1970 British Cohort Study (BCS70) is an ongoing longitudinal study of people born in the United Kingdom in a particular week in 1970. It is the third study in Britain's world-renowned series of birth cohort studies, following the 1946 and 1958 birth cohorts. Whereas the 1958 cohort were part of the 'lucky generation' of post-war baby boomers, the 1970 cohort represent 'Generation X', a term popularised by Douglas Coupland's novel (Coupland, 1991), which portrays this generation struggling with anomie and uncertainty. Compared to the generations born in 1946 and 1958, the 1970 cohort were subject to an era of economic and labour market transformation and turbulence (Ashton and Bynner, 2011), and experienced growing socio-economic inequalities (Goodman and Webb, 1994, Hood and Joyce, 2013). They can be seen as entering middle age with a degree of reluctance, as child-bearing has been delayed to later ages compared to previous generations. The notion that this generation is more troubled than the previous generation gains some support from the higher levels of mental distress reported by the 1970 cohort members at age 42. This special issue asks how this generation is faring in mid-life, and addresses a diverse range of themes including social mobility, obesity and religion.

This introduction is presented in three parts. First, we provide an overview of the social and historical context in which the 1970 cohort have lived and some of the differences between the 1970 cohort and the other British cohorts. The aim here is to provide contextual information to support data analysis and interpretation for researchers who are less familiar with the national historical context in which the BCS70 cohort members have lived their lives. Second, we provide information on the history of the study, and on the data. Third, and finally, we give a summary of the papers presented in this special issue.

Historical context

Politics and Economy

The 1970 cohort were born at a time of low unemployment (3%), but this was also a time of deindustrialisation and industrial strife. Strike action led to a three-day week in 1973-4. Continuing industrial unrest culminated in the 'winter of discontent' in 1978-9. From a child's-eye view, this was a time of blackouts and candlelit evenings.

In 1971, as BCS70 approached their first birthdays, Britain gave up pounds, shillings and pence in favour of a decimal currency. Britain joined the European Community (EC) in 1973. This became the European Union in 1993.

The 1970s also saw the dawn of the troubles in Northern Ireland. Fourteen unarmed protesters were killed in Derry on Bloody Sunday in 1972. The troubles made their mark on BCS70. Northern Ireland was included in the birth survey, but dropped thereafter.

Margaret Thatcher became Britain's first female Prime Minister in 1979, when the cohort members were nine years old. She won three terms in office, remaining in power from 1979 to 1990, when the cohort members were 20 years old. As such, she defined the era politically, and those growing up in the 1980s were often known as 'Thatcher's

children', implying that they were defined by the individualistic and materialistic attitudes summed up in Thatcher's famous statement that there is 'no such thing as society'.

Under the pressures of technological transformation and globalisation the period was also marked by a massive restructuring of employment as traditional manufacturing industries associated with coal and steel lost out to the more efficient high tech driven industries of the 'tiger economies' of the far East. The consequence was a breakdown of the customary transition routes to adulthood to be replaced by a mishmash of extended education and national training schemes. These originated from the Thatcher government's 'New Training Initiative' (1991) promoted to equip young people with the skills employers needed and imbued with risk and uncertainty (Banks et al., 1991). The outcome, finally, was a return in 1986 to the much maligned and discarded (now 'modern') apprenticeship, which has lasted ever since.

The worldwide recession of the early 1980s added to this labour market upheaval, leading to high unemployment, breaking the three million barrier in 1982. Young people in the BCS70 cohort entering adolescence paid a particularly high price in terms of later joblessness. High levels of social unrest were evident at the time, as shown for example by the Brixton riots of 1981.

Nevertheless, the Falklands war helped to propel Thatcher to a landslide election victory in 1983. The Thatcher government marked the end of the post-war corporatism and consensus, and the start of an increasingly economically unequal society. Extensive privatisation of state-run industries took place, including the sell-offs of the electricity, gas and water industries and the railways. Council housing was also sold off and private home ownership promoted (35% of cohort members were living in council housing at age five). House prices rose dramatically both during the 1980s and subsequently, with consequences for the 1970 cohort — at age 42, the number of home owners was lower among BCS70 cohort members than it had been for the cohorts born in 1958 or 1946 (Table 1), with an increase in private renting. The miners' strike of 1984-1985 emphasised Thatcher's determination to break the trade unions. Internationally, this period saw a strong alliance between Thatcher and US President Ronald Reagan, and a reawakening of Cold War tensions with the Soviet Union.

As the 1970 cohort entered their late teens, the economy recovered with the 'Lawson boom' of the late 1980s. A Poll Tax (i.e. a fixed tax per head) was introduced in 1990, replacing local taxes based on property values. This proved extremely unpopular, and difficult to enforce, with high rates of non-payment. Demonstrations against the Poll Tax in 1990 were followed by rioting. Thatcher stood down as Prime Minister in 1990, after losing the support of Conservative MPs. She was replaced by John Major. The period 1990-1993 was marked by another long recession. This came at an important time for the cohort members, many of whom would have been relatively new entrants to the labour market or, in the case of university graduates, entering it for the first time. Despite the economic slump, John Major won the 1992 General Election, which was the first general election that cohort members were eligible to vote in, at age 22. Internationally, 1992 saw the fall of the Berlin wall, symbolising the end of the Soviet bloc. Major initiated the peace process in Northern Ireland, which ultimately culminated in the Good Friday agreement in 1998, marking an end to the Troubles. After nearly twenty years in opposition, the Labour Party, led by Tony Blair took power in 1997, when the cohort members were 27 years old. The leadership had rechristened the party 'New Labour' to emphasise it's turn away from the left-wing policies of the past. The period from 1997-2008 was characterised by economic growth and relatively low unemployment, despite the bursting of the dot-com bubble in 2000. In the early 2000s, the Blair government embarked on costly wars in Iraq and Afghanistan. The international financial crisis of 2008, originating in the banking collapse, and the subsequent recession, ended the long period of economic stability. Labour lost power in the general election of 2010, but the Conservatives did not win an outright majority, and entered government in coalition with the Liberal Democrats. By 2015, real incomes for working-age people were still below their 2007/2008 levels (Cribb, Hood and Joyce, 2015).

Education

Following the Plowden Report (Plowden, 1967) 'progressive' teaching methods were fashionable yet controversial in primary education during the

1970s. Familiar debates about academic selection and 'falling standards' raged during this period, including in a series of 'Black Papers' attacking comprehensives, egalitarianism and progressive teaching methods. Labour Prime Minister Callaghan called for a 'great debate' on the nature and purposes of education in a speech at Ruskin College in 1976. Many of the concerns he raised are still unresolved: high levels of poor basic skills among the British population, poor standards of numeracy among school-leavers, and poor take up of science and technology subjects, especially among girls.

For all the education debates of the 1970s, the school system that the 1970 cohort experienced was remarkably little different from that experienced by the previous generation born in 1958. The Grammar and Secondary Modern schools had continued their slow decline, and the vast majority of the cohort attended comprehensive schools intended for all abilities. Nevertheless, the old two-tier system of examinations at age 16 remained in place. Pupils who were deemed academically able enough took O (Ordinary) levels, while the less able took CSEs (Certificate of Secondary Education). Aligned with the development of youth training, vocational qualifications were overhauled in 1986 leading to the advent of the subsequently marginalised National Vocational Qualifications (NVQs).

The unified GCSE (General Certificate of Secondary Education) system was introduced in 1988, two years after the 1970 cohort would have taken their exams. A high proportion (54%) of the 1970 cohort left school at age 16, and for those who continued, the academic track A (Advanced) level course remained. This highly specialised path meant that pupils typically took only three academic subjects between the ages of 16 and 18.

Teacher pay was eroded in real terms from 1980 onwards. Long-running strike action over pay began in 1984 and culminated in an all-out strike in 1986. The long-term action included a 'work to rule' whereby out-of-hours sporting and cultural activities were not provided. The education of cohort members and their participation in wider activities was affected, although the coverage of the strike varied between schools. Fieldwork for the age 16 wave of BCS70 was unfortunately greatly disrupted by the strike.

Those cohort members who qualified for university entrance would typically have done so at age 18 in 1988. The proportion of the cohort gaining a degree by age 42 seems low by contemporary standards (25%), yet is much higher than it had been for those born in 1958 and 1946 (Table 1). Cohort members faced a two-tier higher education system made up of universities (high status) and polytechnics (lower status). This divide was abolished by the 1992 Education Act, also signalling a shift from elite to mass higher education. So, both at school and at university, the 1970 cohort missed being affected by the major educational reform acts of their era by a few years.

Technology

Televisions were affordable in the 1970s, and the 1970 cohort grew up with television as an accepted part of family life. However, there were only three TV channels until Channel 4 was launched in 1982, and the 1970 cohort did not grow up with TV available all day. Breakfast television was first launched in 1983. The 'test card' featuring a girl and clown which showed during downtime in the BBC schedule remains an iconic image of a 1970s childhood. 1970s children had to find things other than TV to occupy much of their time.

The cohort has experienced extraordinary levels of technological change across many domains. As teenagers they listened to electropop cassettes on their Sony Walkmen. They taped the first episodes of Eastenders on their brand new video recorders. However, the most notable change has been the revolution in computing and IT. Home computers came into the mainstream in the early 80s. The ZX Spectrum computer came onto the market in 1982, and many of the boys (but fewer girls) of the 1970 cohort would have owned one of these inexpensive machines, which were mostly used for games. BCS70 cohort members would have used pen-and-paper for all their schoolwork and for their university essays in the early 1990s. However by age 34 in 2004, 75% of them were using computers at work. The internet became hugely important commercially with the dot-com boom of the late 90s. Almost all (97%) of cohort members had internet access at home at age 42.

Women and families

The sex-discrimination act of 1975 outlawed discrimination against women by employers. Women's labour market participation increased

greatly in the 1970s and 1980s, and mothers of the 1970 cohort members were more likely to have worked during their children's infancy and early childhood than was the case for earlier generations. The girls of the 1970 cohort achieved equal levels of school success as the boys, and were similarly likely to obtain a university education (25% in men and 25.5% in women; Table 1). Compared with earlier cohorts, the BCS70 women had greater labour market opportunities. However, the gender pay gap has remained (Joshi, Makepeace and Dolton, 2007).
The roles of men and women have changed dramatically. The majority of the 1970 cohort grew up in 'intact' two-parent families (80% lived with both of their natural parents at age 16). However, family structures and relationships in mid-life are more heterogeneous for this cohort than for previous cohorts, with increased divorce and single-living. Many of this generation delayed childbearing, so they are more likely to have young children than previous generations in early mid-life.

Lifestyles and health

While advances in medicine have enabled a greater capacity to treat both infectious and chronic diseases, societal changes have affected lifestyles in ways that have not always been beneficial for health. Compared with cohorts born in 1958 and 1946, the 1970 cohort were more frequently exposed to smoking during pregnancy, less likely to have been breastfed, yet less likely to have contracted measles as children, due to the availability of the measles vaccine (Table 1). This cohort experienced the rise of convenience foods during their teenage years in the 1980s, increased car ownership and declining activity levels, increased alcohol consumption (especially among women), but lower levels of smoking (Schoon and Parsons, 2003). During childhood, they were no more likely to be overweight than the previous cohort, born in 1958. However, by the age of 16, BCS70 were fatter than teenagers of previous generations. How adult health in the BCS70 compares with previous cohorts warrants careful empirical study and is likely to depend on the component of health considered. For example, malaise scores (capturing psychological distress) were higher in the BCS70 than the 1958 cohort (Table 1), suggesting that mental health may have worsened, although these differences may be explained by cohort differences in the propensity to report psychological problems.

Demography

Whereas the 1958 and 1946 cohorts were part of the 'baby boomer' generation, by 1970, the total fertility rate had fallen to 2.4. Unlike previous cohorts, the mothers of the 1970 cohort had access to the contraceptive pill and to abortion via the 1967 abortion act. However, the proportion of participants with at least one child by age 42 was only slightly lower in BCS70 than in it had been in the 1958 cohort (Table 1), suggesting that the idea of an increase in 'child-free' living for this generation has been exaggerated.
When the 1970 cohort was born, immigration was low and the proportion of ethnic minorities was very small (4%). Immigration has subsequently risen dramatically, resulting in a very different ethnic profile, even among the general population of the same age as BCS70, but particularly among younger people. The UK population in 1970 stood at 55.66 million. By 2000 it had increased to 58.9 million, but by 2012 it had shot up to 63.7 million (Source: ONS). This increase was driven in part by open borders within the EU, which expanded to include Eastern European countries in 2004.

Table 1.

	NSHD (1946 cohort)			NCDS (1958 cohort)			BCS (1970 cohort)			MCS (2000 cohort)		
	M	F	Total	M	F	Total	M	F	Total	M	F	Total
Child characteristics												
Smoked during pregnancy				33.7%	33.6%	33.6%	46.2%	46.1%	46.2%	36.3%	35.1%	35.7%
Ever breastfed	75.9%	74.7%	75.3%	67.7%	69.0%	68.3%	35.8%	36.6%	36.2%	67.1%	66.8%	67.0%
Measles by 10 or 11	88.2%	89.7%	88.9%	91.9%	92.6%	92.2%	50.0%	49.8%	49.9%	12.4%	11.2%	11.8%
Adult characteristics – Age 42/43 for NSHD												
High malaise (Age 42)*				9.6%	16.3%	13.0%	16.0%	20.2%	18.2%			
Lives with spouse	81.0%	81.3%	81.2%	69.1%	69.3%	69.2%	59.2%	58.7%	59.0%			
Lives with cohabiting partner	6.0%	3.8%	4.9%	11.5%	11.5%	11.5%	17.6%	16.0%	16.7%			
Lives with no partner	13.0%	14.9%	13.9%	19.5%	19.2%	19.3%	23.2%	25.3%	24.3%			
Has at least one child	83.3%	89.1%	86.2%	77.5%	82.8%	80.2%	74.4%	81.1%	77.9%			
Has a degree	11.2%	6.0%	8.7%	19.4%	16.2%	17.8%	25.0%	25.5%	25.3%			
Homeowner (mortgage paid or owned)	80.9%	81.8%	81.4%	81.4%	79.9%	80.6%	73.8%	73.8%	73.8%			

*based on the 9 item malaise scale; high malaise = a score of more than 4 out of 9

Note: estimates for the 1946 cohort were weighted to account for the stratified sampling design.

History of the study

BCS70 is a multi-purpose and multi-disciplinary study which has collected detailed information from cohort members on many different aspects of their lives. The birth survey was medically focused, but with each follow-up the scope of enquiry has broadened. The early surveys at five, ten and 16 tracked the physical, social and educational development through childhood and into adolescence. Adult sweeps have gathered information about all of the key domains of life including housing; cohabiting relationships; fertility and births; children and wider family; family income and wealth; economic activity; life long learning (qualifications achieved and training); health and health behaviours and social participation. Response rates at each follow-up are presented in the first paper in this issue by Mostafa and Wiggins.

BCS70 began as the British Births Survey (BBS), with information collected about the births and social circumstances of over 17 000 babies born in England, Scotland, Wales, and Northern Ireland (Elliott and Shepherd, 2006). A questionnaire was completed by the midwife who had been present at the birth and, in addition, information was extracted from clinical records. The study aimed to examine the social and biological characteristics of the mother in relation to neonatal morbidity. The BBS was sponsored by the National Birthday Trust Fund in association with the Royal College of Obstetricians and Gynaecologists. The founding Principal Investigator (PI) was Roma Chamberlain, and the study was based at the University of Bristol. Fieldwork was funded by the Medical Research Council, National Birthday Trust Fund, Department of Health and Social Security and Royal College of Obstetricians and Gynaecologists. After the initial birth survey, Northern Irish participants were dropped from the sample.

Sub-studies were carried out at 22-months and 42-months as part of a project known as the British Births Child Study, 1972/73. This was designed to explore the effects of foetal malnutrition on the development of the child. These sub-studies involved a 10% random sample alongside all twins, post-mature and growth retarded births to married mothers. A number of publications reported specifically on the sub-studies (Chamberlain and Davey, 1975; 1976; Chamberlain and Simpson 1977; 1979).

The study remained at Bristol for the age five and ten follow-ups, with Neville Butler coming in as PI. In 1975 and 1980 parents of the children in the study were interviewed by health visitors, and information was gathered from the child's class teacher and head teacher, from the school health service, and from the children themselves. The 1975 wave was funded by the Medical Research Council. The age ten survey had a range of funders: Joseph Rowntree Memorial Trust, Department of Education and Science, Department of Health and Social Security, Manpower Services Commission and National Institute of Child Health and Development.

In 1986, Neville Butler left Bristol University and set up his own charity, the International Centre for Child Studies (ICCS). Controversially, he removed all the BCS70 records and control of the study to ICCS. Neville Butler was an energetic and well-connected fundraiser for the survey. His personal friends included Margaret Thatcher, and his fundraising parties at the Dorchester and Claridges were attended by aristocrats such as the Marquess of Bath. A news release (Butler, 1984) featuring the text of a speech given by Neville Butler at the London Stock Exchange provides an insight into these fundraising activities. Robert Maxwell spoke at the event and donated a covenant of £5,000 yearly for seven years.

The 1986 follow-up of BCS70 was known as 'Youthscan' and comprised 16 separate survey instruments, including parental questionnaires, class teacher and head teacher questionnaires, and medical examinations. In addition to completing educational assessments, the cohort members themselves answered questionnaires on a wide range of different topics and were asked to keep two four-day diaries, one on nutrition and one on general activity. It was originally planned to trace cohort members in time to interview them at 15.5 years old, well before the minimum school leaving age. Unfortunately, industrial action by teachers, who were responsible for the educational tests, meant that the survey was delayed, and resulted in incomplete data collection from schools.

Through 1989 Neville Butler was in discussion with John Bynner and colleagues at the Social Statistics Research Unit (SSRU), City University about the future of BCS70, and in 1991 SSRU took over responsibility for the study with John Bynner as director. At this point, no data had been deposited with the Economic and Social Research

Council (ESRC) since the age five wave (the deposit of the age five data had been brought together by Jean Golding, who later became founding director of the Avon Longitudinal Study of Parents and Children (ALSPAC)). With funding from ESRC, SSRU made a commitment to deposit the later data. BCS70 moved from the International Centre for Child Studies (ICCS) to SSRU in 1991. In 1998, SSRU moved from City University to the Institute of Education (now UCL Institute of Education) to become the Centre for Longitudinal Studies (CLS) with Director John Bynner becoming PI of both BCS70 and the National Child Development Study (NCDS) until his retirement in 2003.

A 10% sub-sample follow-up of BCS70 was carried out in 1991 directed at the basic skills of cohort members and funded by the Adult Literacy and Basic Skills Unit and the Paul Hamlin Foundation. A postal follow-up of the full cohort was conducted in 1996 when study members were aged 26. This survey was funded primarily by ESRC, and ESRC has consistently supported the study subsequently.

In 2000, BCS70 combined with NCDS to undertake, for the first time, a joint survey, funded by ESRC. This simultaneous survey of the BCS70 and NCDS cohorts was undertaken to facilitate comparisons between these two groups born 12 years apart (Ferri, Bynner and Wadsworth, 2003). This study restored the BCS70 sample to over 11, 000 and established a baseline for the scientific content of the adult surveys, ensuring that all major life domains were covered. Heather Joshi, Director of CLS from 2003, took over responsibility as PI from 2003-2004, and Jane Elliott took over as PI subsequently.

The age 34 wave carried out in 2004 was a full interview-based survey of the cohort and incorporated a component of basic skills assessment funded by the government-established National Research and Development Centre for Adult Literacy and Numeracy (NRDC) based in CLS. The survey also included an NRDC-funded 'Child Study', involving collection of data from 50% cohort members and their children via self-completion questionnaires and assessments. The age 38 follow-up took the form of a 30 minutes telephone survey. In 2010, Alice Sullivan became PI for the age 42 survey, which consisted of a one hour face to face interview and a self-completion questionnaire conducted in cohort members' homes. As well as collecting details about key experiences and circumstances since the time of the prior interview, the age 42 survey covered a range of new topics including career help from parents and other family and friends, an assessment of vocabulary, sexuality, class identity, childlessness, reading preferences, television watching and belief in God.

There are a number of books that provide a useful general introduction to health data in the early waves of the study. The birth sweep provided a valuable insight into patterns of obstetric and neonatal care in the United Kingdom (Chamberlain, Philipp, Howlett & Masters, 1978; Chamberlain, Chamberlain, Howlett & Claireaux, 1975). The birth and five-year findings relating to health were discussed in Butler, Golding and Howlett (1986), and general findings from the first two sweeps were also outlined in Osborn, Butler and Morris (1984). The age 26 survey is reported in Bynner, Ferri and Shepherd (1997). More recently, Ferri, Bynner and Wadsworth (2003) have provided an overview of intergenerational changes between the 1946, 1958 and 1970 cohorts, including changing health and lifestyles. Wadsworth and Bynner (2011) provide the complementary historical context for all three British birth cohort studies over the period since the end of the second world war.

Future Plans

The next follow-up will take place in 2016 when cohort members will be aged 46 and will take the form of a biomedical survey conducted by nurses. The biomedical follow-up will provide an invaluable resource for investigating the longitudinal predictors of health in mid-life. It will also provide a baseline for research on ageing, as risk factors will be measured prior to functional decline and disease becoming apparent. The survey will address the major public health agendas faced by this generation, including obesity, sedentary lifestyles, and mental health and wellbeing. Many of the planned measures were included in the biomedical follow-up of the 1958 cohort (NCDS) which took place in 2002-3 when cohort members were aged 44-45 meaning that cross-generational comparisons will be possible.

Planned measures include anthropometric measurements (height, weight, waist measurement and hip measurement), physical functioning (grip strength and standing balance), blood pressure and resting heart rate, and cognitive assessments.

Blood will be collected for a full range of future analyses. A 45 minute interview is planned, which will include the core content covered in all adult sweeps and additional detail on mental and physical health and health behaviours.
It is currently planned that future follow-ups will occur at 50 and then every five years thereafter.

Papers

Missing data

Our special issue opens with a consideration of the important issues of attrition and non-response in birth cohort data. Mostafa and Wiggins consider the extent of non-response in BCS70 and its effect on the composition of the sample between 1970 and 2012. They find that, although men from lower social class backgrounds and with less educated parents are less likely to respond, at a statistically significant level, the predictive power of their non-response models is weak. The authors illustrate the use of weights and Multiple Imputation respectively to address the problem of differential non-response. Using predictors of non-response to construct non-response weights does not improve the efficacy of the illustrative models presented in the analysis, largely because the predictors are weak. This paper presents the case that Multiple Imputation provides a superior solution, enabling the researcher to restore sample size with partial information, and coming closest to the 'benchmark' model with complete data in terms of parameter estimates and standard errors.

Mostafa and Wiggins emphasise the importance of making best use of the available data for longitudinal analysis. With each wave of data collection, attrition increases, and for longitudinal analyses exploiting several waves of data, problems due to wave and item non-response can multiply. The option of ignoring cases with partial information therefore becomes increasingly untenable over time, particularly for those who wish to exploit the rich data available and to take a genuine life course perspective. For researchers using BCS70 data, the possibilities of using Multiple Imputation are particularly exciting given the variable levels of response in past waves, and high levels of instrument non-response in the 1986 wave in particular.

Obesity

Costa, Johnson and Viner examine trajectories into overweight and obesity by age 42. Only a minority (30%) of the cohort had never experienced overweight or obesity at age 42. For this generation, only a small group (6%) became overweight or obese by age ten. The most common pattern was for individuals to become overweight or obese in early adulthood, i.e. at age 26 to 30, and remain overweight by age 42 (44%). However, stark gender differences were apparent. A majority of men (63%) became overweight or obese in early adulthood, compared to 31% of women. Only 15% of men never experienced overweight or obesity up to age 42, compared to 40% of women.

Although only a minority of this cohort actually became overweight as children, the childhood roots of adult overweight and obesity were made clear by this analysis. Both the mother's and the father's Body Mass Index (BMI) when the cohort member was a child predicted not just childhood overweight and obesity but also early-adulthood onset overweight and obesity. A high social class position (either parent having a professional or managerial job) during childhood was a protective factor against overweight and obesity, while evidence of early puberty (by age ten) substantially increased the risk of early-adulthood onset overweight and obesity.

Vocabulary

The paper contributed by myself, Alice Sullivan, and Matt Brown, examines vocabulary development between adolescence (age 16) and middle age (age 42). We build on earlier work where we examined the role of reading in progress in vocabulary and mathematics between the ages of ten and 16 (Sullivan & Brown, 2015 in press). The age 16 vocabulary test was repeated at age 42 – the first time that one of the childhood cognitive tests has been repeated in adulthood. We found that, on average, cohort members' vocabulary scores increased substantially between the ages of 16 and 42.

We examined the role of reading habits both in childhood and adulthood in predicting this vocabulary growth. In adulthood, we were able to examine the genres of books that the cohort members read as well as how much they read. Both the quantity and quality of reading varied enormously according to educational level. We were interested to discover a large cultural divide in the types of books that graduates of Russell Group

and other universities read. For example, nearly half (48%) of Russell Group graduates read 'contemporary literary fiction' compared to nearly a third (30%) of graduates of other universities.
We found that reading habits in both childhood and adulthood influenced vocabulary development between the ages of 16 and 42. In adulthood, the type as well as the quantity of reading was important, those who read high-brow novels made the greatest vocabulary gains. High levels of post-16 educational and occupational attainment were also linked to high vocabulary gains between 16 and 42.

Social mobility

Much attention and debate has focussed on the levels of social mobility experienced by the 1970 cohort as compared to past generations. Gutierrez, Micklewright and Vignoles' paper examines one of the potential mechanisms through which parents pass their advantage (or lack thereof) on to their children. To what extent do social networks and assistance from parents, other family and friends make a difference to people's labour market chances?

At age 42, the cohort members were asked a retrospective question on help received from parents, family and friends in getting a job. The types of help that they were asked about included providing advice, recommending the cohort member to an employer, directly employing them, and helping them to get a job via their social networks. The authors found that there was a clear social gradient in the levels of parental help received, with those from higher social class backgrounds receiving more help. However, the results show that those who received parental help did not have higher earnings or higher social class positions at age 42 than those who did not report receiving any help. This may be because not all help is equally beneficial. Those cohort members who reported that the help they had received had been important to their careers did have higher earnings than those who reported that their parents' assistance had not contributed to their careers.

Religion

The final paper in the special issue reports on 'The mysteries of religion and the lifecourse'. For the first time, the 2012 survey asked cohort members questions about belief in God and life after death, alongside questions about religious affiliation and practice, which have been asked previously.
Voas's paper reveals the apparent unreliability over time of cohort members' responses on religion. Between 2004 and 2012, nearly a quarter of people changed their minds about whether they had been raised in a religion. Between 1996 and 2012, more than a quarter changed whether they saw themselves as belonging to a religion. This must give us pause for thought as survey researchers. When measuring something which respondents may have vague ideas about, small changes in question wording or context can have an enormous effect on the apparent findings.

Belief in God was gauged by asking cohort members: 'Which of these statements below comes closest to expressing what you believe about God?'. The responses were: I don't believe in God (22%); I don't know whether there is a God and I don't believe there is any way to find out (21%); I don't believe there is any way to find out (21%); I don't believe in a personal God but I do believe in a Higher Power of some kind (14%); I find myself believing in God some of the time but not others (12%); While I have doubts, I feel that I do believe in God (19%); and I know God really exists and I have no doubts about it (12%). Nearly half of cohort members (48%) believed in life after death, but belief in God and life after death did not always go together. For example, only two thirds of those who believed in God with no doubts also said that there was definitely life after death, while 26% of agnostics believed in the hereafter. Voas also finds that those who say they have a religion do not necessarily mean that they believe in God. Among mainline Protestants (Church of England, etc.) only nearly half (48%) believed in God with the remainder expressing various shades of agnosticism or disbelief. The religious beliefs of the British appear far too muddy to give much comfort to either the established churches or advocates of clear-sighted atheism.

A striking finding from this analysis is the large gender gap in belief. 54% of men reported atheist or agnostic views, compared to 34% of women. Why this should be the case remains an unresolved question. But we can look forward to BCS70 being used increasingly as a resource to help us to understand the predictors of religious belief as well as its consequences later in life.

Conclusions

The papers presented in this special issue give some idea of the wide range of topics and questions that may be addressed using BCS70 data. Rather than presenting the final word on any of these subjects, the aim has been to promote the possibilities of the survey, and encourage as wide as possible a range of researchers from all disciplines to consider how they could exploit this rich resource.

Acknowledgements

Many thanks to Karen Mackinnon and Diana Kuh for invaluable help with the NSHD comparisons presented in Table 1.

References

Ashton, D. and Bynner, J. (2011) 'Labour Market, Employment and Skills', in Wadsworth, M. E.J. and Bynner, J. (2011) *A Companion to Life Course Studies: The social and historical context of the British birth cohort studies*. London: Routledge.

Banks, M., Bates, I., Breakwell, G., Bynner, J., Emler, N., Jamieson, L. and Roberts, K. (1992) *Careers and Identities*. Buckingham: Open University Press.

Brown, M.J. & Elliott, J. (2010). *Family and household profiles: Comparing the 1958 and 1970 Birth Cohorts*. Twenty-First Century Society, 5, 81-101. http://dx.doi.org/10.1080/17450140903483991

Butler, N. (1984). News release: International Centre for Child Studies. *Early Child Development and Care, 16*, 301-315. http://dx.doi.org/10.1080/0300443840160313

Butler, N.R., Golding, J. & Howlett, B.C. eds. (1986). *From Birth to Five: A study of the health and behaviour of Britain's five year olds*. Oxford: Pergamon Press.

Bynner, J., Ferri, E. and Shepherd, P. (eds.) (1997), *Twenty-something in the 90s: getting on, getting by, getting nowhere*. Aldershot: Ashgate Press

Chamberlain, G., Philipp, E., Howlett, B.C. & Masters, K. (1978). *British Births: Vol 2 Obstetric Care*, London: Wm Heinemann.

Chamberlain, R. & Davey, A. (1975). Physical growth in twins, postmature and small-for-dates children. *Archives of Disease in Childhood, 50*, 437-442. http://dx.doi.org/10.1136/adc.50.6.437

Chamberlain, R. & Davey, A. (1976). Cross-sectional study of developmental test items in children aged 94 to 97 weeks: Report of the British Births Child Study. *Developmental Medicine and Child Neurology, 18*, 54-70. http://dx.doi.org/10.1111/j.1469-8749.1976.tb03605.x

Chamberlain, R., Chamberlain, G., Howlett, B.C. & Claireaux, A. (1975). *British Births: Volume 1. The first week of life*. London: Wm Heinemann.

Chamberlain, R.N. & Simpson, R.N. (1977). *Cross-sectional studies of physical growth in twins, postmature and small for dates children*. Acta Paediatrica Scandinavica, 66, 457-463. http://dx.doi.org/10.1111/j.1469-8749.1976.tb03605.x

Chamberlain, R.N. & Simpson, R.N. (1979). *The prevalence of illness in childhood*. Tunbridge Wells: Pitman Medical.

Coupland, D. (1991). *Generation X: Tales for an accelerated culture*. Macmillan.

Cribb, J., Hood, A. & Joyce, R. (2015). *Living Standards: Recent Trends and Future Challenges*. London: IFS.

Elliott, J. & Shepherd, P. (2006). Cohort Profile: 1970 British birth cohort (BCS70). *International Journal of Epidemiology, 35,* 836-843. http://dx.doi.org/10.1093/ije/dyl174

Ferri, E., Bynner, J. & Wadsworth, M. eds. (2003). *Changing Britain: Changing lives*. London: Institute of Education.

Goodman, A. & Webb, S. (1994). *For richer, for poorer: The changing distribution of income in the UK, 1961 - 1991.* Fiscal Studies, 15, 29-62. http://dx.doi.org/10.1111/j.1475-5890.1994.tb00209.x

Hood, A. & Joyce, R. (2013). *The economic circumstances of cohorts born between the 1940s and 1970s*. London: Institute for Fiscal Studies.

Joshi, H., Makepeace, G., & Dolton, P. (2007). More or less unequal? Evidence on the pay of men and women from the British birth cohort studies. *Gender, Work & Organization*, *14*(1), 37-55. http://dx.doi.org/10.1111/j.1468-0432.2007.00331.x

Osborn, A.F., Butler, N.R. & Morris, A.C. (1984). *The Social Life of Britain's Five Year Olds. A report of the child health and education study*. London: Routledge and Kegan Paul.

Plowden, B. (1967). *Children and their Primary Schools: A report of the Central Advisory Council for Education*. London: HMSO.

Schoon, I. & Parsons, S. (2003). Lifestyle and health-related behaviour, in E. Ferri, J. Bynner and M. Wadsworth (eds) Changing Britain, *Changing Lives: Three generations at the turn of the century*. London: Bedford Way Papers.

Sullivan, A.. & Brown, M., (2015 in press) Reading for pleasure and attainment in vocabulary and mathematics. *British Educational Research Journal*

Wadsworth, M. E.J. and Bynner, J. (eds.) (2011) *A Companion to Life Course Studies. The social and historical context of the British birth cohort studies*. London: Routledge (Paperback 2014)

Longitudinal and Life Course Studies 2015 Volume 6 Issue 2 Pp 131 – 146 *ISSN 1757-9597*

The impact of attrition and non-response in birth cohort studies: a need to incorporate missingness strategies

Tarek Mostafa Institute of Education, University College London, UK
T.Mostafa@ioe.ac.uk
Richard D. Wiggins Institute of Education, University College London, UK

(Received June 2014 Revised October 2014) http://dx.doi.org/10.14301/llcs.v6i2.312

Abstract

This paper reveals the extent of attrition in the British Cohort Study begun in 1970 (BCS70) and how it affects sample composition over time. We examine the determinants of response and then construct inverse probability weights (IPWs) to adjust for sample loss. Secondly, we create a hypothetical substantive data set from BCS70 across data collection waves 3 and 4 to illustrate the effectiveness of the use of weights and multiple imputations (MI) in handling the impact of unit non-response and item missingness respectively. Our findings show that when the predictive power of the response models is weak, the efficacy of non-response weights is undermined. Further, multiple imputations are effective in reducing the bias resulting from item missingness when the magnitude of the bias is high and the imputation models are well specified.

Keywords: BCS70, attrition, unit non-response, item non-response, inverse probability weights, multiple imputation.

Introduction

This paper has two objectives. Firstly, to provide a historical account of the extent of unit non-response in the British Cohort Study begun in 1970 (BCS70) across the nine waves of data collection between 1970 and 2012 together with an illustration of the construction of inverse probability weights (IPWs) to adjust for unit non-response. Secondly, we illustrate the impact of using IPWs and the application of multiple imputation (MI) for an artificially constructed set of patterns of missingness under a substantive question for analysis. The construction of IPWs under the first objective are based upon a set of birth characteristics for cohort members (CMs) because these values are available for almost the entire sample of CMs originally recruited into the study. This second objective is to assess the impact of non-response weights and imputation techniques on the bias resulting from unit non-response and item missingness respectively by conditioning on a fully observed subset of data from wave 4. The next section presents a brief overview of the literature on approaches to handle missing data and more background to response patterns in BCS70. We follow this section with a description of wave-by-wave non-response in BCS70 and examine the extent to which the representativeness of the study may be revealed over time in terms of the cumulative loss of CMs according to their birth characteristics. There then follows a section to illustrate the impact of the use of IPWs and MI. The final sections provide a brief discussion and conclusion. Our underlying motivation for this paper is to help raise users' appreciation of the need to incorporate strategies to handle missingness in any longitudinal analysis of BCS70 and, indeed birth cohort studies more generally.

Background

Statistical description and analysis are persistently challenged by the problem of missing data (Little & Rubin 2002). Survey samples are threatened by both unit non-response and individual item missingness where a respondent fails to provide all of the information requested

(also known as partial non-response). In longitudinal surveys, the problem of maintaining co-operation with CMs over time adds another dimension to the problem of non-response. Attrition refers to situations where CMs drop out of a study and never return, and situations where individual CMs have an interrupted response pattern over time. For instance, a respondent may not co-operate during a particular wave of a study and then return subsequently creating an instance of 'wave non-response'. These patterns are distinguished as monotone and non-monotone response, respectively.

Missing data constitutes a problem for two reasons. First, missingness leads to the loss of observations and to the reduction of sample size. For instance, in BCS70 if only CMs who have responded in all nine waves (since 1970) are considered available for any substantive analysis, the resulting sample would represent only 20 per cent of the original sample of 17,284 CMs. Secondly, missingness may lead to selection bias and inaccurate inference. In order to tackle these problems and make best use of all of the available data in any analysis it becomes necessary to make assumptions about how a proportion of our data came to be missing at all. Many authors (for example Carpenter & Plewis, 2011) term this process to be 'the missingness mechanism'. Following Rubin (1976), Little and Rubin (2002), Carpenter and Plewis (2011) and others we adopt a typology of missingness mechanisms described as 'missing completely at random' (MCAR), 'missing at random' (MAR) and 'missing not at random' (MNAR). Briefly, MCAR implies that the probability of not answering a particular question is uncorrelated with the characteristics of the respondent, and in any longitudinal survey it means that the probability of dropping-out from any wave is uncorrelated with the characteristics of the CM. MCAR is a very strong assumption to make since missingness is more likely to be at random (MAR) or not at random (MNAR). Under MAR the probability of non-response to a question or the probability of dropping-out from a particular sweep are related to some of the observable characteristics of the respondent such as gender, social class, or educational level. Under MNAR, the probability of item non-response or the probability of dropping-out from a particular wave is related to characteristics or traits yet to be observed (or simply described as unobserved variables). If the pattern or nature of missingness is related to any observable or unobservable variables then ignoring it would lead to the loss of a particular type of respondents (e.g. men, the less well educated) and hence the sample will no longer be random or representative of the parent population.

Historically, two broad approaches have been typically adopted to tackle the problem of missing data involving the application of weighting adjustments (IPWs) to compensate for unit non-response and/or multiple imputation (MI) to handle item non-response (Kalton (1986) and Lepkowski (1989)). Both adjustment strategies assume MAR. Under MI, there are several different approaches appropriate to the type of data and/or data structure (Carpenter & Plewis, 2011). In this paper we demonstrate that analyses which draw upon the longitudinal history of the cohort require strategies to use all of the available data, including any partial information (incomplete wave responses).

The application of weights or IPWs in longitudinal surveys is typically used to adjust or re-balance the distributions of the responders so that the relative importance of each CM's characteristic in any particular wave is reweighted according to the importance of the characteristics of those who dropped out. In other words, if the survey is losing men over time, then men will be given a relatively higher weight than women (see Hawkes & Plewis (2006), Plewis (2007a) for illustrations using the Millennium Cohort Study (MCS), Plewis (2007b) and Seaman and Wright (2011) for a review). The probability of response at each sweep is estimated using logistic regression models for a binary outcome (response or non-response) or multinomial-logistic regression models (allowing for more than two outcome categories). These models draw upon CM characteristics as covariates to improve the goodness of fit of the underlying model. Additionally, analysts may use external metadata as auxiliary covariates to enhance efficiency. For useful illustrations, see Plewis (2011), Schouten and de Nooij (2005) and Micklewright, Skinner and Schnepf (2012).

In survey research, there is a long tradition of applying weights to adjust for differential probabilities of selection (the sample design) and/or response. It is fairly straightforward to construct weights to adjust for differential non-response, and relatively easy to apply and make them routinely

available as part of an archived dataset. For example the British Household Panel Survey (BHPS) (Taylor, Brice, Buck & Prentice-Lane, 2010) and MCS both provide longitudinal weights to adjust for cumulative drop-out over time as well as wave-specific or cross-sectional weights. However, they have a number of disadvantages:

1. Weights (both longitudinal and cross-sectional or wave specific) adjust for unit non-response at the level of the individual CM. Implicitly the same weights adjust for all item response appertaining to a lost individual. There may be more effective solutions available to handle item missingness, namely MI as described below.
2. If variables x, y and z are used in predicting unit non-response, and thus in the construction of weights, the results of analyses using x, y and z as dependent and independent variables will yield unbiased results. However, if there is a fourth omitted variable, w, that happens to be strongly related to x, y and z, then the inclusion of w could well improve the efficiency of the construction of weights.
3. Under conditional regression applications, if we are regressing an outcome variable from sweep t+1 on a number of independent variables collected during an earlier sweep t where attrition has possibly occurred, the weighted analysis will be constrained to using only the non-missing cases in both waves (Goldstein, 2009). This further undermines the efficacy of non-response weights because they will only adjust for non-response in one wave (usually the wave in which the dependent variable was observed). As an alternative, there is the possibility of combining IPW and MI as a remedy (Wiggins, Schofield, Bartley, Sacker & Head, (2004) McDonald & Ketende (2009) and Seaman, White, Copas & Li (2012)).

Under MI, (Little & Rubin (2002), Schafer & Olsen (1998) and Rubin (1987, 2004)) missing values under MAR assumptions are replaced several times to create filled-in replicates of our data. These replicates of multiply-imputed data are analysed separately and ultimately combined under Rubin's Rules (Rubin, 1987) to provide parameter estimates which take account of the uncertainty introduced into any analysis by filling-in under the MAR assumption. There are a number of approaches to MI which vary according to assumptions made about the type of data to be imputed (e.g.) multivariate Normal together with the application of transformation before and after imputation, for categorical data and/or where data is missing at varying levels of aggregation or clustering, (i.e. for multilevel or hierarchical data), and various software packages offer these solutions. The interested reader is encouraged to consult Carpenter and Kenward (2013) for a valuable overview of MI and its application.

MI approaches appear more complex to use than weights but their application presents two main advantages:

1. MI can handle both the treatment of item and unit non-response. Indeed wave non-response can be considered a special case of item missingness where all variables are missing for the same respondent within a longitudinal record.
2. MIs can be custom-made according to the needs of the researcher. When properly specified, they are robust and generate valid inference. MI can be implemented according to the structure of the data (e.g. for handling a multilevel structure see Goldstein, Carpenter & Browne (2014) and the type of variables (e.g. continuous, ordinal or multinomial variables) see Nathan (1983), Nathan & Holt (1980), Pfeffermann (2001) and Carpenter & Plewis (2011)).

According to Carpenter and Kenward (2013) in their chapter entitled 'Sensitivity analysis: MI unleashed', in order to explore how robust inferences are to the assumption of MAR, analysts have an obligation to impute data under MNAR or at least 'approximate the results of doing so'. One approach that may appeal to the user is known as 'joint modelling' where the substantive model of interest is modelled jointly with a model for missingness (also referred to as 'Heckman modelling', Heckman (1979)). In this way, it is proposed that the unobserved variables that simultaneously influence both the outcome and the missingness are captured by the residuals in the two models, which are allowed to correlate. Technically, the challenge for the analyst is to identify variables (or instruments) for the missingness model which predict the probability of missingness but do not correlate with the substantive outcome (see Carpenter & Plewis (2011) for an illustration using NCDS data).

Clearly, in this brief overview of various approaches to handling missingness we have placed the use of IPW and MI in the foreground of our coverage simply because good illustrations are available in the literature for users to consult as

well as software to match (e.g. STATA). For more recent methodologies which specifically draw upon the longitudinal nature of information databases, users are referred to the application of MI under a 'two-fold fully conditional specification (FCS)' which fills-in data at time t conditional upon data in time t and adjacent time points t-1 and t+1 (Nevalainen, Kenward & Virtanen (2009) and Welch, Bartlett & Petersen (2014)). Additionally, users may prefer to find solutions to maximising their use of available data in the context of the application of a particular analysis, such as structural equation modelling where the emphasis is upon the temporal structure of relationships. Here 'full information maximum likelihood (FIML)' has been developed under well-known SEM algorithms (e.g. AMOS, Arbuckle (1996)) where item missingness is not handled directly but the likelihood function is adjusted so that incomplete data is used in the estimation under MCAR or MAR (Davey & Savla, 2010). The application of FIML also incorporates the use of auxiliary variables (Enders, 2008).

Non-response in the British Cohort Study (BCS70)

All nine available waves of BCS70 are used to examine attrition and model response in terms of several birth characteristics. These covariates include: gender, father's social class, father's and mother's age at completion of education, mother's age at delivery, whether mother lived in London in 1970, whether or not the CM's mother attempted breast-feeding, her marital status and the number of older siblings at the time of the CM's birth.

In table 1 we summarise the pattern of missing data for BCS70 over the nine waves of data collection from 1970 to 2012. Just under 1 in 5 (19.8 per cent, labelled as non-missing) of the CMs participated in all nine waves, whereas over half (52 per cent, labelled as non-monotone) dropped out from at least one wave but returned to the study in a subsequent wave, and nearly a third (27.2 per cent, labelled as monotone) dropped out from the survey after participating in a number of waves without ever returning, to date. The base sample of 17,284 CMs consists of the original birth sample (i.e. excluding immigrants who joined the study later on).

Table 1. Patterns of missing data in BCS70 (1970 to 2012)

Pattern	Frequency	Percentage
Monotone	4,716	27.2
Non monotone	9,153	53.0
Non missing	3,423	19.8
Total	17,284	100

Table 2 below shows that over 42 years, from birth in 1970 up to and including the ninth wave in 2012, 7,930 CMs dropped out of the study for various reasons. Some have died, others have left Great Britain, while some have refused to participate or disappeared from the study for one or more waves, only to reappear again. The category labelled 'dead' describes the total number of deaths over the nine sweeps, while the category 'unproductive' describes all other possibilities for dropout: e.g. permanent and temporary immigrants, refusals, non-contact. One should note that dropout is not always permanent since some respondents return in later waves.

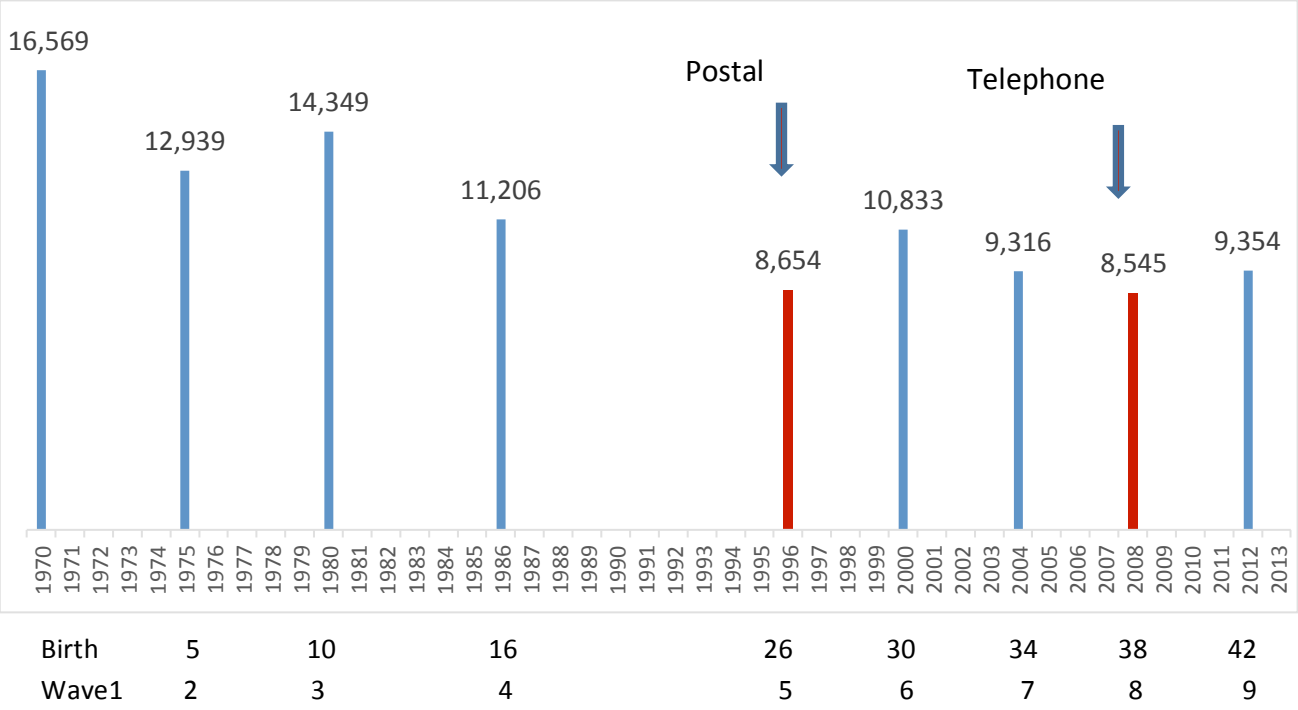
Table 2. Detailed response and non-response categories for BCS70 from 1970 to 2012

Response categories	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8	Wave 9
Age	Birth	**5**	10	16	26	30	34	38	42
Full or partial response	16,569	12,939	14,349	11,206	8,654	10,833	9,316	8,545	9,354
Dead	0	565	585	597	697	748	795	824	853
Unproductive	715	3,780	2,350	5,481	7,933	5,703	7,173	7,915	7,077
Total	17,284	17,284	17,284	17,284	17,284	17,284	17,284	17,284	17,284

Note. The mode of data collection changed between wave four and wave five. In wave five, the data was collected through postal services.

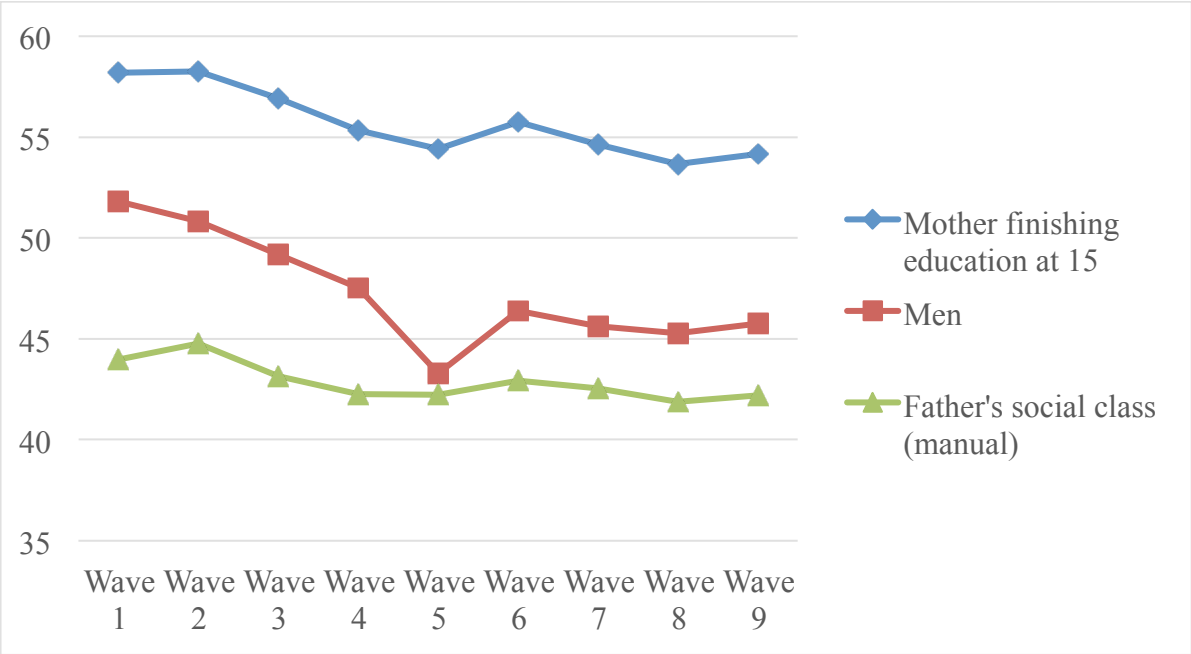
Figure 1 shows that there was a substantial drop in the achieved sample size between age 16 and age 26 years. There are several possible reasons for this drop including the length of the period of 10 years separating the two waves, a teacher's strike at age 16, the use of a self-completion postal survey at age 26, and the fact that the responsibility for CMs to provide consent to participate shifted from their parents to themselves as young adults. Furthermore, the drop in the achieved sample size at age 38 years can also be partially attributed to the use of a telephone survey as a mode of data collection rather than face-to-face interviews. Interestingly, the achieved sample size increased by 4.7 per cent for wave 9 (CM age 42) as some CMs were successfully traced and followed-up. For further information on the data collection, please see the BCS70 Technical Report (Centre for Longitudinal Studies, 2014).

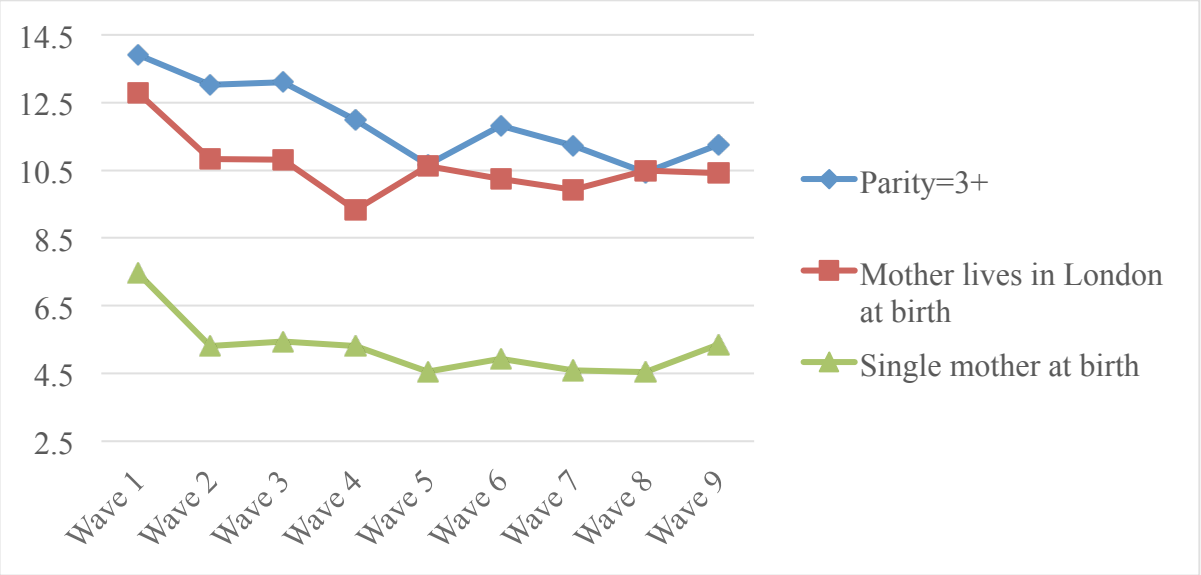
Figure 1. Sample size in the different waves of BCS70



Figures 2 and 3 show the evolution of the sample composition according to CM characteristics recorded at birth. We find that the proportions of male CMs, CMs with fathers who have manual occupations, and CMs with mothers with low levels of education remaining in the study are steadily falling over time. Likewise, the proportions of CMs whose parents were single in 1970, whose mothers were living in London in 1970, and those who have at least three older siblings, have also fallen. It is also worth noting that the rise in sample size in wave 9 (age 42), and the switch from use of the telephone in wave 8 to face-to-face data collection in wave 9, resulted in a change in the point estimates of these proportions. Indeed, they have slightly converged towards their original values at birth. This suggests that the non-response bias in wave 9 will be lower than that in wave 8. Moreover, the differential in the ratio of males to females attained was at a maximum by age 26 (wave 5) and has declined ever since.

Figures 2 and 3. Change in the BCS70 sample composition over time (1970 to 2012). Change is computed in percentages





In general, we can say that men from lower social backgrounds whose parents were single in 1970 are more likely to drop out from the survey. The drop out within these groups could have also been exacerbated by the lack of cohort maintenance. Based on these findings, it is clear that any dropout or lack of co-operation is not a random phenomenon. Differential subject loss according to key birth characteristics will weaken the representativeness of the study, and unless the analysis adjusts for the characteristics which impact upon the probability to respond, inferences will be unreliable. The impact of subject loss may be compounded when we consider the combined effect of these characteristics. We now consider this possibility by using logistic regression analyses for each wave at a time.

What follows in table 3 are the results of a logistic regression of obtaining a response or otherwise for each wave using the birth characteristics as a set of covariates. Response as an outcome is a binary variable taking the value of 1 for those who participated (first category in table 2) and 0 for all other categories including those who died or migrated. Note that sample size is relatively smaller (i.e. 15,270 instead of 17,284) than in table 2 because some CMs had missing birth characteristics.

Table 3. Odds ratios based on logistic regressions of binary response outcome for successive BCS70 waves

	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8	Wave 9
Age	**5**	**10**	**16**	**26**	**30**	**34**	**38**	**42**
Gender (reference: men)								
Women	1.00	1.08	1.26***	1.80***	1.49***	1.48***	1.48***	1.44***
	(0.040)	(0.049)	(0.044)	(0.060)	(0.052)	(0.049)	(0.049)	(0.048)
Marital status (reference: single)								
Married	1.47***	2.18***	1.67***	1.85***	1.89***	1.89***	1.79***	1.42***
	(0.140)	(0.218)	(0.151)	(0.174)	(0.171)	(0.174)	(0.169)	(0.128)
Mother lives in London in 1970 (reference: not in London)								
In London	0.57***	0.55***	0.47***	0.71***	0.61***	0.62***	0.70***	0.67***
	(0.032)	(0.034)	(0.024)	(0.037)	(0.031)	(0.032)	(0.036)	(0.034)
Parity (reference: 0)								
1	0.97	1.02	0.87**	0.92	0.94	0.89**	0.93	0.92*
	(0.050)	(0.059)	(0.039)	(0.039)	(0.042)	(0.038)	(0.039)	(0.039)
2	0.82**	0.89	0.81***	0.79***	0.84**	0.74***	0.75***	0.81***
	(0.053)	(0.065)	(0.046)	(0.042)	(0.047)	(0.040)	(0.040)	(0.044)
3+	0.72***	0.90	0.70***	0.58***	0.65***	0.58***	0.54***	0.61***
	(0.053)	(0.076)	(0.045)	(0.036)	(0.041)	(0.036)	(0.033)	(0.038)
Breastfeeding (reference: attempted)								
Not attempted	0.82***	0.84***	0.85***	0.85***	0.92*	0.87***	0.87***	0.80***
	(0.036)	(0.041)	(0.032)	(0.031)	(0.034)	(0.031)	(0.031)	(0.029)
Mother's age at Delivery (reference: less than 20)								
[20-24]	1.42***	1.17	1.20**	1.31***	1.23**	1.33***	1.28***	1.26***
	(0.105)	(0.098)	(0.080)	(0.085)	(0.080)	(0.085)	(0.083)	(0.081)
[25-29]	1.51***	1.27**	1.28***	1.46***	1.35***	1.50***	1.45***	1.35***
	(0.121)	(0.115)	(0.092)	(0.102)	(0.096)	(0.103)	(0.101)	(0.093)
[30-34]	1.63***	1.36**	1.30**	1.62***	1.44***	1.66***	1.59***	1.39***
	(0.151)	(0.143)	(0.106)	(0.129)	(0.117)	(0.131)	(0.125)	(0.109)
35 or more	1.81***	1.56***	1.40***	1.69***	1.51***	1.81***	1.73***	1.45***
	(0.204)	(0.198)	(0.140)	(0.164)	(0.149)	(0.175)	(0.167)	(0.139)

(Table 3 cont'd)

Mother's age at completion of education (reference: 14 or less)								
15	1.56***	1.81***	1.29**	1.38***	1.20*	1.32***	1.15	1.04
	(0.141)	(0.179)	(0.106)	(0.114)	(0.098)	(0.107)	(0.094)	(0.084)
16	1.63***	1.73***	1.50***	1.50***	1.37***	1.51***	1.34***	1.21*
	(0.164)	(0.190)	(0.137)	(0.135)	(0.124)	(0.134)	(0.119)	(0.107)
17	1.47***	1.42**	1.32**	1.56***	1.26*	1.45***	1.32**	1.18
	(0.172)	(0.180)	(0.138)	(0.160)	(0.131)	(0.148)	(0.134)	(0.120)
18 or more	1.31*	1.34*	1.30*	1.48***	1.14	1.33**	1.24*	1.05
	(0.147)	(0.164)	(0.133)	(0.149)	(0.116)	(0.134)	(0.124)	(0.105)
Father's social class (reference: SC 1)								
Professional	0.94	0.98	0.85	0.94	0.93	0.99	0.95	0.97
	(0.102)	(0.116)	(0.084)	(0.087)	(0.090)	(0.092)	(0.088)	(0.090)
Clerical, non-manual	1.06	1.20	1.04	1.07	1.10	1.13	0.99	1.00
	(0.122)	(0.151)	(0.107)	(0.102)	(0.111)	(0.108)	(0.094)	(0.095)
Skilled manual	0.90	0.94	0.79*	0.79*	0.83*	0.84	0.74***	0.77**
	(0.097)	(0.109)	(0.076)	(0.071)	(0.078)	(0.076)	(0.067)	(0.070)
Unskilled manual	0.87	0.85	0.75**	0.70***	0.76**	0.75**	0.68***	0.69***
	(0.101)	(0.108)	(0.079)	(0.068)	(0.077)	(0.073)	(0.066)	(0.068)
Lowest grade workers	0.70**	0.77	0.69**	0.56***	0.64***	0.65***	0.56***	0.59***
	(0.091)	(0.111)	(0.081)	(0.063)	(0.074)	(0.072)	(0.063)	(0.065)
Other	0.34***	0.60***	0.70**	0.70**	0.69**	0.76*	0.65***	0.70**
	(0.044)	(0.085)	(0.086)	(0.082)	(0.083)	(0.088)	(0.075)	(0.081)
Father's age at completion of education (reference: 14 or less)								
15	1.20*	1.24*	1.11	1.02	1.19*	1.03	1.11	1.03
	(0.102)	(0.119)	(0.083)	(0.076)	(0.089)	(0.076)	(0.082)	(0.076)
16	1.09	1.00	1.14	1.07	1.13	1.00	1.10	0.99
	(0.107)	(0.108)	(0.098)	(0.090)	(0.096)	(0.084)	(0.092)	(0.082)
17	0.92	1.04	1.25*	1.21	1.27*	1.10	1.29*	1.08
	(0.107)	(0.136)	(0.131)	(0.122)	(0.132)	(0.111)	(0.130)	(0.108)
18 or more	0.79*	0.82	0.98	0.96	1.05	1.00	1.06	0.92
	(0.083)	(0.094)	(0.092)	(0.088)	(0.097)	(0.091)	(0.097)	(0.083)
N	15270	15270	15270	15270	15270	15270	15270	15270
pseudo R^2	0.036	0.034	0.026	0.040	0.028	0.031	0.033	0.025

Notes. Exponentiated coefficients; Standard errors in parentheses, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

The pseudo R-squared values for the regressions in table 3 are consistently very small in magnitude, dropping from 3.6 per cent in wave one to 2.5 per cent in wave 9. This indicates that the combined predictive power of birth characteristics is weak even for the early waves. This happens because a large number of variables which affect the probability of response are not accounted for in the model. Such variables may include the characteristics of interviewers and the conditions surrounding the collection of the data. However, metadata are not available in BCS70 and any underlying theory of response as a social process is absent as a means of informing the selection of covariates. On a practical note, of course including additional covariates from other waves will inevitably lead to a reduction in the observations remaining available to model response for a particular wave, due to missingness as the result of previous attrition or item non-response.

Turning our focus to the parameter estimates for the regression results in table 3 we obtain some indicative understanding of the response process. As

expected, women are more likely to respond than men. The effect of gender becomes stronger after wave 3 (CM age 10). Interestingly, the sample becomes more skewed towards women at a time when the responsibility for co-operation is transferred from parents to CMs. Individuals whose parents were married at birth are also more likely to respond than those whose parents were single. In contrast, individuals whose mothers were living in London in 1970, and those whose mothers did not attempt breast-feeding, are less likely to respond. Further, the probability of response drops with parity. The higher the number of older siblings a CM has the less likely he or she is to respond. The probability of response is strictly increasing in relation to the age of mothers at delivery for all waves, and is higher for CMs whose mothers had a longer formal education.

The higher the social class of the CM's father at birth the more likely the CM will respond. However, the effect of social class is only significant for the lowest three social classes (e.g. skilled manual, unskilled manual and lowest grade workers). Father's age at completion of education does not appear to have a significant effect on the likelihood of response. One should note that the effects of the covariates are highly significant in statistical terms except father's social class and father's age at completion of education. However, the explanatory power of these models remains very weak, suggesting that the missingness mechanism is driven by a more complex set of influences than birth characteristics alone.

The results from the logistic regression analysis confirm the findings from the descriptive analysis. In other words, attrition is not a random process and dropout will most likely depend on some of the CM characteristics. Hence, working with only the productive cases from any sweep without any adjustments will lead to bias unless we have MCAR.

In order to make the illustration in the next section, we construct IPWs based on the response model for wave 4 (fourth column in table 3). These weights are to adjust for unit non-response in wave 4 but not for item missingness or unit non-response from previous waves.

Table 4 presents a set of point estimates for the birth characteristics used as covariates in the response probability models (table 3), comparing adjusted and unadjusted estimates with those obtained at birth in wave 1. In wave 1 the sample consists of 15,270 CMs which drops to 10,059 by wave 4 due to unit non-response. The first row gives the percentages for each category at birth (N=15,270), the second row gives the percentages at wave 4 (without the use of IPWs, N=10,059) and the third row gives the percentages at sweep four after adjustment using the IPWs we have constructed (N=10,059).[i]

Table 4. The impact of non-response weights in wave 4

Variables	Men	Father's occupation (skilled manual)	Mother finishing education at 15	Parents are single in 1970
Wave 1 (at birth)	51.87	59.14	45.61	4.78
Wave 4 without weights	49.93	58.59	45.43	3.63
Wave 4 with weights	51.91	59.17	45.59	4.78

Table 4. (continued)

Variables	Mother lives in London	Parity=+3	Number of observations
Wave 1 (at birth)	12.34	13.75	15,270
Wave 4 without weights	9.53	12.44	10,059
Wave 4 with weights	12.31	13.72	10,059

When there is no adjustment for unit non-response the percentage estimates for wave 4 deviate substantially from their original value at birth. This indicates that the sample is biased according to these birth characteristics. In contrast, when the IPWs are applied, the percentages are almost identical to their original values despite the loss of observations due to non-response. Hence, non-response weights are

effective in reducing bias in produced descriptive estimates for our selected birth characteristics. Obviously, the efficacy of these IPWs will be highest when applied to substantive analyses which include the covariates used in the construction of these weights in the explanatory part of the model. Given their weak explanatory power, this will be most unlikely in models which draw upon wider range of predictors.

Next, we will explore the performance of our IPWs and MI for an artificial example based upon BCS70 data. The purpose of the illustration is to assess the effectiveness of weights and imputation techniques in dealing with statistical bias in regression analyses (in terms of both estimates and their standard errors).

An illustration to examine the effectiveness of reweighting and imputations?

The generation of BCS70 datasets used to illustrate the application of re-weighting and multiple imputation

We generate a BCS70 dataset imagining that the analyst wishes to predict vocabulary scores at the age of 16 years (Parsons, 2014) as an outcome based upon gender, the gross family income per week (measured at wave 3 when the CM is aged 10 years) and highest parental qualification (measured at wave 4) as explanatory variables. Income was chosen from a previous wave (wave 3) to the outcome variable in order to illustrate the complexity of working with data collected at more than one time point. Figure 4 below summarises the history of the sample sizes that ultimately generate our example. Sample A refers to the number of CMs in wave 1 (15,270), sample B those CMs available in wave 4 (10,059). The difference between samples A and B is due to cumulative attrition and unit non-response between wave 1 and wave 4. Sample C is a set of complete cases (CC) selected from sample B for the four variables used in our illustration. Finally, we generate sample D from sample C to contain an artificial amount of missing items in a number of distinct steps:

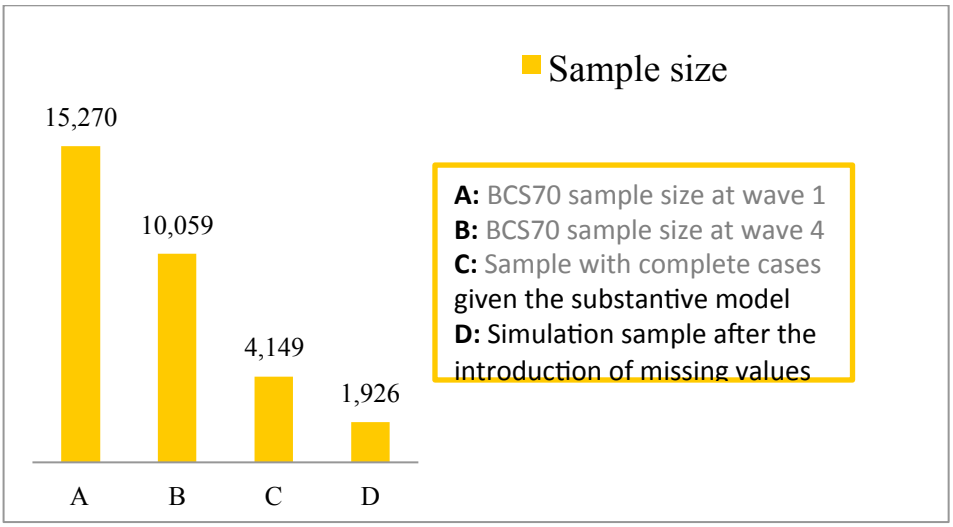
1. For the vocabulary scores we introduce missing values for 10 per cent of the cases in a random or haphazard manner unknowingly related to the values of any other variable.
2. We recode the father's occupation into a binary variable with two categories, manual and non-manual.
3. On income and highest qualification we introduce 40 per cent missing values at random if the father is from the manual group and 10 per cent at random if he is from the non-manual group. We do this for each of the two variables separately. The absolute difference in terms of missing values between the two categories is then 30 per cent. This difference is considered large enough to reveal a bias in the analysis when compared to the CC analysis based on sample C. Father's occupational group is not included in any of our subsequent models.
4. We do not introduce any missing values for gender since it is unlikely to suffer from any item missingness in practice (especially as it is recorded at the birth of a CM).

Finally, the sample size reported for sample D (1926) in figure 4 is that for a reduced set of CC obtained by applying listwise deletion to the modified version of sample C (C_{miss}) after imposing the missingness reported above.

The generated pattern of missingness attempts to provide a semblance of the reality of working with longitudinal data. Even if the application of IPWs were to be efficient in dealing with unit non-response bias, item missingness will still lead to additional bias if it is not completely at random. This bias will vary according to the magnitude of missingness and how much it deviates from MCAR. In this example we introduced item-missingness on income and highest qualification. The magnitude of missingness varied according to the father's social class in 1970 – manual vs. non-manual – with those CMs with fathers from the manual group being less likely to answer the questions. Hence, the item missingness we have introduced is not MCAR, and the extent to which it can be assumed to be MAR or MNAR depends on whether the father's social class is treated as observable (MAR) or unobservable (MNAR) in the imputation procedures. In figure 4, the reduced sample C will artificially represent the 'truth' for purposes of comparing various analyses.

Note that the difference between samples B and C is due to the combination of missing values on the three variables included in the substantive model: vocabulary scores (49.92 per cent item missing for 5,021 CMs), income (17.47 per cent item missing or 1,757), parents highest qualification (11.81 per cent item missing for 1,188 CMs). Note that item missingness is high on vocabulary scores because not all CMs had undertaken this test.

Figure 4. The size of the sample



Modelling strategy

After introducing item missingness into sample data C, we estimate the following models:

Model 1: this model is estimated using sample C (with CC) while applying the non-response weights to adjust for the bias resulting from the unit non-response that is reported between waves 1 and 4 (A-B).This model has no item-missingness because sub-sample C is fully observed or complete for our substantive example. The parameter estimates and standard errors provide a 'benchmark' model to which all other model estimates can be compared.

Model 2: this model is estimated using sample C again but without applying the non-response weights. By comparing this model with model 1 we will be able to ascertain by how much the application of non-response weight adjustment affects the findings irrespective of the need to fill-in any item missingness.

Model 3: this model is estimated using the reduced sample D (with listwise deletion applied after our hypothetical pattern of missingness is imposed on sample C). This model will suffer from both biases (i.e. unit non-response and item missingness) and we are not applying any adjustment technique.

Model 4: this model is estimated using sample D with IPWs applied to adjust for the loss of observations (across samples A to B).

Model 5: this model is estimated under MI (averaging under Rubin's Rules across 20 imputed datasets) which restores the sample size back to that reported for C (4,149). The imputations adjust for the bias resulting from the generation of item missingness (in sample C) but not for any unit non-response.

Model 6: this model is fully adjusted and combines the MI results based upon model 5 in conjunction with the application of IPWs based on the sample loss between waves 1 and 4 (A-B).

It is important to note that none of the models adjusts for the bias resulting from the existing item missingness already present at wave 4 (sample B cases minus C cases). To that extent all models suffer from the same degree of bias, which does not affect their comparability.

The imputations in models 5 and 6 are carried out using MI under a Markov-Chain-Monte-Carlo procedure (Gilks, Richardson & Spiegelhalter, 1996) and chained equations in STATA (Royston (2009) and Royston & White (2011)). We use the MI command with a linear procedure to impute vocabulary scores because the variable is continuous, and ordinal-logit to impute income and highest qualification (the two variables are ordinal). Following the example of Goldstein (2009) we produce 20 imputed datasets.[ii] The explanatory variables used in the imputation model are based on the birth characteristics listed in table 3, with the exception of father's social class at CM's birth: the variables used are gender of the CM, parental marital status, parity, breast-feeding, mother's age at delivery, mother's age at completion of education, and father's age at completion of education. In other words, we treated father's social class as unobservable in the imputation model because we introduced item missingness based on the values of the variable father's social class. Our motivation was to ensure that the data

was more likely to deviate from being fully MAR. In other words, missingness will depend on observable and unobservable factors. In substantive terms, it could be argued that in the analysis model we introduced a proxy for social class by including parental income and education as influences upon a CMs vocabulary score.

We expect that models 5 and 6 will generate the closest results to our benchmark-model 1 in terms of the magnitude of the estimates and their standard errors. Model 3 is expected to generate the least similar results to model 1 since it does not adjust for any type of bias.

Modelling Results

Table 5. Results for regression modelling of vocabulary scores at age 16 years under varying adjustments for unit and item non-response

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Gender						
Women	8.99**	8.41**	5.50	5.93	9.91**	10.4***
	(2.921)	(2.921)	(4.256)	(4.248)	(3.098)	(3.099)
Age 10 gross family income per week (reference: under £50)						
£50 - £99	3.61	2.22	0.90	2.93	1.97	2.78
	(7.464)	(7.538)	(12.054)	(11.919)	(9.508)	(9.532)
£100 - £149	9.40	7.64	2.24	4.45	8.19	8.92
	(7.443)	(7.507)	(11.994)	(11.865)	(8.989)	(9.132)
£150 - £199	14.4	12.9	5.54	7.54	10.0	10.6
	(7.941)	(7.989)	(12.641)	(12.536)	(9.913)	(10.073)
£200 - £249	13.7	11.9	10.5	12.8	17.7	18.7
	(9.184)	(9.195)	(13.898)	(13.820)	(11.240)	(11.468)
£250 or more	28.2**	27.2**	23.4	26.3	27.0*	28.0*
	(9.523)	(9.502)	(14.131)	(14.078)	(11.140)	(11.478)
Parental highest qualification (reference: no qualification)						
Other	24.3*	26.4*	30.6	29.9	20.7	19.8
	(11.816)	(11.897)	(16.687)	(16.410)	(13.989)	(13.807)
Vocational	16.9***	17.2***	23.4**	23.6***	13.8**	14.2**
	(4.502)	(4.545)	(7.256)	(7.163)	(5.262)	(5.265)
O level	32.8***	34.1***	40.1***	38.9***	28.4***	27.9***
	(4.250)	(4.246)	(6.485)	(6.465)	(5.286)	(5.235)
A level	51.3***	51.5***	56.9***	57.3***	44.1***	44.5***
	(5.533)	(5.477)	(8.135)	(8.153)	(7.180)	(7.320)
Nurse	54.5***	56.5***	46.3**	43.1**	47.6***	46.3***
	(9.536)	(9.381)	(14.076)	(14.272)	(10.854)	(11.129)
Teacher	70.3***	69.1***	74.6***	75.4***	59.1***	60.1***
	(9.115)	(8.994)	(12.006)	(12.071)	(9.470)	(9.563)
Higher degree	85.6***	84.5***	90.7***	91.1***	73.1***	74.0***
	(5.288)	(5.250)	(7.434)	(7.444)	(6.000)	(6.030)
Constant	-55.3***	-52.1***	-44.4***	-48.0***	-49.6***	-52.3***
	(8.428)	(8.517)	(13.288)	(13.142)	(9.940)	(9.885)
N	4,149	4,149	1,926	1,926	4,149	4,149

*Notes. Standard errors in parentheses * p < 0.05, ** p < 0.01, *** p < 0.001.* ***Model 1****: complete cases dataset for 'benchmarking' with IPWs applied.* ***Model 2****: complete cases dataset but no IPW adjustment.* ***Model 3****: listwise deleted dataset after missing data pattern imposed.* ***Model 4****: as above with an IPW adjustment.* ***Model 5****: dataset with missing data pattern imposed with MI adjustment.* ***Model 6****: as above together with an IPW adjustment.*

The findings show that models 1 and 2 generate almost identical results even though model 2 does not adjust for unit non-response. Both models use the sample with complete cases (C) and do not suffer from item missingness. This similarity between the two models is a first indication that non-response weights do not improve the parameter estimates from regression analysis by much. That may quite possibly be because the weights have been constructed using a selection of birth characteristics which have very low predictive power (see pseudo R-squared in table 3).

In terms of standard errors, model 1 generates the lowest standard errors on all estimates, as we expected. Comparing model 1 with model 2, the use of IPWs generates slightly higher standard errors for the latter, as is to be expected. Models 5 and 6, which both involve MI with and without the application of IPWs, generate the closest results to model 1. This is the case for all variables without exception. Notably, the standard errors on the estimates in model 6 are almost identical to those in model 5. In contrast models 3 and 4, based upon listwise deletion with and without weight adjustment, generate parameter estimates that deviate substantially from model 1, together with larger standard errors (as a result of a halving of the sample size reported for models 5 and 6) indicating that they are least reliable.

For the parameter estimates, the picture is mixed. Models 1 and 2 provide very similar estimates since neither suffers from item missingness. Model 2 slightly deviates from model 1 because it does not adjust for unit non-response. Models 5 and 6 generate the closest findings to model 1 on gender and most of the modalities of income. However, for parental highest qualification, models 3 and 4 generate closer results to model 1 on 3 out of 7 modalities. Hence, one can say that MI appears to bring the estimates closer to their original values, but with some exceptions.

Discussion

Our illustration used unit non-response weights to adjust for lost cases and multiple imputations to adjust for any item missingness in our substantive model. These imputations adopt an imputation model based on auxiliary birth characteristics. The efficacy of both our adjustment methodologies depend on a number of conditions. The efficacy of the application of IPWs will depend on the predictive power of the response models used in their construction. In our case, the predictive power of the models was weak. Further, the use of wave specific weights is also undermined when variables from previous waves are used in our regression analysis (as in the case of parental income). This would therefore be an instance where the application of longitudinal weights might improve the adjustments. The efficacy of reweighting could also be improved by considering other variables that would better predict unit non-response, in particular metadata which is not currently collected in BCS70 waves.

Considering the impact of item missingness, our findings show that the application of MI improves the precision (standard errors) on all variable estimates, without exception. This happens because imputations increase sample size to its former level and therefore improve accuracy. Further, the analysis of imputed data generated parameter estimates that were closest to our benchmark model (1) with the exception of parental highest educational qualification. Ideally, this would require further investigation perhaps by repeating the generation of our chosen levels of missingness across many analyses (i.e. a simulation study) to assess the stability of these findings. The levels of item missingness were introduced on income and parental qualifications based on father's social class (manual vs. non-manual) and thought to be quite substantial at a 30 per cent difference in loss between the non-manual and manual social class groupings. In this manner, we would expect, as evidenced, that imputation would outperform any estimation based on listwise deletion. In contrast, if the difference in terms of the levels of missingness between the two categories had been very small then the researcher can assume that item missingness is almost MCAR and that carrying out MI prior to analysis will not make much difference to the results. By and large, the efficacy of imputations will depend on the extent of item missingness, how much it appears to deviate from MCAR, and whether or not the researcher is able to anticipate the missingness mechanism under appropriate assumptions (MAR or MNAR). The application of IPW and MI either in combination or as stand-alone adjustments are based on MAR. Under MI, it is important that the imputation procedure is appropriate for the type of the variables to be imputed (e.g. linear procedures for

continuous variables, ordinal and multinomial logit procedures for ordinal and multinomial variables, multilevel models for nested observations), in order to generate robust and valid inference. However, it is also advisable to conduct sensitivity analyses under MNAR as described by Carpenter and Kenward (2013). Instances of such checks are all too rare in the literature. Finally, for further information about missing data that may include interaction or non-linear terms, see Carpenter and Plewis (2011) and Goldstein et al., (2014) as these issues have not been covered in our illustration.

Conclusion

In this paper, we examine the extent of non-response in the British Cohort Study begun in 1970 (BCS70) and its effect on sample composition over the nine available data waves (1970 to 2012). The findings are based on BCS70, but their relevance extends to a wide range of birth cohort studies such as the National Child Development Study of 1958, the 1946 MRC National Survey of Health and Development, and the Millennium Cohort Study (for further information see:
http://www.closer.ac.uk/data-resources/explore
(Cohort & Longitudinal Studies Enhancement Resources, 2014).

We analyse the determinants of non-response using binary logistic regression models with selected birth characteristics as explanatory variables. We find that men from lower social backgrounds and with less educated parents are less likely to respond. However, despite the statistical significance of the estimated regression coefficients the predictive power of the models is weak. In the second section of the paper, we develop a hypothetical substantive model in order to illustrate the impact of the use non-response weights and imputation techniques on our inference.

The construction of IPW is itself a challenge. In our illustration there is an intuitive approach drawing upon CM's birth characteristics simply because these characteristics are available for almost all of the cohort and loss can be traced over time. Using these items to provide weights does not actually improve the efficacy of our models, largely because they have weak predictive power. The judicious inclusion of further items requires a better-developed theory of response co-operation and access to metadata. What is clearly attractive about MI is that it enables the researcher to restore the sample size to include cases with partial information. In longitudinal analysis this implies that it is possible to regard any wave-specific non-response as a set of missing items in a longitudinal record spanning the life of the cohort. In our illustration we were able to demonstrate that model estimates based on MI came closest to our benchmark model. However, not all estimates were in close agreement.

As the life of a cohort continues it will be increasingly important for analysts to make best use of the available data, which implies that it will be unwise to ignore cases over time with partial information arising from attrition, wave-specific dropout and item non-response. In order to exploit the availability of powerful software tools and the range of approaches now available under MAR and MNAR, assumptions will require that sufficient time and effort is made to understand and model missingness mechanisms as an integral part of the research process.

Acknowledgements

The authors would like to thank Professor Harvey Goldstein in his capacity as the journal's Section Editor, the anonymous referees and Dr George Ploubidis (Centre for Longitudinal Studies) for their helpful and constructive comments on an earlier draft of this paper.

References

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumaker (Eds.), *Advanced structural equation modelling: issues and techniques* (pp. 243–277). Mahwah, NJ: Lawrence Erlbaum Associates.

Carpenter, J. R., & Plewis, I. (2011). Analysing longitudinal studies with non-response: issues and statistical methods. In M. Williams & W. P. Vogt (Eds.) *The SAGE Handbook of innovation in social research methods* (pp.498-523). London: SAGE Publications Ltd. http://dx.doi.org/10.4135/9781446268261.n28

Carpenter, J. R. & Kenward, M. G. (2013). *Multiple imputation and its application*. Chichester: John Wiley & Sons Ltd. http://dx.doi.org/10.1002/9781119942283

Centre for Longitudinal Studies (2014). Technical Report of the 1970 British Cohort Study: Age 42 Survey 2012-2013). TNS British Market Research Bureau Ltd.: London

Cohort and Longitudinal Studies Enhancement Resources (CLOSER). (2014). URL:http://www.closer.ac.uk/data-resources/explore/ accessed 15.12 on 23.10.2014.

Davey, A. & Savla, J. (2010). *Statistical power analysis with missing data: a structural equation modelling approach.* New York: Taylor Francis Group.

Enders, C. (2008). A note on the use of missing auxiliary variables in Full Information Maximum Likelihood-based Structural Equation Models. *Structural Equation Modelling: A Multidisciplinary Journal*, *15*(3), 434-448. http://dx.doi.org/10.1080/10705510802154307

Gilks, W., Richardson, S., & Spiegelhalter, D. (1996). *Markov chain Monte Carlo in practice*. London: Chapman and Hall.

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

Goldstein, H., Carpenter, J. R., & Browne, W. J. (2014). Fitting multilevel and multivariate models with missing data in responses and covariates that may include interactions and non-linear terms. *Journal of the Royal Statistical Society A,* 177, Part II, 553-564. http://dx.doi.org/10.1111/rssa.12022

Hawkes, D., & Plewis, I. (2006). Modelling non-response in the National Child Development Study. *Journal of the Royal Statistical Society A, 169*(3), 479-492. http://dx.doi.org/10.1111/j.1467-985X.2006.00401.x

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 149-74. http://dx.doi.org/10.2307/1912352

Kalton, G. (1986). Handling wave nonresponse in panel surveys. *Journal of Official Statistics, 2*(3), 303-314.

Lepkowski, J. M. (1989). Treatment of wave nonresponse in panel surveys. In D. Kasprzyk, G. Duncan, G. Kalton and. M. P. Singh (Eds.) *Panel Surveys* (pp. 348-374). New York: Wiley.

Little, R., & Rubin, D. (2002). *Statistical analysis with missing data,* 2nd ed. Hoboken, N.J: Wiley. http://dx.doi.org/10.1002/9781119013563

McDonald, J. W. and Ketende, S. Nonresponse weight adjustments using multiple imputation for the UK Millennium Cohort Study. Statistics Canada International Symposium Series - Proceedings, Symposium 2009 : Longitudinal Surveys: from Design to Analysis, 27-30 October, Gatineau, Canada.

McDonald, J. W., & Ketende, S. C. (2009). Nonresponse weight adjustment using multiple imputation for the UK Millennium Cohort Study. *Proceedings of Statistics Canada Symposium: Longitudinal Surveys from Design to Analysis*, 27-30 October, Ottawa: Statistics Canada.

Micklewright, J., Skinner, C., & Schnepf, S. (2012). Non-response biases in surveys of school children: the case of the English PISA samples. *Journal of the Royal Statistical Society A*, *175*(4), 915-938. http://dx.doi.org/10.1111/j.1467-985X.2012.01036.x

Nathan, G. (1983). A simulation comparison of estimators for a regression coefficient under differential non-response. *Communications in Statistics: Theory and Methods, 12*(6) 645-659.

Nathan, G., & Holt, D. (1980). The effect of survey design on regression analysis. *Journal of the Royal Statistical Society B*, *42*(3), 377-386. http://dx.doi.org/10.1080/03610928308828485

Nevalainen, J., Kenward, M. G., & Virtanen, S. M. (2009). Missing values in longitudinal dietary data: a multiple imputation approach based on fully conditional specification. *Statistics in Medicine 28*(29), 3657-3669. http://dx.doi.org/10.1002/sim.3731

Parsons, S. (2014). British Cohort Study Data Note. Centre for Longitudinal Studies, University of London.

Pfeffermann, D., Moura, F., & Silva, P. (2001). *Multilevel modelling under informative probability sampling*. S3RI Methodology Working Paper M04/09. University of Southampton.

Plewis, I. (2007a). Non-response in a birth cohort study: the case of the Millennium Cohort Study. *International Journal of Social Research Methodology*, *10*(5), 325-334. http://dx.doi.org/10.1080/13645570701676955

Plewis, I. (2007b). *The Millennium Cohort Study: technical report on sampling.* Centre for Longitudinal Studies, the Institute of Education, University of London.

Plewis, I., Calderwood, L., & Ketende, S. (2009). Sample loss from cohort studies: patterns, characteristics and adjustments. *Proceedings of Statistics Canada Symposium: Longitudinal Surveys: from Design to Analysis.* 27-30 October, Ottawa: Statistics Canada.

Plewis, I. (2011). *Auxiliary variables and adjustments for missingness in longitudinal studies*. Cathie Marsh Institute for Social Research, CCSR working paper No. 2011-04.
Royston, P. (2009). Multiple imputation of missing values: further update on ICE, with an emphasis on categorical variables. *The Stata Journal*, *9*(3), 466-477.
Royston, P., & White, I. R. (2011). Multiple imputation by chained equations (MICE): implementation in Stata. *Journal of Statistical Software, 45*(4), 1-20.
Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(6), 581-92. http://dx.doi.org/10.1093/biomet/63.3.581
Rubin, D. B. (1987) *Multiple imputation for non-response in surveys*. New York: J. Wiley & Sons. http://dx.doi.org/10.1002/9780470316696
Rubin, D. B. (2004). The design of a general and flexible system for handling non-response in sample surveys. *American Statistician, 58*(4), 298-302. http://dx.doi.org/10.1198/000313004X6355
Seaman, S. R., & White, I. R. (2011). Review of inverse probability weighting for dealing with missing data. *Statistical Methods in Medical Research*, *22*(3) 278-295. http://dx.doi.org/10.1177/0962280210395740
Seaman, S. R., White, I. R., Copas, A. J. & Li, L. (2012). Combining multiple imputation and inverse-probability weighting. *Biometrics, 68*(1), 129-137. http://dx.doi.org/10.1111/j.1541-0420.2011.01666.x
Schafer, J., & Olsen, M. (1997). Multiple imputation for multivariate missing-data problems: a data analyst's perspective. *Multivariate Behavioural Research, 33*(4) 545-571. http://dx.doi.org/10.1207/s15327906mbr3304_5
Schouten, B., & de Nooij, G. (2005). *Non-response adjustment using classification trees.* Voorburg: Statistics Netherlands.
Taylor, M., Brice, J., Buck, N., & Prentice-Lane, E. (2010). *British Household Panel Survey User Manual, Volume A: Introduction, technical report and appendices.* Colchester: University of Essex.
Welch, C., Bartlett, J., & Petersen, I. (2014). Application of multiple imputation using the two-fold fully conditional specification algorithm in longitudinal clinical data. *The Stata Journal*, *14*(2), 418-431.
Wiggins, R. D., Schofield, P., Bartley, M., Sacker, A, & Head, J. (2004). Social determinants of changes in minor psychiatric morbidity over time in the British Household Panel Study 1991-1998. *Journal of Epidemiology and Community Health*, *58*, 779-791.
http://dx.doi.org/10.1136/jech.2003.015958

Endnotes

[i] Note that the number of observations in wave 1 (15,270) and wave 2 (10,059) deviate from those in table 2 because some of the CMs included in the category 'participated' have missing birth characteristics. Hence, the observations included in the computation of descriptive statistics and in the logit models are those with non-missing birth characteristics.

[ii] Note that we repeated the same analysis with 100 imputations. The difference in magnitude of the results was very limited which suggests that 20 imputations are enough to generate valid inference.

Additive influences of maternal and paternal body mass index on weight status trajectories from childhood to mid-adulthood in the 1970 British Cohort Study

Silvia Costa Institute of Child Health, University College London, UK
silvia.costa@ucl.ac.uk
William Johnson MRC Human Nutrition Research Unit, Cambridge, UK
Russell M Viner Institute of Child Health, University College London, UK

(Received July 2014 Revised February 2015) http://dx.doi.org/10.14301/llcs.v6i2.301

Abstract

This study aimed to (i) describe the weight status trajectories from childhood to mid-adulthood and (ii) investigate the influence of maternal and paternal body mass index (BMI) on offspring's trajectories in a nationally representative study in Great Britain. The sample comprised 4,174 (43% male) participants from the 1970 British Cohort Study with complete BMI data at ages 10, 26, 30, 34, and 42 years. Individuals' weight status was categorised as overweight/obese or non-overweight/obese at each age, and trajectories of weight status from 10 to 42 years of age were assessed. Sex-stratified multinomial logistic regression models were used to assess associations of maternal and paternal BMI with trajectory group membership, adjusting for potential confounders (e.g. socioeconomic position and puberty). Thirty per cent of individuals were never overweight/obese (reference trajectory), 6%, 44% and 8% had childhood, early- and mid-adulthood onset of overweight/obesity (respectively), and 12% other trajectories. In fully adjusted models, higher maternal and paternal BMI significantly increased the risk of childhood (relative risk ratio: 1.2-1.3) and early adulthood onset (1.2) of overweight/obesity in both sexes. Relative risk ratios were generally higher for maternal than paternal BMI in females but similar in males. Early puberty also increased the risk of childhood (1.8-9.2) and early-adulthood onset (3.7-4.7) of overweight/obesity. Results highlight the importance of primary prevention, as most individuals remained overweight/obese after onset. Maternal and paternal BMI had additive effects on offspring weight status trajectories across 32 years of the life course, suggesting that prevention/intervention programmes should focus on the whole family.

Keywords: Obesity; trajectories; tracking; life course; parental BMI; socioeconomic factors

Introduction

The obesity epidemic is one of the greatest public health concerns worldwide (World Health Organization, 2011; Health and Social Care Information Centre, 2014a). The increase in the prevalence of obesity has been seen in both adults and children (World Health Organization, 2011). In England the prevalence of overweight and obesity in adults has increased from 58% and 49% in 1993 to 67% and 57% in men and women respectively in 2012 (Health and Social Care information Centre, 2014a). Overweight and obesity in children has also increased, with peaking rates of 18% and 19% of 2-15 year-old boys and girls (respectively) in the 2004 Health Survey for England (Health and Social Care Information Centre, 2014a). Although rates in the Health Survey for England seem to be stabilising in recent years (Health and Social Care Information Centre, 2014a), other data sources show different rates and trends of childhood overweight and

obesity. For example, the latest sweep of the National Childhood Measurement Programme has reported that over a fifth of children in reception year (5-6 years) and about a third of those in year six (10-11 years) were overweight or obese, and that these rates were higher than those reported for the previous school year (Health and Social Care Information Centre, 2014b). This is worrying since obesity has been shown to track from childhood through to adulthood (Reilly et al., 2003; Venn et al., 2007; Brisbois, Farmer, & McCargar, 2012), and both childhood and adulthood obesity are associated with adverse health outcomes (Park, Falconer, Viner, & Kinra, 2012; Park, Sovio, Viner, Hardy, & Kinra, 2013; Prospective Studies Collaboration, 2009; Reilly et al, 2003).

Parental body mass index (BMI) is one of the most consistent factors associated with offspring obesity (Monasta et al., 2010; Parsons, Power, Logan, & Summerbell, 1999). For example, using data from the 1958 British birth cohort, Power, Manor and Matthews (2003) found a significant association of higher maternal and paternal BMI with obesity at 33 years, even when controlling for adulthood (in males) and childhood socio-economic position (SEP), and participants' education at 33 years. However, most research to date has only considered obesity at one time-point in adulthood (Parsons et al., 1999), and the influence of parental BMI on different trajectories of weight status (e.g. persistent overweight/obesity from childhood to adulthood) remains largely unknown. This consistent association of higher parental BMI with offspring overweight/obesity has been hypothesised to operate through a variety of different genetic (e.g. heritance of "obesity genes") (Maes, Neale, & Eaves, 1997), biological (e.g. restricted or excessive intrauterine growth), and shared environmental factors (e.g. inadequate diet and sedentary lifestyles) (Monasta et al., 2010; Parsons, et al., 1999). Few studies have investigated the association of maternal and paternal BMI with offspring overweight and obesity later than early-adulthood (Parsons, et al., 1999), and much of the existing literature suffers from potential confounding effects that were not accounted for, such as SEP and early maturation (Monasta et al., 2010).

Identifying the exact mechanisms through which parental obesity leads to later obesity is difficult, because of the many possible mediators and confounders involved in this relationship (Monasta et al., 2010; Parsons et al., 1999). It is important to clarify whether parental BMI has a significant influence in their offspring's lifelong weight status trajectories, given the high prevalence of adult overweight/obesity (World Health Organization, 2011; Health and Social Care Information Centre, 2014a) and that it is potentially easier to directly intervene in parents than in their offspring (e.g. through change in diet and physical activity) (Golan, 2006; Monasta et al, 2010). Studying large samples where basic measures of the risk factors have been collected has been suggested as a potential approach to deal with the inherent difficulty in measuring the risk factors associated with overweight/obesity (Parsons et al., 1999). The 1970 British Cohort Study (BCS70) has recently released the data for the 42-year follow-up; this presents a remarkable opportunity to investigate the influence of parental BMI on offspring's weight status trajectories from childhood to mid-adulthood, addressing some of the main limitations of previous studies – long-term follow-up (birth to 42 years) and data available for several confounders (e.g. SEP, breast-feeding, and maturation).

The aims of this study were to (i) describe weight status trajectories from childhood to mid-adulthood in the BCS70 and (ii) investigate the separate influences of maternal and paternal BMI on offspring's weight status trajectories.

Methods

Sample

The BCS70 has followed just under 17,200 people born in Great Britain (England, Scotland, and Wales) in one week in April 1970, from birth onwards (Centre for Longitudinal Studies, 2014). Nine main data collections have taken place: at birth (1970), 5 (1975), 10 (1980), 16 (1986), 26 (1996), 30 (2000), 34 (2004), 38 (2008), and 42 (2012) years of age. At the most recent sweep 57% (N = 9,842) of the participants were still in the study and they remained broadly representative of the national population of men and women of the same age. Participants were eligible for this study if they provided valid BMI data at age 42 years (n=8,799), and were included if they also had valid BMI data at 10, 26 or 30, 34, and 42 years of age. The final

sample for the present study comprised 4,174 (1,782 male; 2,392 female) participants, representing 47% of the 8,799 eligible participants. This sample was not noticeably different from the remaining 4,625 (53%) eligible participants who did not meet the inclusion criteria for the present study. There were no significant differences in parental BMI, and only small differences between included and excluded participants in some of the other variables: breastfeeding ≥3 months (87% of included versus 90% of excluded); early puberty (17% of included versus 15% of excluded); high SEP (34% of included versus 30% of excluded) (all $p<0.001$).

Participants' body mass index

Weight and height were measured by community medical officers, health visitors or school nurses at 10 and 16 years of age according to standard protocols, and self-reported in questionnaires at 26 years and face-to-face interviews at 30, 34, and 42 years of age. Due to the low response rate, the amount of data collected at age 16 was considerably less than at all other ages (e.g. valid BMI: 5,723 versus ≥7,303 participants). This was due to prolonged industrial action by teachers in 1986, who were responsible for the educational tests; the delayed survey start meant that many children had left school before assessments could be done and information for many of them could not be obtained. Therefore, data from the 16-year follow-up was excluded from this study. Body mass index (weight (kg)/height $(m)^2$) was computed at ages 10, 26, 30, 34, and 42 years. Due to the small number of obese participants at each age, weight status was categorised as either "1 = overweight/obese" or "0 = non-overweight/obese". In adulthood, participants' BMI was classified as overweight/obese when ≥25 kg/m^2, and non-overweight/obese if <25 kg/m^2. At 10 years of age, the International Obesity Task Force BMI cut-off points for overweight (Cole, Bellizzi, Flegal, & Dietz, 2000) were used to classify cohort members as overweight/obese or non-overweight/obese. To maximise sample size, the 26 and 30 year surveys were collapsed into one time-point, as the 26 year survey had substantially less valid BMI data and these two time points were closely spaced. Participants who were overweight/ obese at either age or both ages were coded as overweight/obese, and participants who were normal weight at both ages were coded as normal weight.

Parental body mass index and confounding variables

Maternal and paternal weight (kg or stones and pounds) and height (meters or feet and inches) were reported by the mother when cohort member was 10 years of age. Data in stones/pounds and feet/inches were transformed into kg and meters respectively, and parental BMI was subsequently calculated and used as a continuous variable in all analyses. Birth weight (in grams) and gestational age at birth (in completed weeks) was recorded by the midwife who had undertaken the delivery or the senior midwife. As there is good evidence of a positive relation between high birth weight and higher fatness levels (Parsons et al., 1999), birth weight was classified as "high birth weight" if >4kg, or "medium-low birth weight" if ≤4kg. Duration of breast-feeding was reported by the mother as "never breastfed", "<1 month", "≥1 month but <3 months", and "≥3 months", when participants were aged 5 years. A binary breast-feeding variable with the responses "<3 months" or "≥3 months" was computed. Evidence of puberty at age 10 years was assessed and recorded during the medical examination by a community medical officer, a health visitor or the school nurse, by answering "yes"/"no" to the question "*Did your examination reveal* (…) *Any evidence of puberty?*". Maternal and paternal occupation was reported by the mother at the 10 year follow-up, and classified according to the 1980 national Classification of Occupations (Office of Population Censuses and Surveys, 1980). Because higher parental SEP and education have shown some evidence of being protective of offspring obesity (Monasta et al., 2010), SEP and maternal education variables were dichotomised as follows. A binary SEP variable with the responses "high SEP" (if the mother, father, or both belonged to a professional or managerial social class) or "medium-low SEP" (if both parents belonged to skilled, partly-skilled or unskilled social classes) was derived. Maternal education was also reported by the mother at the 10 year follow-up, and a binary variable comprising "A-levels or above" (A-levels are the requirement for university access) and "below A-levels" was computed.

Weight status trajectories

Weight status trajectories between 10 and 42 years of age were defined manually for males and females separately, using a three-step procedure. We first tabulated all possible combinations of the binary weight status variable across the four ages to produce 16 trajectories (i.e. 2^4) and examined the frequency of participants belonging to each trajectory (see appendix 1). Second, all trajectories comprising <5% of the sample were collapsed into one group named "other trajectories", and those comprising ≥5% of the sample were retained as individual trajectories, namely: "never overweight/obese", comprising individuals who were never classified as overweight/obese between 10 and 42 years; "mid-adulthood onset of overweight/obesity", comprising individuals who were overweight/obese only at age 42 years; "overweight/obesity onset at age 34 years", comprising individuals who were overweight/obese from 34 years onwards; "overweight/obesity onset at ages 26 or 30 years", comprising individuals who were overweight/obese from 26-30 years onwards; and "childhood onset of overweight/obesity", comprising individuals who were always overweight/obese between 10 and 42 years of age. Third, to improve power and because both trajectories represented two timings of early adulthood, the trajectories "overweight/obesity onset at age 34 years" and "overweight/obesity onset at ages 26 or 30 years" were collapsed into a single trajectory re-labelled as "early-adulthood onset of overweight/obesity". The final outcome variable consisted of five trajectories: 1) "never overweight/obese", 2) "mid-adulthood onset of overweight/obesity", 3) "early-adulthood onset of overweight/obesity", 4) "childhood onset of overweight/obesity", and 5) "other trajectories".

Statistical Analyses

Descriptive statistics were used to describe the sample and the trajectories of weight status. All continuous variables were tested for normality of distribution with the Shapiro-Francia test and found to be non-normally distributed. Mann-Whitney U tests were used to assess differences in continuous and ordinal variables between sexes, and between those included and excluded from analyses. Chi-square tests were used to assess differences in categorical variables between sexes, and between those included and excluded from analyses.

Analyses were conducted separately for males and females. This was to test whether parental BMI affects weight status trajectories differently in males and females, and because of the sex inequality in the consequences of overweight /obesity, the potential differences in the association of early life factors (such as SEP) with adulthood overweight/obesity (Heraclides, Witte, & Brunner, 2008; Khlat, Jusot & Ville, 2009; Pudrovska, Logan & Richman, 2014), and the significant differences in the prevalence of weight status trajectories. Univariable multinomial logistic regression was used to assess the associations of maternal and paternal BMI (separately) with weight status trajectory membership. Multivariable multinomial logistic regression was then used to test how these associations change after adjusting for the confounding effect of birth factors (step 1) and puberty (step 2), maternal education and SEP (step 3). This was firstly done for maternal and paternal BMI separately, and subsequently with maternal and paternal BMI in the same model. Interactions between maternal and paternal BMI were also tested. To assess whether the results on the associations of parental BMI with offspring's weight status trajectories may be influenced by the different distributions of maternal and paternal BMI values, post-hoc sensitivity analyses were undertaken using parental BMI z-score (zBMI) instead of absolute BMI. Alpha-value was set at 0.05, with Bonferroni correction applied to multiple comparisons. All analyses were conducted in Stata (v.12) statistical software package.

Results

Details of the 4,174 participants included for analyses can be seen below in Table 1. More females (57.3%) than males were included for analyses. Significantly more females showed evidence of puberty at 10 years than males (p<0.001), whereas males had significantly higher birth weight and BMI at 42 years of age than females (p<0.001). All other variables were broadly similar for both sexes.

Table 1. Description of study sample

	Missing data [N (%[1])]		Total (N = 4,174)	Females (N = 2,392)	Males (N = 1,782)	*p* (sex difference)
BMI at 10 years (kg/m^2)	0 (0.0)	Median (IQR)	16.7 (15.6-18.1)	16.7 (15.5-18.3)	16.7 (15.7-18.1)	0.605
BMI at 42 years (kg/m^2)	0 (0.0)	Median (IQR)	26.6 (23.6-30.1)	25.2 (22.6-29.4)	27.9 (25.5-30.7)	<0.001
Gestational age (weeks)	864 (20.7)	Median (IQR)	40 (39-41)	40 (39-41)	40 (39-41)	0.016
Birth weight (kg)	241 (5.8)	Median (IQR)	3.3 (3.0-3.6)	3.3 (3.0-3.6)	3.4 (3.1-3.7)	<0.001
Medium-low birth weight		N (valid %[2])	3617 (92.0)	134 (5.9)	182 (10.9)	
High birth weight		N (valid %[2])	316 (8.0)	2130 (94.1)	1487 (89.1)	
Breast-feeding duration	569 (13.6)					0.466
<3 months		N (valid %[2])	3,127 (86.7)	1,796 (86.4)	1,331 (87.2)	
≥3 months		N (valid %[2])	478 (13.3)	283 (13.6)	195 (12.8)	
Evidence of puberty at 10 years	18 (0.4)					<0.001
No		N (valid %[2])	3,432 (82.6)	1,735 (72.8)	1,697 (95.7)	
Yes		N (valid %[2])	724 (17.4)	647 (27.2)	77 (4.3)	
Maternal BMI (kg/m^2)	116 (2.8)	Median (IQR)	22.6 (21.0-25.0)	22.5 (21.0-24.9)	22.7 (21.1-25.0)	0.352
Paternal BMI (kg/m^2)	274 (6.6)	Median (IQR)	24.1 (22.5-26.0)	24.1 (22.5-26.0)	24.1 (22.5-26.3)	0.437
Maternal education at 10 years	251 (6.0)					0.271
Less than A-level		N (valid %[2])	3,184 (81.2)	1,825 (80.6)	1,359 (82.0)	
A-level or above		N (valid %[2])	739 (18.8)	440 (19.4)	299 (18.0)	
Parental SEP at 10 years	114 (2.7)					0.453
Medium-low		N (valid %[2])	2,711 (66.8)	1,542 (66.3)	1,169 (67.4)	
High		N (valid %[2])	1,349 (33.2)	784 (33.7)	565 (32.6)	

Legend: BMI - Body mass index; IQR - Inter-quartile range; SEP - Socio-economic position.

[1] percent of total (4,174) participants with trajectories of weight status.

[2] percent of total (4,174) participants with valid data for each variable.

Trajectories of weight status

Significantly more females than males were never overweight/obese and had mid-adulthood onset of overweight/obesity (p<0.001), whereas the prevalence of males having early-adulthood onset of overweight/obesity was roughly double the prevalence of females (62.7% versus 30.6%; p<0.001). Prevalence of childhood onset of overweight/obesity was the same in females and males (see table 2). Importantly, of those overweight/obese in childhood (229 females; 129 males), the wide majority remained overweight/obese in at least two time points during adulthood – 85% of females and 92% of males (see table in appendix 1).

Table 2. Description of trajectories of weight status included for regression analyses [n (%)].

	Total (n=4,174)	Females (n=2,392)	Males (n=1,782)
Never OW/OB	1,230 (29.5%)	962 (40.2%)	268 (15.0%)*
Mid-adulthood OW/OB onset	334 (8.0%)	231 (9.6%)	103 (5.8%)*
Early adulthood OW/OB onset	1,850 (44.3%)	733 (30.6%)	1117 (62.7%)*
Childhood OW/OB onset	258 (6.2%)	148 (6.2%)	110 (6.2%)
Other trajectories	502 (12.0%)	318 (13.3%)	184 (10.3%)

Legend: OW/OB – Overweight or obese; *p<0.001 for difference between sexes.

The influence of parental BMI on trajectories of weight status

In females, both maternal and paternal BMI were significantly and positively associated with increased risk of childhood, early- and mid-adulthood onset of overweight/obesity, and having "other trajectories" versus never being overweight/obese in unadjusted regression models (table 3). In males, higher maternal BMI was associated with increased risk of having childhood and early-adulthood onset of overweight/obesity, and of belonging to "other trajectories" versus never being overweight/obese. Higher paternal BMI was associated with increased risk of having childhood and early-adulthood onset of overweight/obesity versus never being overweight/obese in unadjusted regression models (table 4). These associations remained significant for males and females even when considering both maternal and paternal BMI in the same model (unadjusted for other variables). Interaction between maternal and paternal BMI was not significant (results not shown) and, therefore, was not included in the subsequent multivariable regression models. Because results of the multivariable regression models were similar when analyses where conducted with maternal and paternal BMI separately and in the same model, only the results for the models including both maternal and paternal BMI are presented below (tables 3 and 4). Results for the multivariable regression models with maternal BMI only and paternal BMI separately can be seen in the tables presented in appendices 2-5.

Results from fully adjusted models

In females after adjustment for birth factors, puberty and SEP factors (table 3), higher maternal BMI remained significantly associated with higher risk of childhood, early- and mid-adulthood onset of overweight/obesity, and of belonging to "other

trajectories" compared with never being overweight/obese (relative risk ratio (RRR): 1.08-1.27; all $p \leq 0.017$). Higher paternal BMI significantly increased the risk of childhood and early-adulthood onset of overweight/obesity, and belonging to "other trajectories" compared with never being overweight/obese (RRR: 1.11-1.21; all $p<0.001$). High SEP was associated with lower risk of childhood and early-adulthood onset of overweight/obesity (RRR: 0.50 and 0.75 respectively; both $p<0.04$), whereas early puberty increased the risk of childhood (RRR: 3.68; 95% CI: 2.29-5.92) and early-adulthood onset of overweight/obesity (RRR: 1.75; 95% CI: 1.31-2.34), and of belonging to "other trajectories" (RRR: 2.30; 95% CI: 1.62-3.28) compared to never being overweight/obese (all $p<0.001$). No other significant associations were found.

Table 3. Associations between parental BMI and trajectories of weight status in females.

			Never OW/OB	Mid-adulthood onset of OW/OB		Early-adulthood onset of OW/OB		Childhood onset of OW/OB		Other trajectories	
Risk Factor	Included n (%)	Missing n (%)	*reference*	RRR (95% CI)	*p*	RRR (95% CI)	*p*	RRR (95% CI)	*p*	RRR (95% CI)	*p*
Unadjusted individual models											
Maternal BMI	2,339 (97.8)	53 (2.2)		1.10 (1.05-1.15)	<0.001	1.21 (1.17-1.25)	<0.001	1.31 (1.25-1.37)	<0.001	1.17 (1.13-1.22)	<0.001
Paternal BMI	2,250 (94.1)	142 (5.9)		1.09 (1.03-1.15)	0.002	1.17 (1.12-1.21)	<0.001	1.29 (1.21-1.36)	<0.001	1.15 (1.10-1.21)	<0.001
Mutually adjusted	2,230 (93.2)	162 (6.8)									
Maternal BMI				1.10 (10.4-1.15)	<0.001	1.20 (1.16-1.24)	<0.001	1.29 (1.23-1.35)	<0.001	1.16 (1.12-1.21)	<0.001
Paternal BMI				1.08 (1.02-1.14)	0.006	1.14 (1.10-1.18)	<0.001	1.23 (1.16-1.31)	<0.001	1.13 (1.07-1.18)	<0.001
Adjusted for birth factors, puberty and socioeconomic position	1,567 (65.5)	825 (34.5)									
Maternal BMI				1.08 (1.01-1.14)	0.017	1.17 (1.13-1.22)	<0.001	1.27 (1.20-1.34)	<0.001	1.17 (1.11-1.23)	<0.001
Paternal BMI				1.03 (0.97-1.11)	0.34	1.15 (1.10-1.20)	<0.001	1.21 (1.13-1.31)	<0.001	1.11 (1.05-1.18)	<0.001

Gestational age		0.93 (0.84-1.03)	0.18	0.99 (0.92-1.07)	0.84	0.92 (0.80-1.06)	0.26	0.96 (0.87-1.06)	0.42
Birth weight									
Medium-low birth weight	reference	-		-		-		-	
High birth weight		1.62 (0.83-3.16)	0.16	0.98 (0.57-1.67)	0.93	1.66 (0.73-3.77)	0.23	0.86 (0.42-1.76)	0.67
Breast feeding									
≥3 months	reference	-		-		-		-	
<3 months		1.02 (0.61-1.71)	0.93	1.18 (0.81-1.72)	0.38	1.33 (0.64-2.75)	0.44	0.99 (0.62-1.57)	0.96
Evidence of puberty at 10 years									
No	reference	-		-		-		-	
Yes		1.31 (0.85-2.01)	0.23	1.75 (1.31-2.34)	<0.001	3.68 (2.29-5.92)	<0.001	2.30 (1.62-3.28)	<0.001
Socioeconomic position									
Medium-low	reference	-		-		-		-	
High		0.85 (0.57-1.24)	0.40	0.75 (0.57-0.98)	0.036	0.50 (0.28-0.88)	0.016	0.90 (0.64-1.28)	0.57
Maternal Education									
Less than A-level	reference	-		-		-		-	
A-level or above		0.94 (0.59-1.53)	0.85	0.96 (0.68-1.34)	0.8	0.82 (0.40-1.67)	0.59	0.71 (0.45-1.13)	0.15

Legend: BMI – Body mass index; CI - Confidence interval; OW/OB - Overweight or obese; RRR - Relative risk ratio.

In the fully adjusted model for males (table 4), both higher maternal and paternal BMI significantly increased the risk of childhood and early-adulthood onset of overweight/obesity compared to never being overweight/obese (RRR: 1.15-1.22; all $p<0.001$). Early puberty also increased the risk of childhood (RRR: 9.14; 95% CI: 1.76-47.56) and early-adulthood onset of overweight/obesity (RRR: 4.81; 95% CI: 1.14-20.32) compared with never being overweight/obese (both $p<0.04$). Additionally, having a high (versus normal-low) birth weight significantly increased the risk of early-adulthood onset of overweight/obesity (RRR: 2.67; 95% CI: 1.33-5.34), whereas having a higher gestational age significantly decreased the risk of early-adulthood onset of overweight/obesity (RRR: 0.88; 95% CI: 0.79-0.98) in relation to never being overweight/obese (table 4). No association was found between trajectories of weight status and breastfeeding and maternal education in both sexes.

Table 4. Associations between parental BMI and trajectories of weight status in males

			Never OW/OB	Mid-adulthood onset of OW/OB		Early-adulthood onset of OW/OB		Childhood onset of OW/OB		Other trajectories	
Risk Factor	Included n (%)	Missing n (%)	*reference*	RRR (95% CI)	*p*	RRR (95% CI)	*p*	RRR (95% CI)	*p*	RRR (95% CI)	*p*
Unadjusted individual models											
Maternal BMI	1,719 (96.5)	63 (3.5)		1.05 (0.97-1.13)	0.26	1.16 (1.10-1.22)	<0.001	1.23 (1.16-1.31)	<0.001	1.09 (1.03-1.16)	<0.001
Paternal BMI	1,650 (92.6)	132 (7.4)		1.05 (0.96-1.15)	0.32	1.17 (1.10-1.23)	<0.001	1.27 (1.17-1.37)	<0.001	1.15 (0.99-1.15)	<0.001
Mutually adjusted	1,628 (91.4)	154 (8.6)									
Maternal BMI				1.02 (0.94-1.11)	0.56	1.13 (1.07-1.18)	<0.001	1.19 (1.12-1.27)	<0.001	1.07 (1.01-1.15)	<0.001
Paternal BMI				1.04 (0.94-1.14)	0.44	1.14 (1.08-1.21)	<0.001	1.23 (1.14-1.34)	<0.001	1.06 (0.98-1.15)	<0.001
Adjusted for birth factors, puberty and socioeconomic position	1,140 (64.0)	642 (36.0)									
Maternal BMI				1.01 (0.92-1.12)	0.8	1.15 (1.08-1.22)	<0.001	1.22 (1.12-1.33)	<0.001	1.08 (0.99-1.17)	<0.001
Paternal BMI				1.02 (0.91-1.15)	0.71	1.16 (1.08-1.24)	<0.001	1.22 (1.10-1.36)	<0.001	1.04 (0.94-1.14)	<0.001

Gestational age		0.98 (0.82-1.17)	0.84	0.88 (0.79-0.98)	0.016	0.89 (0.75-1.06)	0.20	0.16 (0.99-1.35)	0.06
Birth weight									
Medium-low birth weight	reference	-		-		-		-	
High birth weight		2.12 (0.79-5.71)	0.14	2.67 (1.33-5.34)	0.006	0.78 (0.20-3.01)	0.72	1.75 (0.73-4.19)	0.21
Breast feeding									
≥3 months	reference	-		-		-		-	
<3 months		0.66 (0.29-1.49)	0.32	1.80 (0.46-1.38)	0.42	1.33 (0.64-2.75)	0.44	0.99 (0.62-1.57)	0.24
Evidence of puberty at 10 years									
No	reference	-		-		-		-	
Yes		2.56 (0.35-18.64)	0.36	4.71 (1.11-19.95)	0.035	9.21 (1.77-48.02)	0.008	2.92 (0.52-16.36)	0.22
Socioeconomic position									
Medium-low	reference	-		-		-		-	
High		1.07 (0.58-1.95)	0.83	0.88 (0.60-1.28)	0.51	0.97 (0.50-1.87)	0.93	0.82 (0.49-1.38)	0.46
Maternal Education									
Less than A-level	reference	-		-		-		-	
A-level or above		0.63 (0.29-1.53)	0.23	0.78 (0.50-1.23)	0.29	0.86 (0.00-6.31)	0.72	0.75 (0.40-1.43)	0.39

Legend: BMI – Body mass index; CI - Confidence interval; OW/OB - Overweight or obese; RRR - Relative risk ratio.

In post-hoc multivariable models using parental zBMI (instead of absolute BMI), maternal zBMI consistently showed stronger associations with childhood and early-adulthood onset of overweight/obesity than paternal zBMI in both males (RRR: 1.71 and 2.14 (respectively) for maternal, versus 1.55 and 1.85 for paternal zBMI; all $p<0.001$) and females (RRR: 1.85 and 2.50 for maternal, versus 1.53 and 1.80 for paternal zBMI; all $p<0.001$)

Discussion

This study sought to (i) describe the trajectories of weight status from childhood to mid-adulthood, and (ii) investigate the influence of parental BMI on weight status trajectories of their offspring. In both sexes, approximately 6% had childhood onset of overweight/obesity, and the vast majority of those overweight/obese in childhood remained overweight/obese in at least two of the three adulthood time-points considered in this study. Maternal and paternal BMI showed additive effects on offspring weight status trajectories across 32 years of the life course, independently of several confounding factors identified in the literature (Monasta et al., 2010; Parsons et al., 1999).

Parental influence on weight status trajectories

In both males and females, maternal and paternal BMI were positively associated with increased risk of childhood and early-adulthood onset of overweight/obesity (versus never being overweight/obese). In females, maternal and paternal BMI were also positively associated with increased risk of "other trajectories". This means that higher parental BMI also increased the risk of females being overweight/obese at ≥1 time point between 10 and 42 years (apart from mid-adulthood onset of overweight/obesity). These results are consistent with the findings of the systematic review by Parsons et al., (1999), who reported that offspring of obese parents were consistently reported (in longitudinal studies) to be at risk of increased fatness from childhood into adulthood. In the fully adjusted models, these associations were slightly higher for maternal than paternal BMI in females (e.g. RRR for childhood onset of overweight/obesity: maternal BMI=1.27, paternal BMI=1.21), but similar for males (see tables 3 and 4). However, when repeating the analyses using parental zBMI, maternal zBMI consistently showed stronger associations than paternal zBMI with childhood and early adulthood onset of overweight/obesity than paternal zBMI in both sexes. This suggests that the same shift (one z-score) in the distribution of parental BMI will have a higher influence in the offspring's weight status trajectories for maternal than paternal BMI.

The influence of parental BMI on offspring's overweight/obesity trajectories may operate through a variety of factors, including genetic predisposition, restricted/excessive intrauterine growth, shared environment and the inheritance of unhealthy behaviours (Monasta et al., 2010; Parsons et al., 1999). One hypothesis put forward for a higher risk of overweight/obesity in the offspring of overweight mothers is that such mothers are more likely not to breast-feed their infants and to be of lower SEP (Monasta et al., 2010), which may indirectly influence their child's BMI through factors like diet and physical activity. Given that maternal BMI remained significant after controlling for these two variables, our data does not support this hypothesis. Our results also do not support a possible mediating role of excessive intrauterine growth and infant feeding for the association of maternal BMI with offspring overweight/obesity (Monasta et al., 2010), as higher maternal BMI remained significantly associated with childhood and early-adulthood onset of overweight/obesity after adjusting for those variables. Although we have controlled for maternal education and childhood SEP (known confounders) (Monasta et al., 2010), the extent to which the observed influence of parental BMI on offspring's overweight/obesity trajectories is due to genetic, environmental or behavioural factors (or a combination of the three) cannot be assessed in the current study and requires investigation in future studies. Nevertheless, because (i) parental BMI was strongly associated with persistent overweight/obesity, (ii) the population rates of overweight are high and (iii) it is potentially easier to intervene in parents than in children to reduce their overweight (Golan, 2006; Monasta et al., 2010), focusing on the family as a whole and targeting overweight parents in obesity interventions may be a successful strategy to reduce both adulthood overweight and

offspring's risk of lifetime overweight/obesity at the population level. Additionally, as the influence of parental BMI may be largely biological (Monasta et al., 2010), it may be necessary to intervene in parents (particularly the mother) during or even before pregnancy.

Other significant risk factors

Early puberty was also a significant predictor of childhood and early adulthood onset of overweight/obesity for both sexes, as well as belonging to "other trajectories" in females. These results are consistent with previous findings (Parsons et al., 1999; Prentice & Viner, 2013), and improve upon existing literature which in general has lacked adjustment for SEP and parental fatness (Parsons et al., 1999). Because the available BCS70 data did not allow us to investigate whether obesity before 10 years may have led to early puberty and in turn to the subsequent onset of overweight/obesity, the possibility of reverse causality cannot be rejected and requires investigation in future studies. Nevertheless, the fact that early puberty was consistently associated with "early adulthood onset of overweight/obesity" (which by definition controls for overweight in childhood) indicates that this association between early puberty and subsequent adulthood overweight/obesity may be independent of preceding increased fatness, in line with the results of a recent systematic review by Prentice and Viner (2013).

Only three other factors emerged as significant predictors of weight status trajectories in the fully adjusted model, with mixed results between males and females. In males, higher gestational age significantly reduced the risk of having an early adulthood onset of overweight/obesity, in relation to never being overweight/obese. Although significant, the slight difference in the distribution of gestational age between sexes (39.6 weeks for males, 39.8 weeks for females; p=0.02) is unlikely to have influenced the lack of the same association for females. On the other hand, a birth-weight ≥4 kg significantly increased the risk of males having early adulthood onset of overweight/obesity (versus never being overweight/obese), an association previously reported in the literature (Monasta et al., 2010; Parsons et al., 1999). The higher prevalence of high birth-weight in males than in females (11% versus 6%; p<0.001), and the very large sample size of males in the early adulthood onset of overweight/obesity trajectory, may have influenced to some extent the ability to detect an association in males but not in females. Although the general consensus is that there is a consistent positive relationship between birth-weight and later obesity (Monasta et al., 2010; Parsons et al., 1999) and plausible mechanisms have been proposed (Parsons et al., 1999), the nature and strength of this association is less clear in studies that account for potential confounders, as in the present study. Both similar and reverse sex-differences in this association have been previously reported (Parsons et al., 1999), and more research is needed to clarify whether this association of birth-weight with later obesity is independent of known confounders (e.g. gestational diabetes, parental fatness and SEP).

The association of high childhood SEP with lower risk for childhood and early-adulthood onset of overweight/obesity only in females cannot be explained by differences in prevalence of high childhood SEP between sexes (p=0.94). This sex-inequality has been observed in previous studies using adulthood obesity as the outcome even after adjustment for potential adulthood cofounders (Heraclides et al., 2008; Khlat et al., 2009; Pudrovska et al., 2014). One potential explanation is that traditionally men tend to move upward on the social ladder, resulting in a closing of the social gap in health outcomes from childhood SEP; whereas women tend to remain in the social class of origin, magnifying the effect of parental social class on adulthood overweight (Heraclides et al., 2008). However, a considerable amount of literature has reported this effect of early life SEP in both sexes (Monasta et al., 2010; Parsons et al., 1999; Power et al., 2003). For example, Power et al. (2003) reported a significant association of early life social class with obesity at age 30 years for both males and females from the British 1958 birth cohort study. Among other factors, differences in sample sizes, timing and type of the childhood SEP variable, the number of potential confounders adjusted for, and using weight status at only one time-point (versus lifelong trajectories) as the outcome could explain the contradictory results. More studies adopting a life course approach, like the present study, and controlling for the main known confounders/mediators (e.g. childhood weight status, early puberty and parental BMI) are needed,

to clarify whether a sex inequality exists for the SEP effect on later obesity or if it is a result of the selected data and methodology of different studies.

Limitations and strengths

This study had some limitations that deserve consideration. Only 47.4% of the eligible cohort members with valid BMI at 42 years were included for analysis and had weight status trajectories computed (due to participants missing BMI at ≥1 time-points), which could have impacted on the results (e.g. reducing statistical power to detect some associations). However, there was no significant difference in parental BMI, gestational age and prevalence of high birth-weight between those included and excluded from analyses. All other variables were significantly different between those with and without complete data to compute trajectories of weight status, but the observed differences were very small (e.g. median BMI at 10 years: 16.7 kg/m^2 in included versus 16.3 kg/m^2 in excluded participants; high SEP: 34% in included versus 30% in excluded). Furthermore, differences between those included and excluded from the fully adjusted models were only significant in SEP (males and females), and breastfeeding in females; all other variables were similar between included and excluded participants (p>0.07 for differences). There was a slightly higher proportion of participants with high SEP in those included versus excluded from the fully adjusted models (36% versus 26% in males; 36% versus 29% in females). As sample sizes for each independent variable varied widely due to missing data and the observed differences were generally minor in size, it is unlikely that these differences would have had an impact on the results and their generalizability to the wider BCS70 cohort. As data were drawn from a national birth cohort, the findings are highly generalizable to the British population relatively contemporaneous in age to those in the BCS70 cohort. It is unclear how relevant the findings are to more recent age cohorts, as populations born in more recent years have been shown to have higher mean BMI and higher rates of overweight/obesity (Clarke, O'Malley, Johnston, & Schulenberg, 2009; Stamatakis, Primatesta, Chinn, Rona, & Falascheti, 2005). However, we would expect the effects of parental BMI on offspring overweight/obesity to be stronger in younger generations than that observed in the BCS70 cohort.

Weight and height at 10 years were measured but adulthood values were reported by participants, which may have led to biased BMI values and consequent classification of overweight/obesity. However, the recent review by Brisbois et al., (2012) reported overall good agreement between the results of studies using measured and reported adult BMI. Taking this into account and the strong significant associations found in the present study, it is unlikely that the possible misreporting of weight and height would significantly impact on the results seen for the association of parental BMI with persistent overweight/obesity and early adulthood onset of overweight/obesity.

The fact that trajectories of weight status were manually computed and not drawn from the data using a statistical procedure is both a strength and a limitation of the current study. Using the "manual process" for classification of trajectories of weight status allows us know exactly how each trajectory is characterised, and not rely on regression model specifications to derive weight status trajectories. However, there may have been other potential trajectories that were missed by, for example, joining the trajectories with <5% prevalence in one only category (i.e. "other"). Using growth-mixture modelling (Muthén, 2001) with the five weight status trajectories identified in this study as the starting point, would (i) enable us to confirm whether the results from this study are confirmed when not using a manual procedure to derive the trajectories, (ii) include the 16-year sweep, as the model will be able to deal with the large amount of missing BMI data at this age, and (iii) address the issue of a reduced sample size in the fully adjusted regression models, by imputing the missing variables. We are currently undertaking this work. Other strengths of this study include investigating offspring weight status from childhood through to mid-adulthood in a large birth cohort sample, using a life course perspective (rather than only one time-point in adulthood), which addresses two main limitations of many previous studies looking the association of parental BMI with offspring obesity (Parsons et al., 1999; Monasta et al., 2010). A final limitation was the fact that we did not include behavioural variables (such as diet and physical activity habits) which can influence participants'

weight status at all ages, as well as their weight status trajectories.

Conclusions

In conclusion, higher paternal and maternal BMI showed additive influences on offspring weight status trajectories across 32 years of the life course after controlling for potential confounding factors identified in the literature (Parsons et al., 1999; Monasta et al., 2010). In particular, maternal and paternal BMI were significantly associated with higher risk of childhood and early-adulthood onset of overweight/obesity, and "other trajectories" (females only) versus never being overweight/ obese. Early puberty was also significantly associated with higher risk of childhood and early adulthood onset of overweight/obesity for both sexes. Other factors such as high SEP (females), gestational age and high birth-weight (males) also showed significant associations with childhood and early-adulthood onset of overweight/obesity, but these associations were weak and not consistent between sexes. The identified weight status trajectories highlight the importance of primary prevention, as most individuals remained overweight/obese after onset at any age and both childhood and adulthood obesity have been consistently associated with increased risks of type 2 diabetes, coronary heart disease and all-cause mortality later in life (Park et al., 2012; Park et al., 2013; Prospective Studies Collaboration, 2009). Early maturing overweight/obese children and those with overweight/obese parents are potential priority targets for interventions, due to their higher risk for persistent overweight/obesity through to mid-adulthood in both males and females. Prevention/intervention programmes should focus on the parents before pregnancy, or later on the family as a whole, since maternal and paternal BMI showed additive effects on offspring weight status trajectories from childhood to mid-adulthood.

Acknowledgements

The authors declare no competing interests. Permission for download and use of the BCS70 data was granted by the UK Data Service.

References

Brisbois, T. D., Farmer, A. P., & McCargar, L. J. (2012). Early markers of adult obesity: A review. *Obesity Reviews, 13*(4), 347-367. http://dx.doi.org/10.1111/j.1467-789X.2011.00965.x

Centre for Longitudinal Studies. (2014). History of the 1970 British Cohort Study [Internet]. Retrieved 06/05/2014, from: http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=944&sitesectiontitle=History+of+the+1970+British+Cohort+Study

Clarke, P., O'Malley, P. M., Johnston, L. D., & Schulenberg, J. E. (2009). Social disparities in BMI trajectories across adulthood by gender, race/ethnicity and lifetime socio-economic position: 1986-2004. *International Journal of Epidemiology, 38*(2), 499-509. http://dx.doi.org/10.1093/ije/dyn214

Cole, T. J., Bellizzi, M. C., Flegal, K. M., & Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: International survey. *British Medical Journal, 320*(7244), 1240-1243. http://dx.doi.org/10.1136/bmj.320.7244.1240

Golan, M. (2006). Parents as agents of change in childhood obesity - from research to practice. *International Journal of Pediatric Obesity, 1*(2), 66-76. http://dx.doi.org/10.1080/17477160600644272

Health and Social Care Information Centre. (2014a). *Statistics on obesity, physical activity and diet: England 2014* (pp. 1-101). London: Health and Social Care Information Centre.

Health and Social Care Information Centre. (2014b). *National child measurement programme: England, 2013/14 school year* (pp. 1-51). London: Health and Social Care Information Centre.

Heraclides, A., Witte, D., & Brunner, E. J. (2008). The association between father's social class and adult obesity is not explained by educational attainment and an unhealthy lifestyle in adulthood. *European Journal of Epidemiology, 23*(8), 573-579. http://dx.doi.org/10.1007/s10654-008-9245-3

Khlat, M., Jusot, F., & Ville, I. (2009). Social origins, early hardship and obesity: a strong association in women, but not in men? *Social Science & Medicine, 68*(9), 1692-1699. http://dx.doi.org/10.1016/j.socscimed.2009.02.024

Maes, H.H., Neale, M.C., Eaves, L.J. (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behavior Genetics*, *27*(4), 325-351. http://dx.doi.org/10.1023/A:1025635913927

Monasta, L., Batty, G. D., Cattaneo, A., Lutje, V., Ronfani, L., Van Lenthe, F. J., & Brug, J. (2010). Early-life determinants of overweight and obesity: a review of systematic reviews. *Obesity Reviews, 11*(10), 695-708. http://dx.doi.org/10.1111/j.1467-789X.2010.00735.x

Muthén, B. (2001). Latent variable mixture modelling. In G.A. Marcoulides, & R.E. Schumacker. (Eds.), *New developments and techniques in structural equation modelling* (pp. 1-33). Mahan, NJ: Lawrence Erlbaum Associates.

Office of Population Censuses and Surveys (1980). *Classification of occupations 1980*. London: Her Majesty's Stationery Office.

Park, M. H., Falconer, C., Viner, R. M., & Kinra, S. (2012). The impact of childhood obesity on morbidity and mortality in adulthood: a systematic review. *Obesity Reviews, 13*(11), 985-1000. http://dx.doi.org/10.1111/j.1467-789X.2012.01015.x

Park, M. H., Sovio, U., Viner, R. M., Hardy, R. J., & Kinra, S. (2013). Overweight in childhood, adolescence and adulthood and cardiovascular risk in later life: pooled analysis of three British birth cohorts. *PLoS ONE, 8*(7), e70684. http://dx.doi.org/10.1371/journal.pone.0070684

Parsons, T. J., Power, C., Logan, S., & Summerbell, C. D. (1999). Childhood predictors of adult obesity: A systematic review. *International Journal of Obesity and Related Metabolic Disorders, 23*(Suppl 8), S1-107.

Power, C., Manor, O., & Matthews, S. (2003). Child to adult socioeconomic conditions and obesity in a national cohort. *International Journal of Obesity, 27*(9), 1081-1086. http://dx.doi.org/10.1038/sj.ijo.0802323

Prentice, P., & Viner, R. M. (2013). Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. *International Journal of Obesity, 37*(8), 1036-1043. http://dx.doi.org/10.1038/ijo.2012.177

Prospective Studies Collaboration, Whitlock, G., Lewington, S., Sherliker, P., Clarke, R., Emberson, J., & Peto, R. (2009). Body-mass index and cause-specific mortality in 900,000 adults: collaborative analyses of 57 prospective studies. *Lancet, 373*(9669), 1083-1096. http://dx.doi.org/10.1016/S0140-6736(09)60318-4

Pudrovska, T., Logan, E. S., & Richman, A. (2014). Early-life social origins of later-life body weight: the role of socioeconomic status and health behaviors over the life course. *Social Science Research, 46*(2014), 59-71. http://dx.doi.org/10.1016/j.ssresearch.2014.02.007

Reilly, J. J., Methven, E., McDowell, Z. C., Hacking, B., Alexander, D., Stewart, L., & Kelnar, C. J. (2003). Health consequences of obesity. *Archives of Disease in Childhood, 88*(9), 748-752. http://dx.doi.org/10.1136/adc.88.9.748

Stamatakis, E., Primatesta, P., Chinn, S., Rona, R., & Falascheti, E. (2005). Overweight and obesity trends from 1974 to 2003 in English children: What is the role of socioeconomic factors? *Archives of Disease in Childhood, 90*(10), 999-1004. http://dx.doi.org/10.1136/adc.2004.068932

Venn, A. J., Thomson, R. J., Schmidt, M. D., Cleland, V. J., Curry, B. A., Gennat, H. C., & Dwyer, T. (2007). Overweight and obesity from childhood to adulthood: A follow-up of participants in the 1985 Australian Schools Health and Fitness Survey. *The Medical Journal of Australia, 186*(9), 458-460.

World Health Organization (2011) *Global status report on non-communicable diseases 2010*. Geneva: World Health Organization.

Appendix 1. Description and prevalence of possible overweight/obesity trajectories from childhood to mid-adulthood (overall and by sex)

	Age (years)				Overall	Females	Males
Trajectory	**10**	**26-30**	**34**	**42**	**N (%)**	**N (%)**	**N (%)**
1	No	No	No	No	1,230 (29.5)	962 (40.2)	268 (15.0)
2	No	No	No	OW/OB	334 (8.0)	231 (9.7)	103 (5.8)
3	No	No	OW/OB	OW/OB	233 (5.6)	145 (6.1)	88 (4.9)
4	No	OW/OB	OW/OB	OW/OB	1,617 (38.7)	588 (24.6)	1,029 (57.7)
5	OW/OB	No	No	No	27 (0.7%)	21 (0.9)	6 (0.3)
6	OW/OB	OW/OB	No	No	12 (0.3)	10 (0.4)	2 (0.1)
7	OW/OB	OW/OB	OW/OB	No	14 (0.3)	11 (0.5)	3 (0.2)
8	OW/OB	OW/OB	OW/OB	OW/OB	258 (6.2)	148 (6.2)	110 (6.2)
9	No	OW/OB	OW/OB	No	55 (1.3)	32 (1.0)	23 (1.3)
10	OW/OB	No	No	OW/OB	10 (0.2)	8 (0.3)	2 (0.1)
11	No	No	OW/OB	No	76 (1.8)	47 (2.0)	29 (1.6)
12	No	OW/OB	No	No	99 (2.4)	55 (2.3)	44 (2.5)
13	OW/OB	OW/OB	No	OW/OB	18 (0.4)	17 (0.7)	1 (0.1)
14	OW/OB	No	OW/OB	OW/OB	13 (0.3)	10 (0.4)	3 (0.2)
15	No	OW/OB	No	OW/OB	172 (4.1)	103 (4.3)	69 (3.9)
16	OW/OB	No	OW/OB	No	6 (0.1)	4 (0.2)	2 (0.1)

Legend: OW/OB – classified as overweight/obese; No – classified as non-overweight/obese.

Appendix 2. Results of unadjusted and adjusted (multivariable) multinomial logistic regression assessing the effect of maternal BMI on females' weight status trajectories.

			Never OW/OB	Mid-adulthood onset of OW/OB		Early-adulthood onset of OW/OB		Childhood onset of OW/OB		Other trajectories	
Risk Factor	**Included n (%)**	**Missing n (%)**	***reference***	**RRR (95% CI)**	***p***	**RRR (95% CI)**	***p***	**RRR (95% CI)**	***p***	**RRR (95% CI)**	***p***
Unadjusted model	2,339 (97.8)	53 (2.2)									
Maternal BMI				1.10 (1.05-1.15)	<0.001	1.21 (1.17-1.25)	<0.001	1.31 (1.25-1.37)	<0.001	1.17 (1.13-1.22)	<0.001
Adjusted for birth factors	1,678 (70.2)	714 (29.8)									
Maternal BMI				1.08 (1.02-1.15)	0.007	1.20 (1.15-1.24)	<0.001	1.31 (1.24-1.38)	<0.001	1.19 (1.13-1.24)	<0.001
Gestational age				0.93 (0.84-1.03)	0.15	0.98 (0.91-1.06)	0.65	0.95 (0.84-1.08)	0.43	0.94 (0.86-1.03)	0.18
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				1.45 (0.77-2.74)	0.25	0.80 (0.48-1.33)	0.39	1.13 (0.52-2.47)	0.75	0.72 (0.36-1.44)	0.35
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				1.04 (0.64-1.68)	0.88	1.28 (0.91-1.80)	0.16	1.56 (0.79-3.09)	0.20	1.02 (0.67-1.57)	0.92
Adjusted for birth factors and puberty	1,672 (69.9)	720 (30.1)									
Maternal BMI			reference	1.08 (1.02-1.14)	0.01	1.19 (1.14-1.23)	<0.001	1.30 (1.24 (1.37)	<0.001	1.18 (1.12-1.23)	<0.001
Gestational age			reference	0.93 (0.84-1.03)	0.15	0.98 (0.91-1.05)	0.57	0.93 (0.82-1.06)	0.30	0.93 (0.85-1.02)	0.14
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				1.45 (0.77-2.75)	0.25	0.81 (0.49-1.34)	0.41	1.16 (0.53-2.55)	0.71	0.73 (0.36-1.47)	0.37
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				1.04 (0.64-1.69)	0.86	1.32 (0.93-1.87)	0.12	1.73 (0.87-3.46)	0.12	1.09 (0.71-1.70)	0.69

Evidence of puberty at 10 years										
No			reference	-		-		-		-	
Yes				1.34 (0.88-2.04)	0.17	1.80 (1.36-2.38)	<0.001	3.63 (2.32-5.69)	<0.001	2.34 (1.66-3.30)	<0.001
Adjusted for birth factors, puberty and socioeconomic position	1,620 (67.7)	772 (32.3)									
Maternal BMI			reference	1.07 (1.01-1.14)	0.018	1.18 (1.13-1.22)	<0.001	1.29 (1.22-1.36)	<0.001	1.17 (1.12-1.23)	<0.001
Gestational age			reference	0.93 (0.84-1.03)	0.19	0.98 (0.91-1.06)	0.60	0.92 (0.80-1.05)	0.20	0.95 (0.86-1.05)	0.29
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				1.54 (0.79-2.98)	0.20	0.93 (0.55-1.56)	0.77	1.53 (0.69-3.43)	0.30	0.83 (0.41-1.70)	0.62
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				1.01 (0.61-1.67)	0.98	1.20 (0.83-1.72)	0.33	1.42 (0.70-2.90)	0.33	0.98 (0.62-1.54)	0.92
Evidence of puberty at 10 years											
No			reference	-		-		-		-	
Yes				1.35 (0.88-2.06)	0.17	1.83 (1.38-2.43)	<0.001	3.89 (2.46-6.15)	<0.001	2.40 (1.69-3.39)	<0.001
Socio-economic position											
Medium-low			reference	-		-		-		-	
High				0.86 (0.59-1.26)	0.44	0.72 (0.55-0.94)	0.015	0.48 (0.28-0.82)	0.008	0.89 (0.63-1.26)	0.51
Maternal education											
Less than A-level			reference	-		-		-		-	
A-level or above				0.93 (0.59-1.48)	0.77	0.82 (0.59-1.13)	0.23	0.72 (0.37-1.42)	0.35	0.62 (0.39-0.98)	0.039

Legend: BMI – Body mass index; CI - Confidence interval; OW/OB - Overweight or obesity; RRR - Relative risk ratio.

Appendix 3. Results of unadjusted and adjusted (multivariable) multinomial logistic regression assessing the effect of maternal BMI on males' weight status trajectories.

			Never OW/OB	Mid-adulthood onset of OW/OB		Early-adulthood onset of OW/OB		Childhood onset of OW/OB		Other trajectories	
Risk Factor	**Included n (%)**	**Missing n (%)**	***reference***	**RRR (95% CI)**	***p***	**RRR (95% CI)**	***p***	**RRR (95% CI)**	***p***	**RRR (95% CI)**	***p***
Unadjusted model	1,719 (96.5)	63 (3.5)									
Maternal BMI				1.05 (0.97-1.13)	0.26	1.16 (1.10-1.22)	<0.001	1.23 (1.16-1.31)	<0.001	1.09 (1.03-1.16)	0.006
Adjusted for birth factors	1,231 (69.1)	551 (30.9)									
Maternal BMI				1.06 (0.96-1.16)	0.25	1.19 (1.12-1.26)	<0.001	1.26 (1.16-1.36)	<0.001	1.12 (1.04-1.21)	0.004
Gestational age				0.99 (0.84-1.17)	0.94	0.87 (0.79-0.96)	0.007	0.89 (0.75-1.05)	0.15	1.13 (0.98-1.30)	0.10
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				1.96 (0.73-5.22)	0.18	3.01 (1.52-5.95)	0.002	0.84 (0.22-3.23)	0.81	1.81 (0.77-4.26)	0.17
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				0.65 (0.31-1.34)	0.25	0.95 (0.58-1.56)	0.83	1.10 (0.44-2.72)	0.84	0.73 (0.38-1.39)	0.34
Adjusted for birth factors and puberty	1,226 (68.8)	556 (31.2)									
Maternal BMI				1.06 (0.96-1.16)	0.23	1.19 (1.12-1.27)	<0.001	1.26 (1.16-1.37)	<0.001	1.11 (1.03-1.20)	0.008
Gestational age				0.99 (0.84-1.17)	0.93	0.87 (0.79-0.97)	0.009	0.89 (0.75-1.05)	0.16	1.14 (0.98-1.31)	0.08
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				1.95 (0.73-5.20)	0.18	2.92 (1.47-5.79)	0.002	0.81 (0.21-3.12)	0.77	1.79 (0.76-4.22)	0.18
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				0.64 (0.31-1.34)	0.24	0.94 (0.57-1.55)	0.80	1.06 (0.43-2.62)	0.91	0.77 (0.40-1.47)	0.43

Evidence of puberty at 10 years										
No			reference	-		-		-		-	
Yes				2.34 (0.32-16.99)	0.40	4.80 (1.14-20.19)	0.032	8.86 (1.72-45.66)	0.009	2.80 (0.50-15.62)	0.24
Adjusted for birth factors, puberty and socioeconomic position	1,181 (66.3)	601 (33.7)									
Maternal BMI			reference	1.04 (0.95-1.15)	0.40	1.18 (1.11-1.25)	<0.001	1.26 (1.16-1.37)	<0.001	1.09 (1.01-1.18)	0.03
Gestational age				1.00 (0.84-1.18)	0.96	0.88 (0.79-0.98)	0.016	0.90 (0.76-1.07)	0.22	1.16 (1.0-1.35)	0.05
Birth weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				2.01 (0.75-5.37)	0.17	2.86 (1.44-5.68)	0.003	0.80 (0.21-3.07)	0.74	1.83 (0.77-4.32)	0.17
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				0.66 (0.30-1.46)	0.31	0.85 (0.50-1.43)	0.53	0.93 (0.36-2.37)	0.87	0.64 (0.32-1.26)	0.20
Evidence of puberty at 10 years											
Yes			reference	-		-		-		-	
No				2.50 (0.34-18.22)	0.37	4.70 (1.11-19.85)	0.035	9.12 (1.76-47.22)	0.008	2.89 (0.52-16.18)	0.23
Socio-economic position											
Medium-low			reference	-		-		-		-	
High				1.05 (0.58-1.89)	0.88	0.82 (0.67-1.18)	0.28	0.86 (0.45-1.65)	0.66	0.75 (0.45-1.25)	0.27
Maternal education											
Less than A-level			reference	-		-		-		-	
A-level or above				0.69 (0.33-1.45)	0.33	0.81 (0.52-1.26)	0.34	0.87 (0.39-1.94)	0.73	0.74 (0.39-1.41)	0.36

Legend: BMI – Body mass index; CI - Confidence interval; OW/OB - Overweight or obesity; RRR - Relative risk ratio.

Appendix 4. Results of unadjusted and adjusted (multivariable) multinomial logistic regression assessing the effect of paternal BMI on females' weight status trajectories.

			Never OW/OB	Mid-adulthood onset of OW/OB		Early-adulthood onset of OW/OB		Childhood onset of OW/OB		Other trajectories	
Risk Factor	Included n (%)	Missing n (%)	*reference*	RRR (95% CI)	*p*	RRR (95% CI)	*p*	RRR (95% CI)	*p*	RRR (95% CI)	*p*
	2,250	142									
Unadjusted model	(94.1)	(5.9)									
Paternal BMI				1.09 (1.03-1.15)	0.002	1.17 (1.12-1.21)	<0.001	1.29 (1.21-1.36)	<0.001	1.15 (1.10-1.21)	<0.001
	1,629	763									
Birth factors	(68.1)	(31.9)									
Paternal BMI				1.05 (0.98-1.12)	0.14	1.18 (1.13-1.23)	<0.001	1.27 (1.19-1.37)	<0.001	1.15 (1.09-1.21)	<0.001
Gestational age				0.93 (0.83-1.03)	0.14	1.00 (0.93-1.07)	0.92	0.94 (0.83-1.07)	0.38	0.96 (0.87-1.05)	0.36
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				1.56 (0.83-2.95)	0.17	0.97 (0.59-1.60)	0.91	1.69 (0.80-3.57)	0.17	0.82 (0.41-1.64)	0.58
Breast- feeding											
≥3 months			reference	-		-		-		-	
<3 months				1.09 (0.67-1.78)	0.73	1.24 (0.88-1.74)	0.23	1.25 (0.65-2.38)	0.50	1.04 (0.68-1.60)	0.86
Adjusted for birth factors and puberty	1,625 (67.9)	767 (32.1)									
Paternal BMI				1.05 (0.98-1.12)	0.17	1.17 (1.12-1.22)	<0.001	1.25 (1.17-1.34)	<0.001	1.13 (1.07-1.20)	<0.001
Gestational age				0.92 (0.83-1.24)	0.13	0.99 (0.92-1.07)	0.84	0.93 (0.81-1.06)	0.27	0.95 (0.86-1.04)	0.28
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				1.57 (0.83-2.97)	0.17	0.98 (0.59-1.62)	0.93	1.75 (0.82-3.72)	0.15	0.84 (0.42-1.68)	0.63
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				1.10 (0.67-1.80)	0.70	1.30 (0.92-1.84)	0.14	1.37 (0.71-2.65)	0.34	1.13 (0.73-1.75)	0.59

Evidence of puberty at 10 years										
No			reference	-		-		-		-	
Yes				1.32 (0.87-2.02)	0.19	1.81 (1.37-2.40)	<0.001	3.69 (2.35-5.78)	<0.001	2.36 (1.67-3.33)	<0.001
Adjusted for birth factors, puberty and socioeconomic position	1,578 (66.0)	814 (34.0)									
Paternal BMI				1.04 (0.97-1.11)	0.28	1.16 (1.11-1.22)	<0.001	1.25 (1.16-1.34)	<0.001	1.13 (1.06-1.20)	<0.001
Gestational age				0.93 (0.83-1.03)	0.16	0.99 (0.92-1.07)	0.78	0.92 (0.80-1.05)	0.19	0.96 (0.87-1.06)	0.45
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				1.68 (0.86-3.26)	0.13	1.13 (0.67-1.90)	0.65	2.30 (1.05-5.00)	0.036	0.98 (0.48-1.98)	0.95
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				1.04 (0.62-1.75)	0.88	1.18 (0.82-1.70)	0.37	1.14 (2.41-6.02)	<0.001	1.00 (0.63-1.58)	1
Evidence of puberty at 10 years											
No			reference	-		-		-		-	
Yes				1.32 (0.86-2.03)	0.2	1.83 (1.37-2.43)	<0.001	3.81 (2.41-6.02)	<0.001	2.41 (1.70-3.41)	<0.001
Socio-economic position											
Medium-low			reference	-		-		-		-	
High				0.80 (0.54-1.17)	0.24	0.66 (0.51-0.87)	0.003	0.42 (0.25-0.73)	0.002	0.81 (0.57-1.14)	0.22
Maternal education											
Less than A-level			reference	-		-		-		-	
A-level or above				0.93 (0.58-1.49)	0.77	0.95 (0.68-1.32)	0.74	0.86 (0.44-1.68)	0.65	0.69 (0.43-1.09)	0.11

Legend: BMI – Body mass index; CI - Confidence interval; OW/OB - Overweight or obesity; RRR - Relative risk ratio.

Appendix 5. Results of unadjusted and adjusted (multivariable) multinomial logistic regression assessing the effect of paternal BMI on males' weight status trajectories

Risk Factor	Included n (%)	Missing n (%)	Never OW/OB	Mid-adulthood onset of OW/OB		Early-adulthood onset of OW/OB		Childhood onset of OW/OB		Other trajectories	
			reference	RRR (95% CI)	*p*	RRR (95% CI)	*p*	RRR (95% CI)	*p*	RRR (95% CI)	*p*
Unadjusted model	1,650 (92.6)	132 (7.4)									
Paternal BMI				1.05 (0.96-1.15)	0.32	1.17 (1.10-1.23)	<0.001	1.27 (1.17-1.37)	<0.001	1.07 (0.99-1.15)	0.08
Adjusted for birth factors	1,190 (66.8)	592 (33.2)									
Paternal BMI				1.03 (0.93-1.15)	0.59	1.19 (1.11-1.27)	<0.001	1.28 (1.16-1.42)	<0.001	1.07 (0.98-1.17)	0.16
Gestational age				0.98 (0.83-1.16)	0.82	0.87 (0.79-0.97)	0.01	0.89 (0.76-1.06)	0.19	1.11 (0.96-1.29)	0.15
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				2.08 (0.78-5.56)	0.15	2.97 (1.50-5.89)	0.002	0.83 (0.22-3.18)	0.79	1.79 (0.75-4.27)	0.19
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				0.68 (0.32-1.44)	0.32	0.97 (0.59-1.59)	0.89	0.95 (0.40-2.28)	0.92	1.80 (0.41-1.55)	0.50
Adjusted for birth factors and puberty	1,185 (66.5)	597 (33.5)									
Paternal BMI				1.03 (0.92-1.14)	0.64	1.18 (1.11-1.27)	<0.001	1.28 (1.16-1.41)	<0.001	1.06 (0.97-1.16)	0.19
Gestational age				0.98 (0.83-1.16)	0.81	0.87 (0.79-0.97)	0.01	0.90 (0.76-1.06)	0.19	1.12 (0.97-1.30)	0.13
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				2.07 (0.77-5.54)	0.15	2.87 (1.45-5.70)	0.003	0.81 (0.21-3.09)	0.75	1.76 (0.74-4.20)	0.20
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				0.68 (0.32-1.43)	0.31	0.97 (0.58-1.59)	0.89	0.93 (0.39-2.22)	0.87	0.84 (0.43-1.64)	0.61

Evidence of puberty at 10 years										
No			reference	-		-		-		-	
Yes				2.39 (0.33-17.37)	0.39	4.77 (1.13-20.09)	0.033	8.83 (1.71 (45.58)	0.009	3.65 (0.69-19.18)	0.13
Adjusted for birth factors, puberty and socioeconomic position	1,150 (64.5)	632 (35.5)									
Paternal BMI				1.02 (0.91-1.14)	0.69	1.18 (1.10-1.26)	<0.001	1.25 (1.13-1.39)	<0.001	1.05 (0.95-1.15)	0.32
Gestational age				0.97 (0.82-1.16)	0.77	0.87 (0.79-0.97)	0.012	0.90 (0.76-1.07)	0.25	1.15 (0.98-1.34)	0.08
Birth-weight											
Birth-weight ≤4 kg			reference	-		-		-		-	
Birth-weight >4 kg				2.19 (0.81-5.88)	0.12	2.88 (1.45-5.74)	0.003	0.83 (0.22-3.19)	0.79	1.80 (0.75-4.31)	0.19
Breast-feeding											
≥3 months			reference	-		-		-		-	
<3 months				0.70 (0.31-1.57)	0.38	0.89 (0.52-1.50)	0.66	0.82 (0.33-2.01)	0.66	0.72 (0.36-1.45)	0.37
Evidence of puberty at 10 years											
No			reference	-		-		-		-	
Yes				2.60 (0.36-18.96)	0.35	4.74 (1.12-20.02)	0.034	9.07 (1.75-46.91)	0.009	3.72 (0.71-19.62)	0.12
Socio-economic position											
Medium-low			reference	-		-		-		-	
High				1.10 (0.60-2.01)	0.75	0.85 (0.59-1.23)	0.38	0.93 (0.49-1.77)	0.82	0.83 (0.50-1.40)	0.49
Maternal education											
Less than A-level			reference	-		-		-		-	
A-level or above				0.64 (0.30-1.37)	0.25	0.74 (0.47-1.16)	0.19	0.74 (0.33-1.65)	0.47	0.78 (0.41-1.47)	0.44

Legend: BMI – Body mass index; CI - Confidence interval; OW/OB - Overweight or obesity; RRR - Relative risk ratio.

Longitudinal and Life Course Studies 2015 Volume 6 Issue 2 Pp 173 – 189 *ISSN 1757-9597*

Vocabulary from adolescence to middle age

Alice Sullivan Institute of Education, University College London, UK
a.sullivan@ioe.ac.uk
Matt Brown Institute of Education, University College London, UK

(Received June 2014 Revised January 2015) http://dx.doi.org/10.14301/llcs.v6i2.310

Abstract

The 1970 British Cohort Study (BCS70) is rich in cognitive measures taken during childhood and adolescence, and also includes adult measures of literacy and numeracy. The Age 42 survey in 2012 included a cognitive scale which had been used previously with the cohort in childhood – a vocabulary test first taken in 1986, when the cohort members were 16 years old. This paper asks how vocabulary scores changed between the ages of 16 and 42, taking account of early social background and childhood reading behaviour, but also examining the influence of educational and labour market attainment and reading for pleasure in mid-life. We find that both educational and occupational attainment, and reading habits in child-hood and adulthood, are linked to the development of vocabulary over time.

Introduction

A substantial literature examines the longitudinal development of cognition in childhood and adolescence. The growth in cognitive inequalities according to socio-economic status during childhood has been documented by analyses of the British cohort studies of 1946, 1958, and 1970 (Douglas, 1964, Feinstein, 2003, Feinstein, 2004, Fogelman, 1983, Fogelman & Goldstein, 1976, Sullivan, Ketende & Joshi, 2013). The BCS70 childhood cognitive scores have been shown to be important predictors of a range of adult outcomes, including employment (Breen & Goldthorpe, 2001) and health (Batty, Deary, Schoon & Gale, 2007). There is also a growing literature examining the decline in cognitive function from mid-life into old age (Richards & Sacker, 2003, Richards & Hatch, 2011, Richards, Shipley, Fuhrer & Wadsworth, 2004, Singh-Manoux, Kivimaki, Glymour, Elbaz, Berr, et al., 2012). The period between adolescence and early mid-life is relatively neglected however. Yet mental development surely continues post-16, and may be expected to be related to mental stimulation in adulthood, including further and higher education, occupational experience and leisure activities. For example, evidence from the 1946 birth cohort has demonstrated that both adult education and physical exercise are linked to cognitive attainment in mid-life (Hatch, Feinstein, Link, Wadsworth & Richards, 2007, Richards, Hardy & Wadsworth, 2003). The present article contributes to the literature by investigating the relationships between adult accomplishments and activities, including educational attainment, social class, and reading, and adult vocabulary, while taking into account parental background, earlier reading, and prior vocabulary. Is vocabulary mainly formed in childhood and adolescence, or do experiences in adulthood also play a role?

The complexity and level of intellectual stimulus provided by an individual's environment influences cognitive performance. As such, increased cognitive demands and rewards to cognitive skills have been put forward as an explanation of massive IQ score gains in populations over time (Flynn, 1987). The relationship between cognition and environmental stimulus is a reciprocal one, meaning that high cognitive functioning leads people into more stimulating environments and activities, which in turn promote cognitive function (Schooler & Mulatu, 2001, Schooler, Mulatu & Oates, 2004). This reciprocal relationship provides a plausible

model for the increased heritability of IQ as children get older (Dickens & Flynn, 2001), and may help to account for the growth of cognitive inequalities during childhood that has been observed in the British birth cohorts.

The relationship between cognition and education is an example of this reciprocity, as cognitive high achievers tend to get more and higher quality education, thus further improving their cognitive attainment (Deary & Johnson, 2010). But the extent of causality in one direction or the other is typically very difficult to unpack, and longitudinal data is clearly vital in tackling such questions.

Jobs which demand self-direction rather than obedience protect against cognitive decline in older workers (Schooler, Mulatu & Oates, 2004). Again, this is a reciprocal relationship. It is well established that high cognitive performance is linked to upward social mobility, and it is also plausible that jobs of higher social class status should be linked to cognitive development between adolescence and mid-life.

There is also a substantial literature on 'cultural capital' and educational attainment (Bourdieu & Passeron, [1977] 1990). While the concept of cultural capital can be somewhat opaque, empirical researchers have operationalised cultural capital in terms of leisure activities, and have found that reading has a distinctive link to educational attainment, whereas other 'cultured' pursuits such as going to art galleries or playing a musical instrument appear to be far less relevant to educational attainment once family background factors are taken into account (Crook, 1997, De Graaf, De Graaf & Kraaykamp, 2000, Sullivan, 2001). Drawing on Ganzeboom's distinction between status-seeking and information-processing versions of cultural reproduction theory (Ganzeboom, 1982) we have argued that reading is distinctive because it introduces the reader to new vocabulary and new concepts (Sullivan, 2002, Sullivan, 2007).

In previous work, we have demonstrated the role of reading for pleasure in the cognitive development of the BCS70 cohort members up to age 16 (Sullivan & Brown, 2015 in press). We were able to demonstrate that it is not just that bright children read more, but reading is also linked to greater cognitive progress for children with the same levels of prior cognitive attainment. Reading was associated with progress both in mathematics and vocabulary, but the link was strongest for vocabulary. To our knowledge, this study is the first to examine the influence of reading on cognitive development up to mid-life.

While our previous work established the importance of time spent reading, we were not able to examine the potential importance of the types of books that individuals read. Some sociologists see cultural tastes as an important expression of class identity (Bourdieu, 1984, Savage, Devine, Cunningham, Taylor, Li, et al., 2013), and social differences in cultural tastes and participation have been documented extensively (see, for example, (Chan & Goldthorpe, 2007a, Miles & Sullivan, 2012). The question of whether exclusive high-brow tastes have been replaced by cultural omnivorousness as a marker of high social status has generated a long-running debate (Peterson, 1997). In this paper we are able to examine the genres that cohort members read, and whether these literary tastes were linked to the development of their vocabularies. We might expect that reading 'high-brow' genres, which typically use more complex language and a wider range of vocabulary, would be more likely to lead to improvements in readers' vocabularies over time.

Vocabulary is a measure of crystallised cognitive ability, capturing an important aspect of literacy. Vocabulary is relatively robust, in that it is more resistant to decline during ageing than other cognitive measures, such as memory (Rabbitt, 1993). Adult vocabulary has been found to be protective against general cognitive decline (Richards, et al., 2004). While vocabulary is just one component of cognition, its distinctive importance comes from the fact that knowledge of words is both an adjunct to knowledge of concepts and assists further learning (Hirsch, 1983).

Research questions

In previous work, we have examined the predictors of vocabulary and math scores at age 16, and found that reading for pleasure in childhood was a powerful predictor of progress in both vocabulary and math, but especially of vocabulary. Inequalities in educational attainment continued to expand between the ages of ten and 16. In this paper we examine the role of both childhood and adult circumstances and childhood and adult reading behaviour on vocabulary change between the ages of 16 and 42.

1. How do adult reading habits vary according to educational status?
2. Do vocabulary scores continue to diverge according to childhood socio-economic circumstances between 16 and 42?
3. Do post-16 educational and occupational experiences have an influence on vocabulary change between 16 and 42?
4. Does reading for pleasure in childhood and adulthood have a continuing role in vocabulary development post-16? If so, does it matter what people read, or simply how much they read?
5. Does reading have a distinctive role in vocabulary development? We will test whether playing a musical instrument is linked to vocabulary development in a similar way to reading.

Data

The 1970 British Cohort Study (BCS70) follows the lives of more than 17,000 people born in England, Scotland and Wales in a single week of 1970 (Elliott & Shepherd, 2006). Over the course of cohort members' lives, the BCS70 has collected information on health, physical, educational and social development, and economic circumstances among other factors. Since the birth survey in 1970, there have been eight surveys (or 'waves') at ages 5, 10, 16, 26, 30, 34, 38 and 42. This paper is part of a Special Issue of Longitudinal and Life Course Studies Journal, which showcases the possibilities of the age 42 data.

Of the 17,284 study members who took part in the original birth survey, a total of 9,354 (54%) took part in the Age 42 survey, as shown by Mostafa and Wiggins' paper in this volume (Mostafa & Wiggins, 2015), which provides a full account of attrition and changes in sample composition over time. In addition, a further 488 immigrants who joined the study at 5, 10 or 16 also participated at 42, giving a total of 9,842 participants of whom 9,432 (96%) completed the vocabulary test. This group forms the analytical sample for this paper.

The 1970 cohort study is rich in measures of various aspects of cognition, covering reading, spelling, numeracy, verbal and non-verbal reasoning throughout the early years. The childhood test scores (up to age 10) have been analysed extensively, including influential work by Feinstein (2003; 2004). A comprehensive guide to the cognitive scores used in childhood is available (Parsons, 2014). The age 16 scores have been used less frequently (Duncan, Bergman, Duckworth, Kokko, Lyyra, et al., 2012, Sullivan & Brown, 2015 in press). Work using BCS70 to assess adult literacy and numeracy at age 34 has also been influential (Bynner & Parsons, 2006).

Tests of vocabulary were included in the surveys at age 5, 10, 16 and 42. At age 5, vocabulary was assessed using the English Picture Vocabulary Test (EPVT)(Brimer & Dunn, 1962), a test of verbal vocabulary in which the child selected the picture from four options which corresponded to a given word. At age 10, study members completed a 100 item Pictorial Language Comprehension Test which was based on the EPVT. Of the 100 items, 71 were used to assess vocabulary. At age 16, vocabulary was assessed using a 75 item test where each item was a word followed by a list of five other words; the respondent was required to pick the one with the same meaning as the first word. The measure included at age 42 is a shortened 20-item version of the test used at 16.

It is important to acknowledge that people's levels of motivation and compliance, as well as potential stereotype-threat (anxiety due to the potential to confirm a negative stereotype about a group one belongs to) (Croizet & Claire, 1998, Spencer, Steele & Quinn, 1999) will affect their scores in cognitive tests. We also acknowledge that multiple-choice tests do not capture the full range of academic skills, and girls tend to fare worse in multiple-choice tests than in other forms of assessment (Gipps & Murphy, 1994). We do not interpret the tests used here as providing an estimation of innate intelligence. They are simply tests of attainment based on the capability and motivation to complete a particular task under given conditions.

Analytical strategy

Our analysis investigates the influence of childhood, adolescent and adult characteristics on vocabulary at age 42. We use linear regression, with percentage scores in vocabulary as the outcome variable. In order to examine influences on longitudinal change in vocabulary between 16 and 42, we condition on vocabulary percentage scores at 5, 10 and 16 from model two onwards. Our approach is appropriate to the question at hand, which is whether various predictors have a net effect on vocabulary in mid-life over and above

vocabulary in childhood and adolescence. Our aim in this paper is not to provide a formal mediation analysis or growth model, and when we use the term 'growth' in the text we use it in the common-parlance sense of an increase or gain rather than in the sense of a formal growth model.

Because we exploit data from all of the childhood waves of the study, including the age 16 wave, the problem of missing data must be addressed. The age 16 survey employed sixteen separate survey instruments, and unfortunately coincided with a teachers' strike which affected the completion of those instruments, including cognitive tests, that were administered via schools (Dodgeon, 2008). This led to substantial instrument non-response, though, as shown by Mostafa and Wiggins' paper in this volume, the overall response and representativeness of the sample at this wave was good. Our response to the question of missing data is informed by Mostafa and Wiggins' work (Mostafa and Wiggins, 2015). Levels of missing data for the variables used in our analysis are provided in table 1. As list-wise deletion was not a practical option, we use multiple imputation to 'fill-in' values of any missing items in the variables selected for our analysis, adopting Schafer's algorithm under the assumption of 'missing at random' (MAR). In order to strengthen the MAR assumption and to protect against departures from multivariate normality we included a set of auxiliary variables in our imputation model (Schafer, 1997). The parameter estimates are valid under the MAR assumption, which in this case implies that the variables included in our models are the only major predictors of missingness in BCS70. All reported analyses are averaged across twenty replicates based upon Rubin's Rule for the efficiency of estimation under a reported degree of missingness across the whole data set of just under 20% (Little & Rubin, 1987). The analytical sample consists of the 9,432 cohort members who completed the age 42 vocabulary test.

Model 1: Social origins

Model 1 focuses solely on social background which is captured by parental social class and education. Social class is based on the NS-SEC (National Statistics Socio-Economic Classification), which groups occupations according to their employment relations and conditions (Goldthorpe, 1997). NS-SEC at age 10 has been derived recently for BCS70 (Gregg, 2012). Parental education is based on the highest qualification obtained by the mother or father (whichever is the highest). We use four categories: Degree (either a first degree or a higher degree); A levels or equivalent; lower than A levels (which includes school leaving qualifications taken at age 16); and no qualifications. A levels (and in Scotland the equivalent 'Highers') were (and remain) the academic-track qualification taken in British schools at age 18. The child's sex is also included in Model 1.

Model 2: Childhood reading and cognitive and educational attainment

Model 2 adds information regarding childhood reading and childhood cognitive and educational attainment. Frequency of reading was recorded at age 10 and 16. Information provided by parents at 16 has been used to summarise the availability of newspapers in the home (classified as tabloids, broadsheets, both or no national papers). The terms 'broadsheet' and 'tabloid' reflect the fact that, until the early 2000s, 'quality' newspapers in Britain were much larger than the standard tabloid format. Newspaper readership was (and remains) a strong cultural identifier (Chan & Goldthorpe, 2007b). The prose style of popular tabloids is simpler and geared towards a lower reading age and smaller vocabulary than the 'quality' papers or broadsheets.

We also include measures of whether the cohort member played a musical instrument at the ages of 10 and 16 in this model. Playing an instrument is seen as a classic 'beaux arts' measure, and children who play musical instruments are part of the stereotype of a 'cultured' middle class family. Playing an instrument may share some characteristics in common with reading, for example it demands a degree of self-directed effort. Playing a musical instrument therefore acts as a useful control variable. To the extent that reading matters specifically because it introduces individuals to new words and concepts, we would hypothesise that reading should be more powerfully linked than playing an instrument to progress in vocabulary.

Educational attainment is based on examination results at age 16 (1986). This cohort took public examinations in a range of subjects at age 16. Higher attaining pupils took O Level (Ordinary Level) examinations, while lower achievers took CSEs

(Certificate of Secondary Education). We derive a total point score from all O level and CSE examinations. An O Level grade A is awarded 7 points, grade B 6 points, continuing to a grade E being awarded 3 points. A CSE grade 1 is equivalent to an O Level grade C and is awarded 5 points, a grade 2 4 points, etc. The lowest CSE grade is grade 6, which is awarded 1 point.

Cognitive attainment is based on the vocabulary assessments completed at 5, 10 and 16 described above. The childhood vocabulary scores are correlated at 0.4 or below, i.e. not strongly enough for concerns regarding colinearity to arise.

Model 3: Educational and occupational attainment at 42

Model 3 adds information relating to post-16 educational attainment and adult occupation. Educational attainment is based on the highest qualification obtained by the age of 42. At the top end of the scale we use new data from the 2012 wave of the study to distinguish between 'elite degrees' from highly selective universities (operationalised as Russell Group[i]) and other degrees. A levels were the academic qualification taken by pupils at age 18 (for those who stayed on in academic-track education post-16). Any qualification lower than A level is classed as 'other' (this includes O levels and CSEs). The comparison category is 'no qualifications'. Occupational social class is based on the National Statistics Socio-Economic Classification (NS-SEC).

Model 4: Reading at 42

Finally, Model 4 adds information on the frequency of reading at age 42 along with summary measures of the types of books and newspapers read. As for childhood, we also include information on whether the cohort member played a musical instrument at 42.

Analysis

Who reads what?

In order to derive a classification of reading genres for our analysis, we first examined patterns of reading according to the educational status of respondents. The paper self-completion questionnaire at age 42 asked respondents 'How often do you read books in your spare time, not for work or study (including in electronic format)?'. This was followed up by two questions on preferred genres: 'Which of the following types of fiction books do you usually read?' and 'Which of the following types of factual books do you usually read?'. This was followed by a list of genres taken from standard bookshop section classifications which we expected would be familiar and meaningful to cohort members.

Table 1. Genres of books 'usually read' by highest qualification

	None	Other	A levels	Degree	Elite degree	All
FICTION						
Action / Adventure / War Fiction	17.5%	18.4%	22.0%	21.8%	25.1%	19.8%
Comics / Graphic Novels	4.7%	5.1%	6.2%	6.1%	5.9%	5.4%
Crime / Thrillers / Mystery	36.2%	39.8%	47.3%	48.9%	54.6%	43.0%
Classic Fiction	10.5%	13.0%	20.5%	29.0%	43.3%	18.7%
Contemporary Literary Fiction	5.2%	5.4%	15.8%	29.8%	47.5%	14.6%
Historical Fiction	9.0%	11.1%	17.3%	22.2%	29.8%	15%
Humour	17.8%	23.3%	30.5%	32.5%	33.7%	25.5%
Horror	11.6%	11.6%	12.8%	8.4%	7.7%	10.9%
Poetry	1.9%	2.1%	3.4%	4.6%	7.0%	3.1%
Romance	22.5%	27.7%	28.6%	21.8%	18.5%	24.7%
Science Fiction / Fantasy	17.6%	16.5%	23.5%	25.1%	27.2%	20.3%
Other Fiction	7.5%	9.0%	8.3%	9.6%	7.2%	8.5%
Do Not Read Fiction Books	33.0%	28.1%	14.6%	14.5%	8.7%	23.3%
Base	**2159**	**2889**	**1257**	**1742**	**573**	**8620**
FACTUAL						
Art / Photography	5.0%	5.9%	12.3%	13.5%	14.7%	8.7%
Autobiography	35.1%	40.0%	41.4%	40.7%	39.0%	39.1%
Biography	24.0%	27.8%	32.1%	35.5%	39.5%	29.8%
Career-specific / Professional	11.0%	12.9%	26.5%	42.8%	44.2%	22.5%
Computing / Technology	7.0%	7.7%	11.6%	13.0%	13.5%	9.5%
Cookery / Food and Drink	31.9%	36.9%	45.1%	46.2%	44.8%	39.2%
Family and Parenting	6.2%	8.1%	12.3%	14.0%	16.8%	10%
Health, Wellbeing, Self-help	12.7%	15.2%	23.2%	25.6%	23.3%	18.4%
DIY / Interiors / Gardening	17.9%	20.4%	26.7%	22.9%	22.9%	21.4%
Music	7.5%	8.5%	9.0%	9.9%	8.6%	8.6%
Religion / Philosophy	4.2%	4.5%	7.5%	11.3%	13.8%	6.9%
Science	5.7%	6.1%	8.2%	15.2%	24.3%	9.4%
Sport	14.8%	16.8%	14.5%	18.0%	16.3%	16.2%
Politics/Economics/Current	3.0%	2.7%	4.5%	11.9%	18.5%	5.9%
Travel	14.1%	18.7%	24.7%	28.8%	33.2%	21.4%
History	14.6%	17.6%	20.4%	27.0%	36.4%	20.4%
Other factual	8.5%	8.4%	8.8%	9.9%	10.1%	8.9%
Do Not Read Factual Books	25.8%	19.1%	8.4%	6.8%	4.7%	15.8%
Base	**2158**	**2894**	**1257**	**1746**	**572**	**8627**

The descriptive tables are based on raw rather than imputed data. Table 1 shows that the pattern of readership of different genres of fiction and factual books varies widely according to educational status. In terms of fiction, respondents with higher levels of education are generally more likely to read each genre, with some exceptions: horror was less popular with graduates than with non-graduates, and romance was most popular with respondents with intermediate qualifications (O levels and A levels) and less popular with graduates and those with no qualifications. Genres which would typically be classified as 'high culture' had the strongest educational gradients. This was most marked in the case of 'contemporary literary fiction' which was read by nearly half (48%) of respondents with elite degrees, 30% of those with other degrees, and only 5% of those with no qualifications. The difference between those with elite and other university degrees is particularly striking, and is also evident in the case of classic fiction, which was read by 43% of elite graduates and 29% of other graduates. The most popular genre across all educational categories is 'crime, thrillers and mystery', which encompasses texts at a wide range of levels of sophistication, and has a less sharp educational gradient than the high-brow genres.

Turning to factual books the readerships of genres such as sport and autobiography are fairly undifferentiated by educational status, while more intellectual topics such as science and politics, economics and current affairs show a marked educational gradient in readership, including a divide between elite graduates and other graduates.

We derived summary variables for fiction and non-fiction genres, classing genres as high-brow if the ratio of readers between the highest and lowest education categories was 2.5 or more, low-brow if the ratio was 1.5 or less, and middle-brow for those in between. We made an exception for the 'family and parenting' genre, which would have been categorised as high-brow on this basis, but which we chose to classify as middle-brow. We considered that the relationship between educational attainment and reading parenting books at age 42 is likely to be partly due to the relationship between educational attainment and age at first birth.

We classed readers as high-brow if they read any high-brow books, middle-brow if they read middle-brow books and no high-brow books, and low-brow if their reading was exclusively low-brow. We considered dividing high-brow univores, with exclusively high-brow tastes, from high-brow omnivores, but there were too few high-brow univores in the fiction category (4% of the sample) to justify this approach. We acknowledge the drawback that the categories that our derived variable is based on are likely to contain a wide range of books of different registers, e.g. some crime fiction is also literary fiction. Nevertheless, the analysis we have presented here shows that these categories are quite well differentiated in terms of the educational gradients of their readerships (Table 2). Study members were also asked about which newspapers they had read in the last month (including online newspapers) and this information was used to classify them as reading broadsheets only, tabloids only, broadsheets and tabloids or neither. Overall three quarters had read at least one newspaper in the last month and table 2 makes clear the strong relationship between education and the types of newspaper read. The difference between those with 'elite degrees' and other degrees is again striking with over three quarters (76%) of those with an elite degree reading a broadsheet (including 21% who also read a tabloid) compared with 57% of other graduates. A minority of non-graduates read broadsheets.

Table 2. Derived summary reading variables by highest qualification

	None	Other	AS/A-Levels/Diploma	Degree	Elite degree	All
Fiction						
None	32.1%	27.3%	14.5%	14.0%	8.2%	22.6%
Low brow	12.1%	11.8%	9.5%	5.6%	2.8%	9.7%
Middle brow	36.9%	38.3%	39.4%	31.2%	22.8%	35.6%
High brow	18.9%	22.5%	36.5%	49.2%	66.3%	32.0%
Base	**2092**	**2834**	**1238**	**1704**	**575**	**8443**
Factual						
None	24.7%	18.6%	8.2%	6.6%	4.2%	15.2%
Low brow	21.6%	20.4%	14.0%	7.8%	5.1%	16.2%
Middle brow	20.9%	23.6%	23.4%	15.4%	10.5%	20.4%
High brow	32.8%	37.4%	54.3%	70.2%	80.3%	48.3%
Base	**2093**	**2839**	**1239**	**1708**	**574**	**8453**
Newspapers						
Broadsheets and tabloids	10.7%	13.2%	21.0%	22.4%	20.9%	16.1%
Broadsheet only	6.8%	7.5%	14.5%	34.4%	55.7%	17.0%
Tabloid only	54.6%	53.7%	39.8%	21.5%	8.9%	42.4%
No newspaper	27.9%	25.7%	24.7%	21.7%	14.5%	24.6%
Base	**2119**	**2859**	**1245**	**1705**	**573**	**8501**

How does vocabulary vary according to respondents' characteristics?

Table 3 shows the percentage response for the categorical variables to be used in our regression analyses, and means for continuous variables. A comparison of mean percentage scores for vocabulary at age 16 and 42 is provided for each categorical variable, and correlations are provided for the continuous variables.

Table 3. Background characteristics and vocabulary at 16 and 42

		Imputed%	Original N	% missing	Age 16 - Mean vocab score (%)	Age 42 - Mean vocab score (%)
All			9,432		54.7	63.0
Sex	Male	48.0	4523	-	54.0	64.0
	Female	52.0	4909	-	55.3	62.1
	Missing	-	0	0		
Parental social class	Managerial/professional	32.2	2431	-	59.5	69.1
	Intermediate	28.2	2100	-	54.4	63.1
	Routine/Semi-routine	35.5	3107	-	60.0	58.4
	Long-term unemp/never worked	4.1	359	-	49.9	54.6
	Missing	-	1435	15.2		
Highest parental qual	No qualifications	44.2	3942	-	49.2	56.7
	Lower than A levels etc	34.9	3172	-	56.3	65.4
	A levels etc	6.0	518	-	60.6	69.4
	Degree	14.9	1398	-	64.4	73.6
	Missing	-	402	4.3		
Newspapers in home 16	Broadsheets and tabloids	5.2	317	-	58.9	68.5
	Tabloids only	57.2	3270	-	52.3	60.5
	Broadsheets only	9.3	616	-	65.0	74.1
	No papers in home	28.3	1682	-	55.1	63.5
	Missing	-	3547	37.6		
Book reading 16	More than once a week	26.7	1176	-	61.2	70.2
	Once a week	23.1	516	-	54.9	63.4
	Less than once a week	25.4	843	-	52.6	60.9
	Rarely/Never	24.8	1290	-	49.5	57.0
	Missing	-	5607	59.4		
Child reading (10)	Often	61.4	4792	-	58.4	66.9
	Sometimes	33.9	2942	-	49.5	57.8
	Never/hardly ever	4.7	437	-	43.4	50.8
	Missing	-	1261	13.4		
Plays musical instrument (10)	Yes	50.3	4072	-	57.3	65.7
	No	49.7	4068	-	52.0	60.3
	Missing	-	1292	13.7		
Plays musical instrument (16)	Yes	22.0	925	-	57.2	66.4
	No	78.0	2875	-	53.9	62.1
	Missing	-	5632	59.7		
Highest qual by 42	No qualifications	26.2	2467	-	47.0	53.5
	Lower than A levels etc	33.2	3128	-	51.6	59.7
	A levels etc	14.5	1369	-	57.5	66.6
	Degree	19.9	1874	-	62.9	72.9
	Elite degree	6.3	594	-	70.3	80.7
	Missing	-	0	0		

		Imputed%	Original N	% missing	Age 16 - Mean vocab score (%)	Age 42 - Mean vocab score (%)
Social class at 42	Managerial/professional	41.6	3899	-	60.3	70.4
	Intermediate	22.3	2075	-	53.3	61.9
	Routine/Semi-routine	24.6	2309	-	48.9	55.5
	Long-term unemp/never	11.5	1081	-	49.2	54.6
	Missing	-	68	.7	-	-
Frequency of reading 42	Read books every day	25.7	2265	-	60.4	70.9
	Several times a week	12.5	1052	-	58.5	67.7
	Once or twice per week	10.5	831	-	55.6	64.1
	At least once a month	9.6	737	-	53.8	61.8
	Every few months	13.8	1148	-	53.4	61.7
	At least once a year	10.6	894	-	50.6	58.5
	Never/less often	17.3	1530	-	46.8	51.8
	Missing	-	975	10.3	-	-
Reading fiction 42	No fiction	23.9	1912	-	47.9	53.2
	Low-brow	12.3	817	-	49.7	55.5
	Mid	34.3	3009	-	54.6	63.6
	High-brow	29.5	2705	-	62.3	73.4
	Missing	-	989	10.5	-	-
Reading factual 42	No fiction	16.0	1283	-	47.4	51.5
	Low-brow	18.0	1367	-	50.2	57.0
	Mid	21.2	1721	-	53.7	62.1
	High-brow	44.7	4082	-	59.5	70.0
	Missing	-	979	10.4	-	-
Reading newspapers 42	Broadsheets and tabloids	15.5	1366	-	58.8	69.1
	Broadsheets only	16.2	1445	-	64.5	76.0
	Tabloids only	41.6	3603	-	50.7	57.2
	No newspaper	26.7	2087	-	52.4	60.7
	Missing	-	931	9.9	-	-
Plays instrument	Yes	12.1	1026	-	60.9	71.4
	No	87.9	7392	-	53.8	61.9
	Missing	-	1014	10.8		
		Imputed Mean	**Original N**	**% missing**	**Corr. Age 16 vocab score (%)**	**Corr. Age 42 vocab score (%)**
Age 16 exam score	Score (max score 59)	13.6	3985	-	0.53	0.58
	Missing		5447	57.8	-	-
Prior vocabulary scores	Age 5 (%)	66.5	7257	-	0.34	0.39
	Missing	-	2175	23.1	-	-
	Age 10 (%)	58.3	7590	-	0.34	0.46
	Missing	-	1842	19.5	-	-
	Age 16 (%)	54.7	7349	-	-	0.61
	Missing	-	2083	22.1	-	-

First of all, we can see that respondents made substantial progress in vocabulary scores between age 16 (mean score 55%) and 42 (mean score 63%). Men scored slightly lower than women at 16, but achieved marginally higher scores than women at 42 (64% compared to 62%).

In terms of childhood characteristics, we see strong gradients in the vocabulary scores at both 16 and 42, as expected, with a general upward shift in scores from 16 to 42 across the board. This pattern is repeated for adult characteristics. For example, people with no qualifications scored 47% on the vocabulary test at age 16, rising to 54% at 42, while those with elite degrees scored 70% at 16, rising to 81% at 42.

Over a quarter (26%) said they read books in their spare time every day, and nearly half (49%) read for pleasure at least weekly, while 28% read books once a year or less. We can see a clear association between reading frequency and vocabulary scores at 16 and 42, showing that those with strong vocabularies at 16 were more likely to be frequent readers at 42. Both fiction and factual book genres were clearly associated with vocabulary scores, with respondents who read high-brow genres gaining higher scores. 16% of the sample read broadsheet newspapers only, a further 16% read both broadsheets and tabloids, 42% read tabloids only, and the remaining 27% had not read any newspapers in the last month. Those who read tabloids exclusively gained the lowest vocabulary scores at 16 and 42, lower even than those who did not read newspapers. Those who read a mixture of broadsheets and tabloids scored lower than those who read broadsheets exclusively. At age 42, tabloid readers attained vocabulary scores of 57% compared to 76% for broadsheet readers.

Playing a musical instrument, both in childhood and particularly in adulthood was also associated with higher scores – the 12% who played an instrument at 42 achieved average scores of 71%, compared to 62% amongst those who did not.

The correlation matrix of cognitive scores shows that age 5 and age 10 vocabulary are both correlated 0.34 with age 16 vocabulary. Correlations between age 5, age 10 and age 16 vocabulary and age 42 vocabulary were 0.39, 0.46 and 0.61 respectively. The examination score at age 16 is correlated 0.53 with vocabulary at 16 and 0.58 with vocabulary at 42.

Regression analysis of vocabulary at 42

Table 4 shows a series of linear regressions (OLS) predicting vocabulary scores at age 42. Our dependent variable is treated as a percentage score, as using standardised scores can mask increases in social differences due to increased variance in absolute skills over time (Magnuson, Waldfogel & Washbrook, 2012). Model 1 includes the child's sex and childhood social origins, captured via parental social class and educational qualifications. Men scored nearly 2 percentage points higher than women in this model. Social class differences were apparent, with a 5 percentage point advantage for the children of managers and professionals, and a 3 percentage point disadvantage for the children of the long-term unemployed compared with those whose parents had routine or semi-routine occupations. Parental education was substantially more powerful than social class in this model, with a 14% advantage for respondents whose parents had a university degree.

Model 2 includes both childhood reading behaviour, educational attainment at age 16, and prior vocabulary scores at the ages of 5, 10 and 16. In this and subsequent models we are essentially examining the predictors of change in vocabulary between the ages of 16 and 42. Vocabulary scores at 5, 10 and 16 and educational attainment at 16 are powerfully predictive of vocabulary at 42. Broadsheets and tabloids in the home are not significant, but reading books for pleasure at the ages of ten and 16 strongly predicts vocabulary at age 42, with a 4 percentage point advantage for those that read 'more than once a week' at 16, and a 5 percentage point advantage for those that read 'often' at age 10. We interpret this as showing that the advantage of childhood reading is not fully captured by attainment up to age 16, as childhood reading has a continued link with subsequent progress. As an indicator of 'beaux arts' cultural participation, Model 2 also includes whether musical instruments were played at 10 and 16. Playing a musical instrument at age 10 was a significant predictor of vocabulary at 42 but the coefficient is much smaller than those associated with regularly reading books, and playing an instrument at 16 is negative and non-significant.

Table 4. Vocabulary at 42: Linear regression models (N= 9, 432)

		Model 1		Model 2		Model 3		Model 4	
		B	Sig	B	Sig	B	Sig	B	Sig
	(Constant)	54.5	0.00	14.0	0.00	12.8	0.00	12.8	0.00
Sex(Ref = F)	Male	1.9	0.00	3.2	0.00	2.9	0.00	3.9	0.00
Parental social class	Managerial/professional	5.3	0.00	0.4	0.31	0.2	0.56	0.1	0.77
(Ref = Routine/Semi-routine)	Intermediate	2.4	0.00	0.2	0.62	0.0	0.92	0.1	0.76
	Long-term unemployed	-3.3	0.00	-1.8	0.02	-1.7	0.03	-1.5	0.05
Highest parental qual	Lower than A levels etc	7.5	0.00	1.6	0.00	1.3	0.00	1.2	0.00
(Ref=No quals)	A levels etc	10.7	0.00	1.0	0.14	0.7	0.30	0.7	0.30
	Degree	13.9	0.00	1.3	0.02	0.8	0.17	0.5	0.38
Newspapers in home 16	Broadsheets and tabloids			1.2	0.13	1.1	0.14	0.7	0.33
(Ref =No papers in home)	Tabloids only			0.1	0.88	0.3	0.59	0.4	0.35
	Broadsheets only			0.1	0.83	0.0	0.96	-0.7	0.28
Book reading 16	More than once a week			3.9	0.00	3.9	0.00	1.8	0.00
(Ref=Rarely/Never)	Once a week			2.2	0.00	2.2	0.00	0.9	0.09
	Less than once a week			1.1	0.06	1.0	0.07	0.3	0.55
Child reading (10)	Often			4.9	0.00	4.8	0.00	3.5	0.00
(Ref=Never/Hardly ever)	Sometimes			2.8	0.00	2.7	0.00	2.0	0.00
Musical instrument	Age 10			0.8	0.03	0.7	0.06	0.6	0.08
	Age 16			-0.3	0.66	-0.4	0.52	-0.6	0.37
Age 16 exam score	Score			4.9	0.00	4.1	0.00	3.5	0.00
Prior vocab scores	Age 5			0.1	0.00	0.1	0.00	0.1	0.00
	Age 10			0.3	0.00	0.3	0.00	0.2	0.00
	Age 16			0.4	0.00	0.3	0.00	0.3	0.00
Highest qual by 42	Lower than A levels etc					1.8	0.00	1.7	0.00
(Ref=No quals)	A levels etc					2.7	0.00	1.8	0.00
	Degree					3.3	0.00	1.9	0.00
	Elite degree					4.2	0.00	2.2	0.01
Social class at 42	Managerial/professional					2.9	0.00	2.4	0.00
(Ref = Routine/Semi-routine)	Intermediate					1.8	0.00	1.7	0.00
	Long-term unemployed					-0.7	0.15	-1.0	0.05
Frequency of reading 42	Read books every day							3.5	0.00
(Ref=Never/less often)	Several times a week							2.3	0.00
	Once or twice per week							1.5	0.04
	At least once a month							1.0	0.14
	Every few months							1.1	0.12
	At least once a year							0.4	0.59
Fiction at 42	Low-brow							0.8	0.23
(Ref = None)	Middle-brow							3.4	0.00
	High-brow							5.3	0.00
Factual books at 42	Low-brow							1.1	0.05
(Ref = None)	Middle-brow							2.3	0.00
	High-brow							3.0	0.00
Reading newspapers 42	Broadsheets and tabloids							0.2	0.64
(Ref=No newspapers)	Broadsheets only							1.2	0.03
	Tabloids only							-1.3	0.00
Musical instrument	Age 42							1.2	0.02
R-squared		0.13		0.52		0.53		0.56	
R-square change				0.39 (F=698.8, p<0.001)		0.01 (F=28.1, p<0.001)		0.03 (F=41.4, p<0.001)	

The gender difference increases in this model, confirming that men made more progress in vocabulary between the ages of 16 and 42 than women did. The influence of social origins is greatly attenuated in this model. Parental social class becomes largely insignificant, with the exception of a two percentage point disadvantage for cohort members who grew up with a long-term unemployed parent. The advantage due to a graduate parent declines from just under 14 percentage points to just over one percentage point, showing that this advantage is largely captured by factors which are already apparent by age 16. The model fit is substantially improved, increasing from R2= 0.13 in model 1 to 0.50 in model 2 (a change in R^2 of 0.39, F=698.8, p<0.001).

Model 3 includes adult educational and occupational attainment, measured at age 42. Post-16 educational attainment is clearly linked to progress between the ages of 16 and 42. Compared to those with no qualifications, respondents with higher qualifications achieved higher vocabulary scores, with an advantage in vocabulary gains amounting to around three percentage points for those with ordinary degrees, and four percentage points for those with elite degrees. Managerial and professional social class status was associated with an advantage of three percentage points, and intermediate social class was associated with a two percentage point advantage. But the influence of childhood reading is barely attenuated in this model, and childhood reading remains an important predictor of progress in vocabulary scores. The improvement in model fit is very small but statistically significant, R^2 increased from 0.52 to 0.53 (F=28.1, p<0.001).

Model 4 includes reading for pleasure at age 42. The frequency of reading for pleasure is clearly linked to vocabulary scores, with the greatest gains for those who read every day (four percentage points). Reading less often than once a week led to no advantage in the vocabulary test. Reading high-brow fiction is linked to an advantage in vocabulary gains of five percentage points, with a smaller advantage (three percentage points) for middle-brow fiction, and no advantage for reading low-brow fiction compared to not reading at all. Reading high-brow factual books was linked to an advantage of around three percentage points, while middle-brow factual reading made a smaller difference, and low-brow reading presented no advantage compared to not reading factual books at all. Reading broadsheet newspapers was linked to an advantage in progress in vocabulary of one percentage point, while reading tabloids was linked to a small but significant disadvantage (one percentage point). Reading both tabloids and broadsheets was not significantly different from reading no newspapers. Playing a musical instrument at 42 was associated with a small but significant advantage of one percentage point. Model fit was only marginally improved, although the improvement was statistically significant (R^2 increased from 0.56 (F=41.3, p<0.001).

Interestingly, the effect of the highest qualification achieved by age 42 is partially mediated by reading behaviour (the elite degree coefficient declines from 4.2 to 2.2). The social class coefficients on the other hand are only marginally reduced in this model. The male advantage rises to four percentage points in this model. This presents a puzzle, as men clearly make more progress in vocabulary between 16 and 42 despite reading less for pleasure than women. It may be that men's typically greater labour market experience and other forms of reading (e.g. reading at work) fill the gap for men. The influence of childhood social origins has been almost entirely attenuated in this model, with the exception of a small persistent disadvantage due to long-term parental unemployment, and a small advantage for respondents whose parents had intermediate qualifications. The advantage due to childhood reading is persistent, especially reading 'often' at age ten, which is linked to a four percentage point advantage in progress in vocabulary.

Conclusions

First, we asked how reading habits in mid-life varied according to educational status. We found a strong educational gradient in reading habits, with the most highly educated respondents more likely to read high-brow genres such as contemporary literary fiction. The divide between elite graduates and other graduates in this regard is interesting, and suggests that, just as broad social class categories can mask distinctive 'class fractions' (Savage, 1992), broad educational categories can also mask important differences. The patterns of reading behaviour shown here are consistent with both the 'status seeking' hypothesis regarding cultural participation as a way of asserting social

superiority (Bourdieu, 1984), and the 'information processing' hypothesis, according to which intellectually able and/or highly educated individuals prefer intellectually stimulating material (Ganzeboom, 1982, Sullivan, 2001, Sullivan, 2002, Sullivan, 2007). It may well be that both of these processes are at work, but the strong link between reading habits and both adolescent and mid-life vocabulary points towards the reciprocal relationship between intellectual performance and intellectual stimulus noted elsewhere in the literature. We found that the link between reading and vocabulary growth was far greater than the link between playing a musical instrument and vocabulary growth. This supports the 'information processing' hypothesis, and adds weight to the view that reading is distinctive, and is important to learning in a way that other forms of cultural participation are not.

Second, we asked whether social origins had a persistent influence on vocabulary, extending into mid-life. For the most part, the influence of social origins was fully mediated by factors apparent by age 16, including measured vocabulary at five, ten and 16, examination results at 16, and childhood reading behaviour. The small differences that remained according to parents' social class and education were in turn largely mediated by adult status attainment and reading patterns. However, unlike childhood social origins and playing an instrument in childhood, childhood reading habits appeared to exert a long-term influence, even when adult characteristics were included in the model. The greater progress in vocabulary made by frequent childhood readers was only partially mediated by mid-life reading habits. This long-term influence could be explained by frequent childhood readers continuing to read throughout their twenties and thirties, but unfortunately we do not have measures of reading for these waves of the study.

Third, we asked whether post-16 educational attainment and achieved social class in mid-life were linked to the development of vocabulary between 16 and 42. We found that both social class at 42 and qualifications attained by 42 were linked to vocabulary growth. Our analysis further supported the distinction between elite degrees and other degrees, as elite degree holders made the most progress.

Fourth, we asked whether adult reading habits were linked to progress in vocabulary between 16 and 42. We found that the frequency of reading for pleasure was positively linked to progress in vocabulary. However, what people read mattered as much as how often they read. Those who read high-brow fiction made greater vocabulary gains than those who read middle-brow fiction, and low-brow fiction readers made no more progress than non-readers. The gains linked to reading factual books were smaller than those for fiction, and similarly, low-brow factual reading was not linked to any vocabulary gain. Readers of broadsheets made more progress in vocabulary than people who didn't read newspapers, while tabloid readers actually made less progress than non-readers of newspapers. This negative tabloid link is in line with our previous work showing the presence of tabloid newspapers in the home during childhood was linked to poor cognitive attainment at age 16 (Sullivan & Brown, 2015 in press). Granted, there is an issue of endogeneity in interpreting these results, as people with larger vocabularies are likely to be more attracted to reading and to more difficult reading materials. Although we control for vocabulary up to age 16, we cannot entirely rule out the possibility that cohort members' reading behaviour may have been driven by a change in their vocabularies post-16 which was independent of their reading behaviour. A limitation of the data is that we have no data points for vocabulary or reading between the ages of 16 and 42.

In summary, our previous work showed that reading for pleasure was linked to cognitive progress up to age 16, especially in vocabulary scores, and this paper suggests that reading for pleasure both in childhood and adulthood continue to be linked to progress in vocabulary post-16. On a positive note, learning certainly doesn't stop at 16, and average scores increased substantially from 16 to 42. In addition, while childhood scores were powerful predictors of vocabulary at 42, post-16 educational and occupational attainment and leisure reading at 42 were all relevant, suggesting that learning continues to be influenced by adult activities.

How do the findings presented here contribute to the development of a theory of lifelong vocabulary growth? First of all, we address the issue of stability and change. Our results show the importance of childhood learning to adult

vocabulary. A high proportion of the variance in vocabulary at age 42 is attributable to the earlier test scores. The relationship with earlier scores solidifies with age. So, the correlations between vocabulary at 5 and 16 and at 10 and 16 are both 0.34, but the correlation between vocabulary at 10 and 42 is 0.46, and between 16 and 42 it is 0.61. This suggests a reduction in fluctuation as people get older. However, it does not imply that vocabulary is fixed by age 16. Large vocabulary gains occur between adolescence and mid-life, and these gains vary according to both childhood and adult characteristics and activities.

Secondly, we have engaged critically with alternative interpretations of the role of cultural capital, and brought this sociological perspective together with ideas from psychology regarding the reciprocal relationship between cognitive performance and exposure to intellectual stimulation. We argue that reading is a distinctive form of cultural participation in the extent to which it drives learning. Children from 'cultured' middle-class homes are relatively likely both to play an instrument and to be exposed to lots of books, but these exposures do not have the same influence on learning, either in the short or the longer term. This view is in line with our previous work on education, but extends the application of these ideas beyond success or failure during the school years, to learning throughout the life course. The absence of a substantial direct 'long shadow' of childhood disadvantage on adult vocabulary and the presence of a direct long-term advantage associated with childhood reading suggests food for thought in the development of our understanding of inequalities in education, learning and the life course.

Acknowledgements

We thank the BCS70 cohort members for their contribution over many years. Thanks to George Ploubidis and to the participants at a symposium for this Special Issue for advice and comments. Thanks also to three anonymous reviewers.

References

Batty, G.D., Deary, I.J., Schoon, I., & Gale, C.R. (2007). Mental ability across childhood in relation to risk factors for premature mortality in early life: the 1970 British Cohort Study. *Journal of Epidemiology and Community Health*, *61*, 977-1003. http://dx.doi.org/10.1136/jech.2006.054494

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*, London: Routledge and Kegan Paul.

Bourdieu, P., & Passeron, J.C. ([1977] 1990). *Reproduction in education, society and culture*, 2nd Edition, London ; Beverly Hills: Sage Publications.

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

Brimer, M., & Dunn, L. (1962). English Picture Vocabulary Test, London: Education Evaluation Enterprises.

Bynner, J., & Parsons, S. (2006). *New light on literacy and numeracy*: National Research and Development Centre for Adult Literacy and Numeracy London.

Chan, T.W., & Goldthorpe, J.H. (2007a). Social stratification and cultural consumption: The visual arts in England. *Poetics*, *35*, 168-190. http://dx.doi.org/10.1016/j.poetic.2007.05.002

Chan, T.W., & Goldthorpe, J.H. (2007b). Social Status and Newspaper Readership1. *American Journal of Sociology*, *112*, 1095-1134. http://dx.doi.org/10.1086/508792

Croizet, J.-C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, *24*, 588-594. http://dx.doi.org/10.1177/0146167298246003

Crook, C.J. (1997). *Cultural Practices and Socioeconomic Attainment: The Australian Experience*, Westport, Connecticut: Greenwood Press.

De Graaf, N.D., De Graaf, P.M., & Kraaykamp, G. (2000). Parental Cultural Capital and Educational Attainment in the Netherlands: A Refinement of the Cultural Capital Perspective. *Sociology of Education*, *73*, 92-111. http://dx.doi.org/10.2307/2673239

Deary, I.J., & Johnson, W. (2010). Intelligence and education: causal perceptions drive analytic processes and therefore conclusions. *International Journal of Epidemiology*, *39*, 1362-1369. http://dx.doi.org/10.1093/ije/dyq072

Dickens, W.T., & Flynn, J.R. (2001). Heritability estimates versus large environmental effects: the IQ paradox resolved. *Psychological review*, *108*, 346.

Dodgeon, B. (2008). *Guide to the Dataset: BCS70 16 year follow up: APU arithmetic test*, London: Centre for Longitudinal Studies.
Douglas, J.W.B. (1964). *The Home and the School*, London: MacGibbon and Kee.
Duncan, G.J., Bergman, L., Duckworth, K., Kokko, K., Lyyra, A.-L., Metzger, M., Pulkkinen, L., & Simonton, S. (2012). The Role of Child Skills and Behaviors in the Intergenerational Transmission of Inequality: A Cross-National Study, in J. Ermisch, M. Jäntti and T. Smeeding (eds) *From Parents to Children: The Intergenerational Transmission of Advantage*, New York: Russell Sage Foundation.
Elliott, J., & Shepherd, P. (2006). Cohort Profile: 1970 British birth cohort (BCS70). *International Journal of Epidemiology*, *35*, 836-843. http://dx.doi.org/10.1093/ije/dyl174
Feinstein, L. (2003). Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica*, *70*, 73-97. http://dx.doi.org/10.1111/1468-0335.t01-1-00272
Feinstein, L. (2004). Mobility in pupils' cognitive attainment during school life. *Oxford Review of Economic Policy*, *20*, 213-229. http://dx.doi.org/10.1093/oxrep/grh012
Flynn, J.R. (1987). Massive IQ Gains in 14 Nations: What IQ Tests Really Measure. *Psychological Bulletin*, *101*, 171-91. http://dx.doi.org/10.1037/0033-2909.101.2.171
Fogelman, K.R. ed. (1983). *Growing up in Great Britain*, London: Macmillan.
Fogelman, K.R., & Goldstein, H. (1976). Social Factors Associated with Changes in Educational Attainment between 7 and 11 Years of Age. *Educational Studies*, *2*, 95-109. http://dx.doi.org/10.1080/0305569760020201
Ganzeboom, H. (1982). Explaining Differential Participation in High-Cultural Activities - A Confrontation of Information-Processing and Status-Seeking Theories, in W. Raub (ed) *Theoretical Models and Empirical Analyses: Contributions to the Explanation of Individual Actions and Collective Phenomena*, Utrecht: E.S. Publications.
Gipps, C., & Murphy, P. (1994). *A fair test? : assessment, achievement and equity*, Milton Keynes: Open University Press.
Goldthorpe, J. (1997). The "Goldthorpe" Class Schema: Some Observations on Conceptual and Operational Issues in Relation to the ESRC Review of Government Social Classification, in D. Rose and K. O'Reilly (eds) *Constructing Classes: Towards a New Social Classification for the UK*, Swindon: ESRC/ONS.
Gregg, P. (2012). Occupational Coding for the National Child Development Study (1969, 1991-2008) and the 1970 British Cohort Study (1980, 2000-2008). SN 7023., Colchester: UK Data Archive.
Hatch, S.L., Feinstein, L., Link, B.G., Wadsworth, M.E.J., & Richards, M. (2007). The continuing benefits of education: adult education and midlife cognitive ability in the British 1946 birth cohort. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *62*, S404-S414. http://dx.doi.org/10.1093/geronb/62.6.S404
Hirsch, E.D. (1983). Cultural Literacy. *The American Scholar*, *52*, 159-169.
Little, R.J.A., & Rubin, D.B. (1987). Statistical analysis with incomplete data, New York: Wiley.
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. Jantti and T. Smeeding (eds) *From Parents to Children: The Intergenerational Transmission of Advantage, edited by* New York: Russell Sage Foundation.
Miles, A., & Sullivan, A. (2012). Understanding participation in culture and sport: Mixing methods, reordering knowledges. *Cultural Trends*, *21*, 311-324. http://dx.doi.org/10.1080/09548963.2012.726795
Mostafa, T., & Wiggins, D. (2015). Handling attrition and non-response in the 1970 British Cohort Study. *Longitudinal and Life Course Studies, 6,* 131-146.
Parsons, S. (2014). Childhood cognition in the 1970 British Cohort Study. *CLS Data Note*.
Peterson, R.A. (1997). The rise and fall of highbrow snobbery as a status marker. *Poetics*, *25*, 75-92. http://dx.doi.org/10.1016/S0304-422X(97)00013-2
Rabbitt, P. (1993). Does it all go together when it goes? The Nineteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology*, *46*, 385-434. http://dx.doi.org/10.1080/14640749308401055
Richards, M., & Sacker, A. (2003). Lifetime antecedents of cognitive reserve. *Journal of clinical and experimental neuropsychology*, *25*, 614-624. http://dx.doi.org/10.1076/jcen.25.5.614.14581
Richards, M., & Hatch, S.L. (2011). A life course approach to the development of mental skills. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *66*, i26-i35. http://dx.doi.org/10.1093/geronb/gbr013
Richards, M., Hardy, R., & Wadsworth, M.E.J. (2003). Does active leisure protect cognition? Evidence from a national birth cohort. *Social science & medicine*, *56*, 785-792. http://dx.doi.org/10.1016/S0277-9536(02)00075-8
Richards, M., Shipley, B., Fuhrer, R., & Wadsworth, M.E.J. (2004). Cognitive ability in childhood and cognitive decline in mid-life: longitudinal birth cohort study. *Bmj*, *328*, 552. http://dx.doi.org/10.1136/bmj.37972.513819.EE
Savage, M. (1992). *Property, bureaucracy and culture : middle-class formation in contemporary Britain*, London: Routledge.

Savage, M., Devine, F., Cunningham, N., Taylor, M., Li, Y., Hjellbrekke, J., Le Roux, B., Friedman, S., & Miles, A. (2013). A new model of social class? Findings from the BBC's Great British Class Survey Experiment. *Sociology*, *47*, 219-250. http://dx.doi.org/10.1177/0038038513481128

Schafer, J.L. (1997). *Analysis of incomplete multivariate data*, Florida: Chapman and Hall/CRC press. http://dx.doi.org/10.1201/9781439821862

Schooler, C., & Mulatu, M.S. (2001). The reciprocal effects of leisure time activities and intellectual functioning in older people: a longitudinal analysis. *Psychology and aging*, *16*, 466. http://dx.doi.org/10.1037/0882-7974.16.3.466

Schooler, C., Mulatu, M.S., & Oates, G. (2004). Occupational Self-Direction, Intellectual Functioning, and Self-Directed Orientation in Older Workers: Findings and Implications for Individuals and Societies. *American Journal of Sociology*, *110*, 161-197. http://dx.doi.org/10.1086/385430

Singh-Manoux, A., Kivimaki, M., Glymour, M.M., Elbaz, A., Berr, C., Ebmeier, K.P., Ferrie, J.E., & Dugravot, A. (2012). Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. *BMJ: British Medical Journal*, *344*.

Spencer, S.J., Steele, C.M., & Quinn, D.M. (1999). Stereotype threat and women's math performance. *Journal of experimental social psychology*, *35*, 4-28. http://dx.doi.org/10.1006/jesp.1998.1373

Sullivan, A. (2001). Cultural Capital and Educational Attainment. *Sociology*, *35*, 893-912. http://dx.doi.org/10.1177/0038038501035004006

Sullivan, A. (2002). Bourdieu and Education: How Useful is Bourdieu's Theory for Researchers? *Netherlands' Journal of Social Sciences*, *38*, 144-166.

Sullivan, A. (2007). Cultural capital, cultural knowledge and ability. *Sociological Research Online*, *12*. http://dx.doi.org/10.5153/sro.1596

Sullivan, A., & Brown, M. (2015 in press). Reading for pleasure and children's progress in vocabulary and mathematics. *British Educational Research Journal*.

Sullivan, A., Ketende, S., & Joshi, H. (2013). Social class and inequalities in early cognitive scores. *Sociology*, *47*, 1187-1206. http://dx.doi.org/10.1177/0038038512461861

Endnote

[i] We also include University of Bath and University of St Andrews in our elite category, as they have been consistently as highly selective as Russell Group institutions.

Longitudinal and Life Course Studies 2015 Volume 6 Issue 2 Pp 190 – 211 *ISSN 1757-9597*

Social mobility, parental help, and the importance of networks: evidence for Britain

Oscar Marcenaroo Gutierrez Department of Statistics and Econometrics, University of Malaga, Spain
odmarcenaro@uma.es
John Micklewright Institute of Education, University College London, UK
Anna Vignoles Faculty of Education, University of Cambridge, UK

(Received June 2014, Revised November 2014) http://dx.doi.org/10.14301/llcs.v6i2.313

Abstract

Greater levels of social mobility are widely seen as desirable on grounds of both equity and efficiency. Debate in Britain and elsewhere has recently focused on specific factors that might hinder social mobility, including the role of internships and similar opportunities that parents can sometimes secure for their children. We address the help that parents give their children in the job market using data from the recently collected age 42 year wave of the 1970 British Cohort Study. We consider help given to people from all family backgrounds and not just to graduates and those in higher level occupations, who have tended to be the focus in the debate in Britain. Our data measure whether respondents had ever had help to get a job from (i) parents and (ii) other relatives and friends, and the form of that help. We first assess the extent and type of help. We then determine whether people from higher socio-economic status families are more or less likely to have help. Finally we investigate whether help is associated with higher wages and higher occupation levels. The paper provides insight into whether the link between parental socio-economic background and the individual's own economic success can be explained in part by parental assistance to their children in getting jobs. We find parental help to have a strong social gradient. But we are unable to identify a clear link between any particular type of help – advice, help through contacts etc. – and individuals' wages or occupations. The data on parental assistance has some limitations, potentially causing measurement error, and for future research on this topic, better data need to be collected.

Keywords: social mobility, networks, family, wages

Introduction

The desire to achieve a socially mobile society, in which every person can succeed, regardless of who their parents are, is the professed aim of every major political party in the UK (Cabinet Office, 2011). Despite this, levels of intergenerational social mobility remain relatively low (Ermisch, Jantti & Smeeding, 2012). As a consequence, there is a lot of interest in whether help in the job market from families and their networks strengthens the inter-generational transmission of advantage and hence is a brake on social mobility, see for example Cabinet Office (2009, 2012) and Macmillan, Tyler and Vignoles (2013). There is interest in the issue in other countries too e.g. Bingley, Corak and Westergård-Nielsen (2011). Much of the discussion in the UK focuses on the upper end of the socio-economic distribution. For example, there are concerns about access to the professions, especially of graduates, and about the role of unpaid or poorly paid internships that affluent parents can sometimes secure for their children. Viewed in this light, personal connections and social networks provide the brake on mobility described and the only issue is how big a brake this represents.

However, connections and networks may operate at all levels in the labour market so that their effects on economic advantage and social mobility may be both more widespread and less obvious than at first sight. In contrast to much of the literature for the UK, we investigate the help given by parents, other relatives, and friends to people who work in the full range of occupations and who come from all family backgrounds. We draw on data from the newly available age 42 year wave of data collection in the 1970 British Cohort Study (BCS70), which contains information on help received in the job market for a sample of over 9,000 individuals. We relate this help to the data collected on parental circumstances when the cohort members were aged 10 years. As a by-product, we provide new evidence on the extent of social mobility in Britain, comparing the socio-economic status of the 42 year-olds measured in 2012-13 to that of their parents over 30 years earlier in 1980.[1] But our focus is on assessing whether the use of help, connections and networks can assist in explaining the strength of this relationship between parental socio-economic background and an individual's own success in the labour market. Our study is descriptive and draws on economic ways of thinking about networks, though mindful that the increased interest in networks in economics has drawn on the large sociological literature on this topic (Davern, 1997).

Some studies within the literature on barriers to access to professional careers have suggested that the use of networks is a possible mechanism by which social advantage is reproduced in the labour market.[2] The empirical evidence on the role of networks is limited but shows that they are generally effective ways in which individuals can generate job offers, with positive impacts on their labour market success (Loury, 1977, 2006; Holzer, 1988). As networks are a form of social capital it is also clear that access to networks is likely to be socially graded (Coleman, 1990). Individuals will have different types of networks depending on their family situation, their past work experience and the current industry that they work in (Ioannides & Loury, 2004; Fernandez & Castilla, 2001; Marsden & Gorman, 2001). To this extent, networks can potentially reinforce social and economic advantage.

Within this broader literature, key papers have examined a specific form of network, closely related to the focus of this paper, namely the support that parents give their children in securing employment. For example, Bingley et al. (2011) consider the extent to which children are employed in the same organisations as their parents and find that wealthier fathers are more likely to have sons who have worked for the same employer, implying a social grading to this form of assistance. Corak and Piraino (2010) suggest that in Canada around 70% of the sons of the wealthiest fathers have worked in the same firm as their father at some point (see also Kramarz and Skans, 2006). Unfortunately our data do not have this detail but they do contain information on the types of help parents have provided. These types of help vary enormously. For instance, they include 'advice'. Clearly one might argue that it is the role of all parents to provide advice to their children and that this may not constitute something that is socially undesirable. By contrast, if a parent provides an internship for their child this may be problematic from a social mobility perspective. In our discussion of the data below we consider carefully the importance of how help from parents and friends is measured and hence the meaning of our findings.

The first section explains the BCS70 data that we use, contrasting this with other types of data that have or could be used to investigate the issues. The second section begins by documenting the extent and type of help received in the job market that is recorded in the data, before analysing the association of this help with parental socio-economic status. In the third section we focus on help from parents. We report first on individuals' own subjective valuation of the help they have received before then, using regression models to obtain an objective valuation in terms of wages and occupations. The fourth section concludes.

The 1970 British Cohort Study data

Our data are drawn from the age 42 years sweep of BCS70, conducted in 2012-13. Some 9,800 cohort members responded – about 60% of the original cohort.[3] Respondents were interviewed face-to-face and the questions put to them included four on help in the job market from parents, other relatives and friends. Respondents were first asked:

'Have your parents ever done any of the things on this card to help you to get any job you have ever had? Please include internships and placements, even if unpaid.'

Provided advice
Helped with application forms
Acted as a reference

Recommended me to an employer that they work for or had worked for
Recommended me to an employer that they did not work for
Directly employed me
Helped me find a job through people that they know
Other (specify)
No – none of these

All types of help that were indicated were coded separately. Respondents were next asked:
'In your judgement, how much has the help [of all types] that you received from your parents to get a job contributed towards your current occupation or career?'[4]

A lot
A little
Or not at all?

Unfortunately the second question does not enable us to link the respondent's answer to the specific help provided. So if a parent provided many forms of help to their offspring, the second question only provides an indication of whether the help provided overall was considered useful or not.

Cohort members were then asked the same two questions in respect of help from 'your friends or other relatives'. We make less use of the data resulting from these questions. The definition of a friend is less clear-cut than a parent and, in part associated with this, there is likely to be a greater problem of recall; e.g. a friend who gave help 20 years ago by providing some advice could be long forgotten. The problem of recall is no doubt there for both sources of help however. There is also a question mark over the exogeneity of the reported help from friends. A mere acquaintance or friend-of-a-friend who mentioned an opening with their employer may have become a friend only later.

In both cases, parents and friends/other relatives, the 'help' may not always have a positive impact. Help could limit an individual's chances of moving up a career ladder by encouraging restricted aspirations. The questions are not nuanced enough to deduce whether respondents would have reported receiving help if they perceived it to have limited their ambitions, although we do have their own valuation of the value of all the help they have received. In any case, respondents may not realise that 'help' actually limited their options, though we test in our modelling whether the help is correlated with their subsequent economic success.

There are a number of additional limitations to the data. The questions on help cover the entire working lives to date of persons who entered the job market up to a quarter century prior to interview. Consequently, first, there is the issue of recall bias. Those who received parental help and were successful in their endeavours may be more (or perhaps less) likely to remember that they received help. Second, there is no information on the timing of the help. One might expect help on entry to the labour market to be particularly important. Third, we need to recognise that the data are very different to those that have been used in other studies of social networks, contacts and social mobility. For example, the data drawn on by Macmillan et al. (2013) from the *Destinations of Leavers from Higher Education* (DLHE) survey, which contain information on parental help, refer only to very recent graduates. Our BCS70 data refer to a much wider group of individuals, which is an advantage. But as a record of past help over 25 years to people now in mid-career they do not necessarily provide insight into what parents, other relatives and friends are currently doing in the UK job market to help young people, whether graduates or those who have never been to university. The nature of the current job market, including its structure of occupations, differs from that first faced by the BCS70 respondents in the late 1980s. The much greater prominence of white-collar internships is just one example. The decline in blue-collar jobs in manufacturing is another.

The BCS70 data also refer to help of various types that respondents report, which contrasts with the information from the administrative registers used by Bingley, Corak and Westergård-Nielsen (2011) and by Corak & Piraino (2010), that simply recorded whether parents and children had ever worked for the same employer. These administrative data do not suffer from recall bias and have the advantage of measuring something specific that we might say is clearly going to act as a brake on social mobility. If children secure jobs in their parents' firms, this is a mechanism which would undoubtedly encourage intergenerational social immobility and reinforce, by definition, the status quo. By contrast, the BCS70 questions ask about a range of help that might be provided by parents, other relatives and friends, not all of which may be considered deleterious to society. As has been described, we have data on whether the individual

received advice from his or her parents or had their help with application forms. These are perhaps desirable things for parents to do. The extent to which they are done and the impact that such assistance provides may be socially graded, in a similar way to other parent activities, such as reading aloud with a child. However, this does not mean that we want to discourage parents from doing these things. When we interpret our results we pay close attention to the meaning of the measures of help that we are using, and we largely focus on those that are arguably ways in which families may in some sense secure 'unfair' advantage for their children, recognising that this is a highly contested notion. In particular, we focus on specific measures which indicate the potential use of social or economic advantage to secure particular employment opportunities for a child, namely whether or not the parent provided the individual with help to get a job through the use of their contacts, and whether or not the parent recommended their child to a current or previous employer for a job.

It is clear from the literature that in many secondary data sets the measures of how individuals use networks to secure their jobs are potentially problematic. For example, the UK Labour Force Survey (LFS) asks respondents about the use of networks, and around a quarter of men aged 21-64 say they secured their job through "hearing from someone who worked there"[5]. But again the question is a bit imprecise and further it is presented as number 6 in a list where the interviewer is asked to code the first answer that applies (e.g. 'replied to a job advertisement' comes higher in the list). This might mean that 25% is a lower bound estimate if individuals use multiple methods to find their jobs. Whilst this LFS measure has been used in the migration literature to understand job search[6], it is not without its problems and we too must be mindful that our BCS measure may be subject to error.

Our strategy is to relate the help that BCS70 members report to their family socio-economic status (SES) before they entered the job market. By contrast, current SES is best seen as an outcome variable that the help may have furthered. To do this, we link the age 42 year data to the data collected from parents in the 1980s when the cohort members were aged 10 years. (The mean age of parents when the cohort members were aged 10 years was 34years for fathers and 31 years for mothers). The age 10 years data provide information on two measures of SES: family income and parental socio-economic group (for the latter we take father's group or the mother's if the father's is missing). All the analysis that follows is conducted on the sample of individuals present in both the age 42 year sweep and the earlier age 10 year sweep. This means that we lose 753 cases of the total 9,841 in the age 42 year sweep, giving us an analysis sample of 9,078 individuals. The cases lost include those not responding at age 42 years, immigrants to the UK since the age of 10 years, and individuals not present in the age 10 year survey who then responded to subsequent sweeps.[7]

Table 1 describes the social mobility that help in the labour market may partly explain. It shows the distribution of cohort members' socio-economic group (SEG) at age 42 years by parents' SEG at age 10 years. Reading across a row gives the percentages of children from the age 10 years parental SEG who are in each SEG at age 42 years. The SEG classifications at the two ages are not identical despite the categories having the same names, but they are sufficiently indicative for our purposes.[8] SEG is missing for 9% of the sample at age 10 years and 16% at age 42 years (the latter includes a small number classified as 'other') but we include these groups in the table as an additional row and column respectively.

The figures in brackets in the final row and column show the distribution for the full sample at each age. These show that, on average, there has been a clear upwards shift in the occupational levels of the children relative to those of their parents, reflecting the decline in blue collar jobs noted earlier. For example, there are many fewer cohort members in the 'skilled manual' group at age 42 years than at age 10 years (14% compared to 37%), while the opposite is true for the managerial-technical group where there has been a big rise in its importance (from 23% to 38%). The percentages in the interior of the table show the transition probabilities. For example, 68% of children with professional parents at age 10 years were either in the professional group themselves at age 42 years or were in the managerial-technical group. The figures for children with unskilled, partly-skilled and skilled-manual parents reaching the professional or managerial technical level were only 25%, 33% and 36% respectively. The transition probabilities differ somewhat for men and women. For example, while 75% of men with professional parents were in the professional /technical group themselves at age 42 years, this is the case for only 62% of women (not shown).[9]

Table 1. Parental SEG when cohort member aged 10 years and own SEG at age 42 years (row percentages)

	Age 42 own SEG								
Age 10 parental SEG	Unskilled	Partly Skilled	Skilled-Manual	Skilled N-Man	Manage-Tech	Prof.	Missing	Total	(All)
Unskilled	3.5	15.3	14.7	16.6	22.6	1.9	25.5	100.0	(3.5)
Partly Skilled	2.1	13.5	17.3	17.1	31.2	2.2	16.7	100.0	(11.6)
Skilled Manual	2.2	10.8	18.0	16.6	32.3	3.7	16.5	100.0	(36.8)
Sk. Non-Man.	1.1	8.0	11.0	18.2	41.8	5.8	14.0	100.0	(10.3)
Manage-Tech.	0.8	7.6	10.5	14.1	46.4	7.4	13.2	100.0	(23.3)
Professional	0.4	4.8	4.4	12.0	52.3	15.8	10.2	100.0	(5.7)
Missing	2.6	12.4	14.9	11.8	35.2	4.1	18.9	100.0	(8.8)
(All)	(1.7)	(10.1)	(14.3)	(15.5)	(37.5)	(5.3)	(15.7)	(100.0)	(100.0)

Notes. n=9,078. The SEG categories are V, IV, IIIM, IIINM, II and I and refer to the Registrar General's Social Groups classification. The parental SEG is the father's occupation and if missing the mother's.

Blanden, Gregg and Macmillan (2013) used the same data set to examine the relationship between age 10 years parental SEG and the individual's own SEG at age 33 years, some ten years earlier than our data. Although the SEG categories are not completely consistent due to differences in the classification of socio-economic groups, Blanden et al (2013) found that 70% of children with professional parents at age 10 years (category 7, Table 4) were either in the professional or the lower managerial groups by age 33 years (categories 6 and 7, Table 4). Thus when we consider these individuals further on in their careers at age 42 years, we continue to find that the vast majority of children from professional backgrounds also achieve higher level occupations.

Who gets help in the job market?

Table 2 shows the percentages of our sample who report receiving help in the job market at some point in their working lives by type of help received, distinguishing between (i) parents and (ii) other relatives and friends. Just over half of men (55%) report receiving help from parents compared with just under half of women (47%). And about a half of both sexes (48%) report help from other relatives and friends. The reporting of parental and relative/friend help tend to go together although the overlap is far from complete: about 30% of the sample – a bit more for women and a bit less for men – report neither source of help. This figure indicates that the use of family and other networks is not ubiquitous, as proposed earlier, at least as measured in the BCS70 data.

Table 2. Prevalence of help from parents and other relatives/friends (cell percentages)

a) Men

Help from parents	*Help from other relatives or friends*		
	No	Yes	Total
No	28.5	16.2	44.7
Yes	23.2	32.1	55.3
Total	51.7	48.3	100.0

b) Women

Help from parents	*Help from other relatives or friends*		
	No	Yes	Total
No	33.2	19.6	52.8
Yes	18.7	28.5	47.2
Total	51.9	48.1	100.0

What types of help did people receive? As already noted, some forms of help are potentially more worrying from a social mobility perspective than others. Table 3 is restricted to those individuals who do report help of some sort. The percentages sum to more than 100 as more than one type can be received. Advice is easily the most common source of help from parents, reported by over two-thirds of men and nearly three-quarters of women. The second most common assistance from parents is help with application forms, which again features more prominently for women. Next comes help through contacts, reported by 29% of men getting any help from parents and 21% of women. This is a form of parental support that we have argued should be seen in a different light to just having advice. Given that about half of respondents report parental help of any type, these figures imply that about 1 in 6 of all men and 1 in 10 of all women say they have been helped by parents in this way. Recommendation to a former or current employer, which we have argued is another type of help that is particularly important to consider, is next in importance: 19% of men and 15% of women reporting any help.

The pattern is different for help from other relatives and friends. Advice is still a very prominent source of help but is reported by less than half of both men and women. Acting as a referee is almost as common as advice for men and equally so for women. Recommendation to a current or former employer is reported by about 1 in 3 people getting any help, much more than for help from parents, which reflects the fact that work and social life intersect for many people.

Table 3. Percentage receiving each type of help among those receiving any help

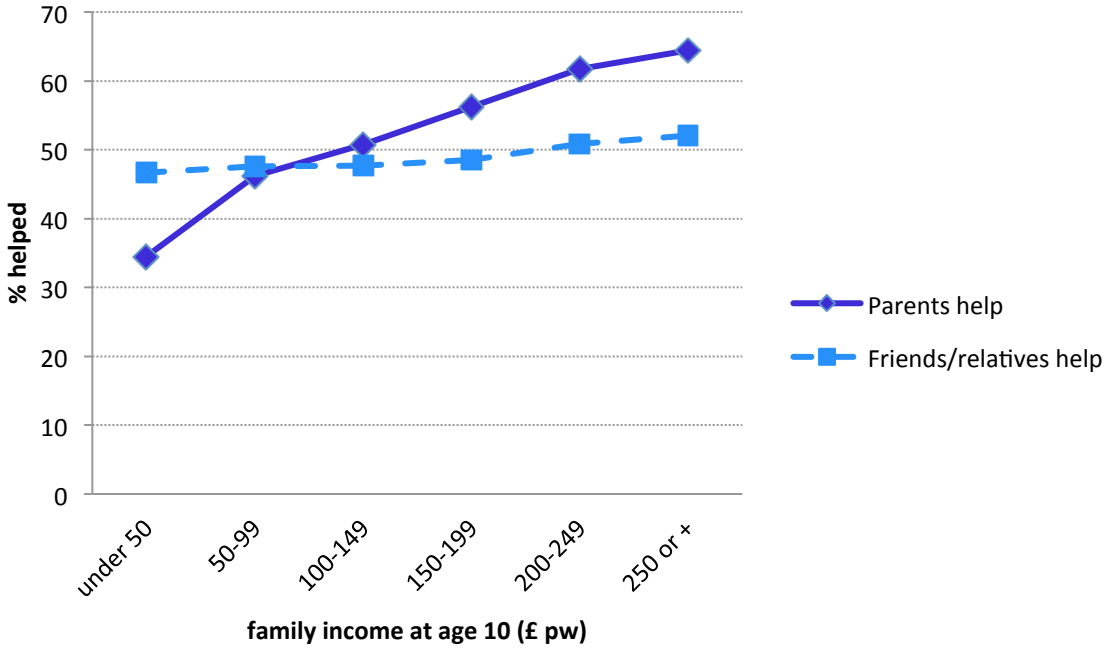
Type of help	Men		Women	
	Parents	*Other relatives/ friends*	*Parents*	*Other relatives/ friends*
Provided advice	67.9	40.6	73.0	42.1
Help with application form	31.9	14.8	38.8	25.3
Acted as a referee	12.6	35.2	7.8	42.6
Recommended to current/former employer	18.9	37.0	14.6	27.4
Recommended to another employer	9.7	13.4	5.3	7.7
Directly employed me	16.7	14.0	10.4	8.4
Helped find job through contacts	29.3	25.4	20.6	15.8
Financial	0.5	0.3	0.2	0.1
Other	0.8	0.6	1.1	0.8
Missing	0.5	0.6	0.5	0.4

Notes. Sample sizes in the four columns are 2,408, 2,105, 2,229, and 2,270. Percentages do not sum to 100 as more than one type of help may have been received. The base for the calculation in each case is all persons receiving any type of help.

We now show how the probability of receiving help varies with parental SES. We first consider income. Figure 1 shows the percentages of the sample receiving parental help and help from other relatives and friends of any kind by categories of total net family income in £ per week at age 10 years.[10] Information on income is missing for 13.4% of the sample. Those missing income information are similar, in terms of the help they receive, to those with complete income data[11]. The income groups are very unequal in size – the intervals £50-£99, £100-149, and £150-199 contain about 80% of the sample.

The probability of reporting help from parents rises substantially with income, from 34% in the bottom group of under £50 to 64% in the top group of £250 or more (each of these two groups contains about 6% of the sample).[12] But there is a much flatter relationship for help from other relatives and friends with only a modest rise across the income distribution from 47% to 52%.

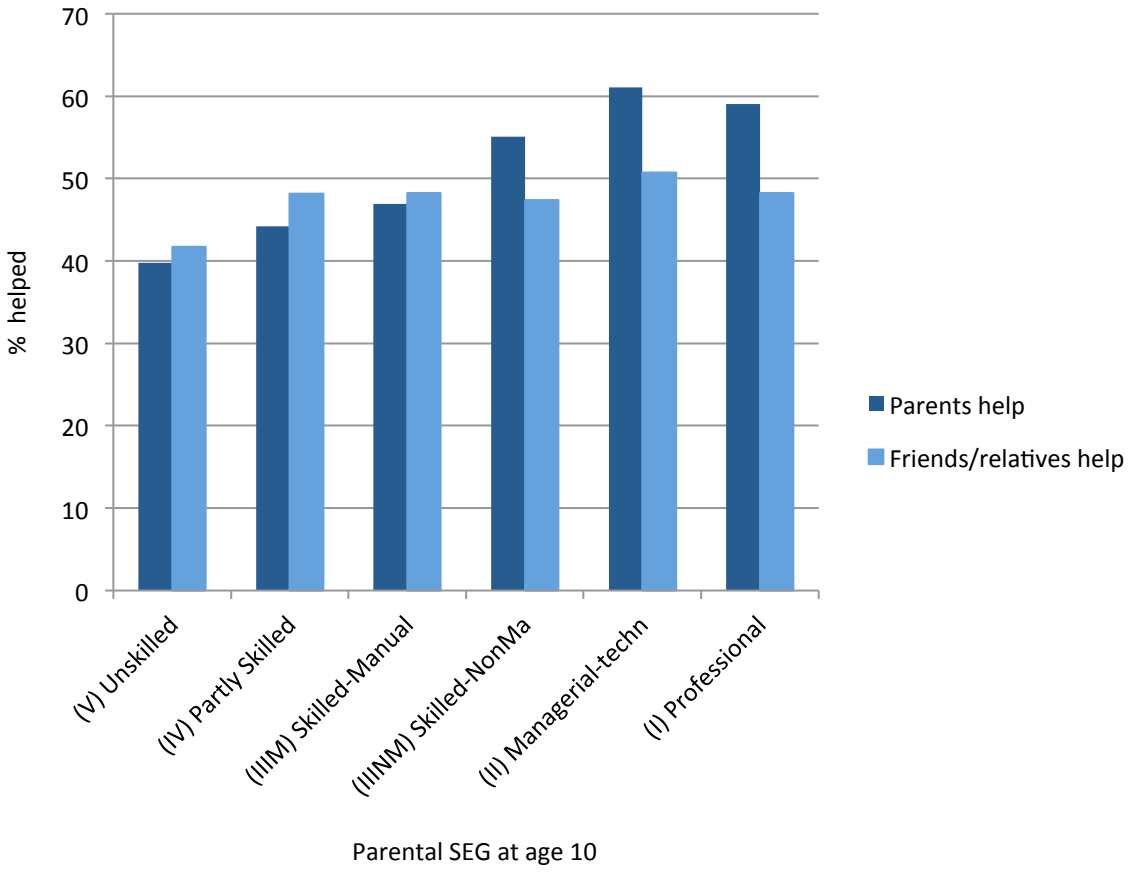
Figure 1. Percentage of cohort members receiving help from parents and from other relatives or friends by family income at age 10 years



Note. n = 7,862

Figure 2 shows results for our other measure of SES at age 10 years, parental socio-economic group. A similar pattern holds to that for income, at least in terms of the sign of the SES gradient. There is a higher probability of receiving parental help in higher SEGs: it is 40% and 44% in the bottom two groups and 61% and 60% in the top two groups. But the probability of receiving help from other relatives and friends differs little between different occupational groups.[13] Broadly speaking therefore, the picture obtained of the social gradient in both types of help is robust to the SES measure used. There is little difference in the pattern by gender with either measure.

Figure 2. Percentage of cohort members receiving help from parents and from other relatives or friends by parental SEG at age 10 years



Note. n = 8,280

We now probe the social gradient for different types of help, restricting the analysis to help from parents as we believe it is likely to be better measured. We take parental SEG as our measure of family background but the findings are similar using family income. Table 4 shows the percentages reporting (i) advice, (ii) help with application forms, (iii) recommendation by a parent to a current or former employer, and (iv) help to find a job through a parental contact ('people that they know'). In contrast to Table 3, we now take the base for the calculations as all individuals in the parental SEG group concerned, whether they had any help or not: the figures show the percentage of all individuals in each group receiving that particular kind of help.

The results for advice are not surprising – this is the most common form of help from parents and, since it is also socially graded, it is a major driver in the SEG gradient already seen in Figure 2. A quarter of the sample with unskilled parents reports advice from them, rising to a half for those with at least one professional parent. The probability of help with applications – the second most frequent form of parental assistance – also rises significantly with parental SEG. But it is notable that neither help in finding a job through contacts nor a recommendation to a current or former employer appears strongly associated with higher parental SEG. Broadly speaking, these forms of parental help seem equally common for people from different family backgrounds (there is just a small increase in the prevalence of use of contacts in managerial and professional families.)

Table 4. Percentage of cohort members receiving help from parents of various types by parental SEG at age 10 years

Parental SEG at age 10:	*Provided advice*	*Help with applications*	*Recommend to current/ former employer*	*Helped find job through contacts*
Unskilled	24.8	12.7	6.7	11.8
Partly Skilled	29.2	14.8	8.7	11.7
Skilled Manual	31.2	14.8	8.8	12.2
Skilled Non-Manual	39.9	21.5	8.7	12.3
Managerial-Technical	45.6	23.3	9.1	15.6
Professional	48.8	26.6	7.1	14.3

Notes. n=9,078. The values for the group with parental SEG missing are 30.7%, 13.9%, 7.8% and 9.4%.

What is the value of the help received from parents?

Thus far we have simply considered the relationship between whether a person received help from their parents and their socio-economic background. We have shown that some but not all types of help provided by parents are socially graded. Of course even if all parents provide help to their children in a number of different ways, such help may not be equally valuable. Parents who are wealthier and have higher occupational status may be more able to provide the kind of help that leads to greater economic success. Whilst we do have information on whether the individuals themselves think the help was useful, this is not necessarily a good way to determine whether more advantaged parents provide the kind of help that leads to greater labour market success. An individual may have found their parents' help useful, perhaps in securing them a job, but it does not necessarily mean that this enabled them to make as much career progress as the help provided by wealthier parents. To judge the average effectiveness of the help provided by parents we estimate the impact of receiving that help on individuals' eventual labour market success at age 42 years, measured by wages and by occupational group.

We start, however, by first considering the subjective valuation by the cohort member: the question on 'how much has the help to get a job contributed towards your current occupation or career?'. As discussed in the first section, the information generated from this question is limited by the fact that although we know whether the help provided was deemed useful, we cannot be sure exactly what type of help the individual was referring to. Table 5 shows the distribution of answers for those receiving any help by parental SEG at age 10 years. Broadly speaking, a bit more than a third of the sample views the help they received as having contributed not at all, a third think it contributed a little, and a bit less than a third see it as having been contributed a lot. The figures vary little by parental SEG: for example the percentage responding 'a lot' is 27% for those with partly skilled manual parents and 25% for those with professional parents.

Table 5. Own valuation of parental help by parental SEG at age 10 years for all receiving any help (row percentages)

	Contribution of help to current occupation or career			
Parental SEG age 10 years	*None*	*A little*	*A lot*	*Total*
Unskilled	36.8	25.6	37.6	100.0
Partly Skilled	38.2	34.1	27.7	100.0
Skilled Manual	37.7	30.5	31.8	100.0
Skilled Non-Manual	37.2	35.0	27.8	100.0
Managerial-Technical	35.3	33.4	31.3	100.0
Professional	38.9	36.0	25.1	100.0
Missing	39.3	28.5	32.2	100.0
All	37.2	32.3	30.5	100.0

Notes. n=9,078

We now attempt an objective valuation of the help provided. We estimate a wage regression, where the dependent variable is log of the gross monthly wage of the individual at age 42 years, and we relate this to an individual's socio-economic background, their years of education and their ability, as measured by maths test score at age 10 years (standardised to a mean of zero and standard deviation of one). To ensure a parsimonious specification, to ease interpretation and to avoid multi-colinearity between different measures of parental socio-economic background, we only include parental SEG at age 10 years (results are robust to using parental income instead). We estimate the model separately for men and women. To this standard wage equation we then add terms indicating whether the individual received a particular kind of help from their parent. We interpret a positive significant coefficient on the help variables as indicative that the help provided had some labour market value. Any impact of parental SEG on wages that comes through the child achieving more years in education or having higher ability at age 10 years (for a discussion of the mechanisms see e.g. Haveman & Wolfe, 1995) will be picked up by the coefficients on those variables – we wish to measure only the impact of parental help in the labour market over and above that coming through education and ability. We interpret any significant reduction in the magnitude of the coefficients on the parental SEG variables once the network variables are included, as indicative evidence that parental networks and help are one route through which individuals from more socio-economically advantaged backgrounds secure labour market advantage. We then consider whether similar findings hold when we use SEG at 42 years as opposed to earnings as the dependent variable. This latter model also provides a rather larger sample size – we have 6,131 observations in the wage regressions but 7,657 in the models of SEG.

Our work comes with caveats. First, we cannot interpret the coefficient on parental help as a causal parameter. Clearly it may be that individuals who are likely to do well in the labour market are also more likely to secure assistance from their parents. We therefore note that our results are associative only. Second, our estimates will suffer from bias if the help variables are measured with error, either due to recall bias or because they are not posed in such a way as to obtain accurate information about how the individual was assisted by his or her parents. We cannot determine the direction of the bias since the measurement error may not be 'classical' in form (i.e. independent of the true measure) if some types of individuals are more or less likely to misreport the help they received, e.g. those who come from advantaged backgrounds or those who secure very good jobs.

Table 6 reports the results from the wage regression for men and Table 7 for women. The model in column 1 includes background variables only, including parental SEG. Column 2 reports on a model including just a single dummy variable indicating that any help from parents was received. Column 3 adds additional controls to distinguish the type of help reported.

Table 6. The relationship between parental help and age 42 years log gross monthly wages: regression coefficients for men

	1	2	3
Years of education	0.046***	0.046***	0.045***
	(0.005)	(0.005)	(0.005)
Maths score aged 10 years (std.)	0.102***	0.102***	0.102***
	(0.017)	(0.017)	(0.017)
Maths score missing	-0.007	-0.008	-0.008
	(0.037)	(0.037)	(0.037)
Professional	0.254***	0.252***	0.253***
	(0.062)	(0.063)	(0.063)
Managerial-technical	0.171***	0.168***	0.169***
	(0.040)	(0.040)	(0.040)
Skilled Non-Manual	0.161***	0.159***	0.159***
	(0.050)	(0.050)	(0.050)
Partly Skilled	-0.100**	-0.100**	-0.098**
	(0.050)	(0.050)	(0.050)
Unskilled	-0.075	-0.073	-0.072
	(0.085)	(0.085)	(0.085)
SEG Missing	0.104*	0.104*	0.102*
	(0.057)	(0.057)	(0.057)
Advice			-0.152
			(0.357)
Help with application form			-0.170
			(0.358)
Referee			-0.205
			(0.364)
Recommended to current/former			-0.188
			(0.361)
Recommended to another employer			-0.147
			(0.373)
Directly employed me			-0.165
			(0.360)
Helped find job through contacts			-0.207
			(0.357)
Financial			
Other			-0.326
			(0.418)
Parents help typemissing			
Receive any parental help		0.023	-0.124
		(0.030)	(0.221)
Constant	7.311***	7.299***	7.626***
	(0.064)	(0.066)	(0.424)
n	2,966	2,966	2,966
R-squared	0.089	0.090	0.090

*Notes. Standard errors in brackets. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively. The excluded category of parental SEG at age 10 years is Skilled Manual.*

Table 7. The relationship between parental help and age 42 years log gross monthly wages: regression coefficients for women

	1	2	3
Years of education	0.059***	0.059***	0.059***
	(0.005)	(0.005)	(0.005)
Maths score aged 10 years (std.)	0.162***	0.162***	0.162***
	(0.018)	(0.018)	(0.018)
Maths score missing	0.037	0.038	0.038
	(0.038)	(0.038)	(0.038)
Professional	0.181***	0.176**	0.178**
	(0.069)	(0.070)	(0.070)
Managerial-Technical	0.051	0.046	0.046
	(0.040)	(0.041)	(0.041)
Skilled Non-Manual	0.052	0.049	0.051
	(0.053)	(0.053)	(0.053)
Partly Skilled	-0.063	-0.061	-0.058
	(0.049)	(0.049)	(0.049)
Unskilled	-0.084	-0.082	-0.088
	(0.090)	(0.090)	(0.090)
SEG Missing	-0.046	-0.045	-0.046
	(0.058)	(0.058)	(0.058)
Advice			0.051
			(0.597)
Help with application form			-0.018
			(0.597)
Referee			-0.201
			(0.606)
Recommended to current/former			0.017
			(0.600)
Recommended to another employer			0.184
			(0.611)
Directly employed me			-0.030
			(0.601)
Helped find job through contacts			0.026
			(0.597)
Financial			-
Other			0.006
			(0.634)
Parents help type missing			0.264
			(0.730)
Received any parental help		0.038	0.030
		(0.030)	(0.218)
Constant	6.474***	6.460***	6.459***
	(0.069)	(0.069)	(0.637)
n	3,187	3,187	3,187
R-squared	0.100	0.100	0.102

*Notes. Standard errors in brackets. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively. The excluded category of parental SEG at age 10 years is Skilled Manual. Financial help contains no observations.*

The results for men in column 1 show a substantial association of wages with parental SEG, conditional on years of education and ability at age 10 years: having a professional parent is associated with wages at age 42 years that are about 28% higher on average than if the individual had a skilled manual parent (the excluded base category)[14]. A managerial-technical parent is associated with an average 19% wage premium and a skilled non-manual parent 17%. By contrast, for women Table 7 generally shows no statistically significant association of wages with parental SEG except that women with professional parents have earnings that are on average 19% higher than women with a skilled manual parent, conditional on the education and ability measures, which themselves have estimated coefficients that are notably larger than for the men (by between 20% and 60%). For women, the association of parental occupation at age 10 years and offspring wages at age 42 years appears largely accounted for by the measures of years of education and ability.

The dummy variable in column 2 in each table is statistically insignificant – we find no evidence for either men or women of higher average wages for those reporting parental help of any type. Column 3 then reports whether there are any additional wage premia associated with the different types of help. This specification seeks to determine whether, even if there is no overall impact from receiving help, perhaps some specific types of help might be associated with greater labour market success. Again however, we find no significant impact of any type of help on wages at age 42 years for men or for women. [15]

Table 8 contains the results of a specification using the subjective valuation of parental help documented earlier in Table 5 and considering its association with wages. The table reports results for both men and women, and their comparison underlines the comments made above about the different impact of the parental SEG measures and the control variables by gender. The results provide some validation of the subjective valuations and at the same time are consistent with the explanation of the insignificance of the single dummy for parental help of any type in Tables 6 and 7. Parental help that is reported as having contributed 'a lot' to the individual's career or current occupation is associated with an approximately 10% average wage premium for both men and women, although the coefficients are not that well determined (t=2.5 for men and 2.6 for women). Help that contributed 'not at all' has a negative but insignificant association, while help that contributed 'a little' has a positive but insignificant association. The significant parental SEG coefficients are practically unchanged for men from those in Table 6 – the parental help that contributes 'a lot' or 'a little' does not help explain the association of higher parental occupations with wages.

Table 8. Self-valuation of parental help and its association with age 42 years log gross monthly wages: regression coefficients

	Men	Women
Years of education	0.046***	0.058***
	(0.005)	(0.005)
Maths score aged 10 years (std.)	0.104***	0.164***
	(0.017)	(0.018)
Maths score missing	-0.003	0.038
	(0.037)	(0.038)
Professional	0.252***	0.182***
	(0.063)	(0.069)
Managerial-Technical	0.165***	0.046
	(0.040)	(0.041)
Skilled Non-Manual	0.159***	0.051
	(0.050)	(0.053)
Partly Skilled	-0.096*	-0.059
	(0.050)	(0.049)
Unskilled	-0.068	-0.085
	(0.085)	(0.089)
SEG missing	0.108*	-0.044
	(0.057)	(0.058)
Help contributed a lot	0.100**	0.117***
	(0.042)	(0.045)
Help contributed a little	0.048	0.074*
	(0.040)	(0.043)
Help contributed not at all	-0.049	-0.055
	(0.039)	(0.042)
Constant	7.288***	6.465***
	(0.066)	(0.069)
n	2,966	3,187
R-squared	0.091	0.103

*Notes. Standard errors in brackets. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively. The excluded category of parental SEG at age 10 years is Skilled Manual.*

In Tables 9-11 we repeat the exercise for the individual's SEG at age 42 years as the outcome of interest. [16] We again use OLS regression despite the ordinal nature of the dependent variable (coding unskilled as 1 through to professional as 6 and dropping cases with missing SEG). Column 1 in Table 9 (men) and Table 10 (women) contains the parental SEG variables but no controls – the results here measure all direct and indirect association of parental and offspring SEG, just as in the transition matrix reported earlier in Table 5. In column 2 we add the education and ability variables, in column 3 the single parental help variable, and then adding additional type of help variables in column 4. Table 11 is the analogue of Table 8, with results for both men and women of a model with three dummy variables for subjective valuation of any help received.

Table 9. The relationship between parental help and age 42 years SEG: regression coefficients for men

	1	2	3	4
Years of education		0.121***	0.121***	0.121***
		(0.006)	(0.006)	(0.006)
Maths score aged 10 years (std.)		0.258***	0.259***	0.256***
		(0.024)	(0.024)	(0.024)
Maths score missing		-0.032	-0.032	-0.035
		(0.045)	(0.045)	(0.045)
Professional	0.955***	0.460***	0.462***	0.451***
	(0.083)	(0.080)	(0.080)	(0.080)
Managerial-Technical	0.540***	0.299***	0.302***	0.300***
	(0.051)	(0.048)	(0.048)	(0.048)
Skilled Non-Manual	0.483***	0.296***	0.298***	0.296***
	(0.066)	(0.062)	(0.062)	(0.062)
Partly Skilled	-0.119*	-0.067	-0.068	-0.065
	(0.066)	(0.061)	(0.061)	(0.061)
Unskilled	-0.369***	-0.263**	-0.264**	-0.260**
	(0.114)	(0.106)	(0.106)	(0.106)
SEG missing	0.156**	0.095	0.095	0.090
	(0.073)	(0.068)	(0.068)	(0.068)
Advice				0.040
				(0.165)
Help with application form				-0.042
				(0.168)
Referee				0.087
				(0.184)
Recommended to current/former				-0.107
				(0.175)
Recommended to another employer				-0.176
				(0.198)
Directly employed me				0.089
				(0.170)
Helped find job through contacts				-0.164
				(0.164)
Financial				0.567*
				(0.335)
Other				-0.053
				(0.271)
Parents help type missing				-1.646*
				(0.990)
Received any parental help			-0.024	-0.034
			(0.036)	(0.249)
Constant	3.865***	2.497***	2.510***	2.562***
	(0.032)	(0.078)	(0.081)	(0.282)
n	3,890	3,877	3,877	3,877
R-squared	0.070	0.199	0.199	0.205

*Notes. Standard errors in brackets. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively. The excluded category of parental SEG at age 10 years is Skilled Manual.*

Table 10. The relationship between parental help and age 42 years SEG: regression coefficients for women

	1	2	3	4
Years of education		0.093***	0.093***	0.092***
		(0.007)	(0.007)	(0.007)
Maths score aged 10 years (std.)		0.255***	0.255***	0.253***
		(0.025)	(0.025)	(0.025)
Maths score missing		-0.011	-0.011	-0.008
		(0.048)	(0.048)	(0.048)
Professional	0.727***	0.315***	0.309***	0.301***
	(0.087)	(0.086)	(0.086)	(0.086)
Managerial-Technical	0.380***	0.147***	0.139***	0.132***
	(0.051)	(0.050)	(0.051)	(0.051)
Skilled Non-Manual	0.201***	0.067	0.062	0.052
	(0.069)	(0.067)	(0.067)	(0.067)
Partly Skilled	-0.089	-0.063	-0.062	-0.060
	(0.065)	(0.062)	(0.062)	(0.062)
Unskilled	-0.257**	-0.108	-0.103	-0.097
	(0.116)	(0.111)	(0.111)	(0.111)
SEG missing	-0.107	-0.137*	-0.134*	-0.137*
	(0.075)	(0.072)	(0.072)	(0.072)
Advice				-0.275
				(0.573)
Help with application form				-0.219
				(0.574)
Referee				-0.527
				(0.588)
Recommended to current/former				-0.430
				(0.579)
Recommended to another employer				-0.155
				(0.597)
Directly employed me				-0.252
				(0.579)
Helped find job through contacts				-0.432
				(0.575)
Financial				-
Other				-0.583
				(0.632)
Parents help type missing				-0.829
				(0.809)
Received any parental help			0.055	-0.212
			(0.038)	(0.271)
Constant	3.979***	2.951***	2.930***	3.522***
	(0.032)	(0.085)	(0.087)	(0.638)
n	3,767	3,759	3,759	3,759
R-squared	0.039	0.126	0.127	0.130

*Notes. Standard errors in brackets. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively. The excluded category of parental SEG at age 10 years is Skilled Manual. Financial help contains no observations.*

Table 11. Self-valuation of parental help and its association with age 42 years SEG: regression coefficients

	Men	Women
Years of education	0.120***	0.093***
	(0.006)	(0.007)
Maths score aged 10 years (std.)	0.257***	0.256***
	(0.024)	(0.025)
Maths score missing	-0.030	-0.012
	(0.045)	(0.048)
Professional	0.458***	0.311***
	(0.080)	(0.086)
Managerial-Technical	0.301***	0.138***
	(0.048)	(0.051)
Skilled Non-Manual	0.296***	0.065
	(0.062)	(0.067)
Partly Skilled	-0.072	-0.060
	(0.061)	(0.062)
Unskilled	-0.263**	-0.105
	(0.106)	(0.111)
SEG missing	0.096	-0.137*
	(0.068)	(0.072)
Help contributed a lot	-0.057	0.142**
	(0.050)	(0.056)
Help contributed a little	0.035	0.066
	(0.050)	(0.054)
Help contributed not at all	-0.024	-0.024
	(0.048)	(0.052)
Constant	2.513***	2.930***
	(0.081)	(0.087)
n	3,877	3,759
R-squared	0.200	0.128

*Notes. Standard errors in brackets. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively. The excluded category of parental SEG at age 10 years is Skilled Manual.*

The results suggest a strong relationship between parental SEG and the cohort member's own SEG for both men and women. Much of this correlation is explained by the fact that higher parental SEG is associated with the cohort member having more education and higher levels of ability. However, once these factors are controlled in the model (column 2), a strong relationship between parental SEG and an individual's SEG at age 42 years persists, particularly for men. This strong link between parental SEG and own SEG is therefore not entirely explained by higher levels of education and ability. We then explore whether parental assistance can explain this strong relationship. However, just as in the wage models above, we find that individuals who report receiving help from their parents do not achieve higher occupational status in the labour market at age 42 years (column 3). The use of parental support does not therefore explain the strong residual relationship we observe between parental and own SEG.

In contrast to the results for wages, for men there is no indication in Table 11 that the individual's subjective valuation that help received has contributed 'a lot' to the SEG attained at age 42 years, indeed the estimated coefficient is negative albeit insignificant. But for women the same pattern is found of help that is reported as important being associated with a statistically significant increase in the value of the outcome variable (t=2.5), controlling for other factors. The estimated coefficient (0.142) may seem modest in size when viewed against the range of the dependent variable (1 to 6) but is similar in magnitude to that on the dummy variable for a parent in the

managerial-technical group when the cohort member was aged 10 years.

Conclusions

We have produced new evidence from the 1970 British Cohort Study on the prevalence of help in the job market from parents, other relatives and friends. In doing so, we also report new estimates of the extent of social mobility in Britain today, specifically on the strong association between parental socio-economic group when cohort members were aged 10 years and their own socio-economic group and earnings at aged 42 years. Our findings are stark: a person's socio-economic background continues to have a persistent and large impact on their eventual occupation and earnings. Much of this impact is via the effect of parental SEG on an individual's ability and education. However, beyond that it remains true for this cohort that parental SEG is strongly associated with own SEG, even after allowing for the impact of family background on ability and education. Inter-generational social mobility continues to be constrained in England, consistent with the large body of economic and sociological literature that has documented this for earlier cohorts.

Our focus, however, was on an important route by which parental SEG might influence an individual's labour market success: through use of help, networks and specific assistance in the labour market from parents. Help in the job market is widespread – about a half of the cohort report help from parents and about a half report help from other relatives and friends. Advice is easily the most widely reported type of help, especially from parents, but a substantial proportion of respondents report that parents helped them find a job through their contacts (29% of men receiving any help and 21% of women) or that other relatives or friends helped in this way (25% of men and 16% of women). Recommendation to a current or former employer is also quite widely reported.

Parental help has a clear social gradient, being more common for individuals with parents from higher socio-economic groups and higher levels of family incomes, while help from other relatives and friends has little association with family socio-economic background. However, it is help in the form of advice and assistance with application forms that drives this social gradient of parental help. Help from parents in the form of recommendation to a current or former employer or in finding a job through their contacts has little apparent relationship with parental socio-economic status.

Our results indicate that individuals who report receiving help from their parents do not have higher earnings than those who do not receive such help, nor do they have higher status occupations. It would seem therefore that parental help is not necessarily beneficial in the labour market, and certainly parental help cannot explain the very strong relationship between parental SEG and own SEG. One might imagine this is because some forms of help are more beneficial than others. About a third of people receiving parental help report that they believe it contributed a lot to their career or current occupation. Individuals who report help that contributed a lot, do appear to have significantly higher wages at age 42 years, controlling for other factors. This is the case for both men and women but only for women could we also find an association of parental help of this type with the individual's occupation (defined in broad categories) at age 42 years. For both genders and for both wages and occupations, help viewed by the respondent as having made no contribution to their career or occupation was confirmed as having no association with their wage or broad occupation. The implication here is that not all help is valuable but where help is perceived as having value by the cohort member (CM), the CM is more likely to have higher earnings.

In conclusion we were unable to identify a clear causal link between any particular type of help – advice, help through contacts etc. – and CMs' wages or occupations. This suggests that other factors and mechanisms must explain the strong intergenerational relationship between parent and CM wages and occupational status. This is an important finding, given that the work also highlights the persistent strong relationship between parental occupation and the CM's own occupation, particularly for males. It is striking that the role of parental SEG in predicting future labour market success is far stronger for males and also goes beyond an impact from parental SEG on the CM's ability and education.

Future work could focus on two features of our findings. First, we need better measures of both the extent to which a parent is supporting their child in

the labour market, and the form of that help. All our results are conditional on the quality of the measures of help collected in the age 42 year sweep of BCS70. Indeed, error in these measures could explain the insignificant relationship between help and labour market outcomes. We discussed the problems with our measures, including the long period of recall. We recommend that cohort studies ask specifically about the *current* job and that information is sought at each wave of data collection. We also see much merit in asking more specific questions to elicit clearer indicators of the extent of help a parent provides. For example, one might ask whether the individual works (or has worked) for the same employer as a parent, although we recognise that there are important definitional issues to overcome (e.g. is the entire public sector one 'employer'?).

Second, more research is needed to explore differences in the way that parental SEG impacts on men and women's eventual labour market success. We have noted that for women the association of parental occupation at age 10 years and the CM wages at age 42 years appears largely accounted for by the CM's ability and years of education – but this is not the case for men. There are differences also when the outcome variable is the age 42 years SEG. We need further research to know whether these gender differences are peculiar to Britain, how they have changed over time, and – more fundamentally – why they occur.

Acknowledgements

We are grateful to Alice Sullivan and Matt Brown for their help and to the referees and editor for their comments on an earlier version.

References

Battu, H., Seaman, P., & Zenou, Y. (2011). Job contact networks and the ethnic minorities. *Labour Economics* 18(1), 48-56. http://dx.doi.org/10.1016/j.labeco.2010.07.001

Bingley, P., Corak, M., & Westergård-Nielsen, N. (2011). *The intergenerational transmission of employers in Canada and Denmark*. IZA Discussion Paper No. 5593.

Blanden, J., Goodman, A., Gregg, P., & Machin, S. (2004). Changes in intergenerational mobility in Britain. In Corak, M. (Ed.), (pp. 121-146). Cambridge: Cambridge University Press.

Blanden, J., Gregg, P., & Macmillan, L. (2013). Intergenerational persistence in income and social class: the impact of within-group inequality. *Journal of the Royal Statistical Society Series A: Statistics in Society, 176 (2),* 541-563. http://dx.doi.org/10.1111/j.1467-985X.2012.01053.x

Bukodi, E., & Goldthorpe, J. H. (2011). Social class returns to higher education: chances of access to the professional and managerial salariat for men in three British birth cohorts. *Longitudinal and Life Course Studies, 2(2)*, 1-71.

Cabinet Office (2009). *Unleashing aspirations: the final report of the panel on Fair Access to the Professions*. London: Information Policy Team.

Cabinet Office (2011). *Opening doors, breaking barriers: a strategy for social mobility.* London: Information Policy Team.

Cabinet Office (2012*). Fair access to professional careers: a progress report by the independent reviewer on social mobility and child poverty*. London: Information Policy Team.

Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: Harvard University Press.

Corak, M., & Piraino, P. (2010). *The intergenerational transmission of employers*. IZA Discussion Paper No. 4819.

Davern, M. (1997). Social networks and economic sociology. *American Journal of Economics and Sociology*, 56: 287–302. http://dx.doi.org/10.1111/j.1536-7150.1997.tb03359.x

Erikson, R., & Goldthorpe, J. (2010). Has social mobility in Britain decreased? Reconciling divergent findings on income and class mobility. *British Journal of Sociology, 61(2)*, 211-230. http://dx.doi.org/10.1111/j.1468-4446.2010.01310.x

Ermisch, J., Jantti, M., & Smeeding., T. (2012). *The intergenerational transmissions of advantage.* New York: Russell Sage Foundation.

Fernandez, R. M., & Castilla, E. J. (2001). How much is that network worth? Social capital in employee referral networks. *Social Capital: Theory and Research*, 85-104.

Goldthorpe, J. (2013). Understanding – and misunderstanding – social mobility in Britain: the entry of the economists, the confusion of politicians and the limits of educational policy. *Journal of Social Policy*, *42(3)*, 431-450. http://dx.doi.org/10.1017/S004727941300024X

Goldthorpe, J., & Mills, C. (2008). Trends in intergenerational class mobility in modern Britain: evidence from national surveys, 1972-2005. *National Institute Economic Review, 205(1)*, 83-100. http://dx.doi.org/10.1177/0027950108096591

Gorard, S. (2008). A re-consideration of rates of social mobility in Britain: or why research impact is not always a good thing. *British Journal of Sociology of Education, 29(3)*, 317–24. http://dx.doi.org/10.1080/01425690801966402

Green, A.E., de Hoyos, M., Li, Y., & Owen, D. (2011). *Job search study: literature review and analysis of the Labour Force Survey*. London: Department for Work and Pensions Research Report No 726.

Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: a review of methods and findings. *Journal of Economic Literature*, *33(4)*, 1829-78.

Holzer, H. J. (1988). Search method use by unemployed youth. *Journal of Labour Economics, Vol. 6(1),* 1-20. http://dx.doi.org/10.1086/298172

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

Kramarz, F., & Skans, O. (2006). Nepotism at work? Family networks and youth labour market entry. Unpublished manuscript, Symposium of Labor Economics (ESSLE).

Langlands, A. (2005). *The gateways to the professions report ('The Langlands Report')*. Nottingham: DfES.

Loury, G. C. (1977). A dynamic theory of racial income differences. In P.A. Wallace & A. Le Mund (Eds.), *Women, minorities and employment discrimination,* (pp.153-188). Lexington, MA: Lexington Books.

Loury, L. (2006). Some contacts are more equal than others: informal networks, job tenure, and wages. *Journal of Labor Economics, 24 (2)*, 299-318. http://dx.doi.org/10.1086/499974

Marsden, P., & Gorman, E. (2001). Social networks, job changes, and recruitment. In I. Berg & A. L. Kalleberg (Eds.), Sourcebook of labor markets: evolving structure and processes. (pp. 467-502). New York: Kluwer Academic/Plenum Publishers. http://dx.doi.org/10.1007/978-1-4615-1225-7_19

Macmillan, L. (2009). *Social mobility and the professions*. Centre for Market and Public Organisation.

Macmillan, L., Tyler, C., & Vignoles, A. (2013). Who gets the top jobs? The role of family background and networks in recent graduates' access to high status professions. DoQSS Working Paper 13-15. London: Institute of Education.

Mostafa, T., & Wiggins, R. (2014). Handling attrition and non-response in the 1970 British Cohort Study. Centre for Longitudinal Studies Working Paper 2014/2, Institute of Education, London.

Sutton Trust (2005). The educational background of the UK's top solicitors, barristers and judges. London.

Sutton Trust (2006). The educational backgrounds of leading journalists. London.

Endnotes

[1] For earlier evidence, see e.g. Blanden, Goodman, Gregg, & Machin (2004), Blanden, Gregg & Mcmillan (2013), Bukodi and Goldthorpe (2011), Goldthorpe & Mills (2008), Gorard (2008), Erikson & Goldthorpe (2010) and Goldthorpe (2013).

[2] See Langlands (2005), Macmillan (2009), Macmillan, Tyler & Vignoles (2013), and Sutton Trust (2005, 2006) for examples of this literature.

[3] Mostafa and Wiggins (2014) discuss attrition in BCS70 to age 42 years. Not surprisingly they find attrition varies with individual characteristics: 'men from lower social backgrounds whose parents were single in 1970 are more likely to drop out' (p9). However, their logistic regression model of attrition to age 42 has very low predictive power (pseudo-r^2 of 0.025). It also cannot be estimated for the full cohort due to missing data in the first sweep of the cohort at birth. In this situation, as the authors note, the efficacy of any non-response weights based on this model of attrition is 'undermined' (p18). We therefore chose not to adjust for the probability of attrition.

[4] Asked only of those reporting any help. The question was not asked in respect of each type of help: the subjective valuation of 'the help that you received' therefore refers to all help of all types.

[5] We are grateful to Stuart Campbell (IOE) for providing this estimate.

[6] Battu et al. (2011) is an example. See also Green, de Hoyos, Li & Owen (2011) for a literature review of this topic.

[7] Immigrants who arrived between 1970 and 1980 may also not be present in our chosen sample as they were not all identified for the age 10 year sweep. We are grateful to Matt Brown for this information.

[8] The age 10 years classification refers to the 1970 version of the Registrar General's Social Groups classification and the age 42 years classification refers to the 1990 version.

[9] The t-statistic for the test of the difference is 3.2. Note however that all transition probabilities to higher level occupations are lower for women on account of the higher percentage with SEG missing at age 42 years: 20% compared to 11% for men.
[10] This is a measure of household earned and unearned income, consisting of the combined gross income of mother and father. Child benefits were not included but other sources of earned and unearned income were.
[11] Parental help was received in 51.3% of cases where age 10 years parental income is missing and help from other relatives and friends in 52.1% of cases.
[12] We have combined two intervals, under £35 and £35-£49, containing 122 cases and 360 cases respectively, to form the bottom group shown in the graph. The percentages receiving parental help differ sharply in these two groups but the imprecision in the estimate for the under £35 interval due to the small sample size led us to combine the groups as described.
[13] Parental help was received in 45.5% of cases where age 10 years parental SEG is missing and help from other relatives and friends in 44.6% of cases.
[14] To determine the actual association, bearing in mind that the dependent variable is in logs, we have taken the exponential (anti-log) of the coefficient. Thus, e.g., in the case of the variable "professional parent" the differential effect on wages is computed as "e^0.254- e^0", which provides the figure stated (28%).
[15] The column 3 coefficients for the type of help variable are still insignificant if we remove the dummy variable for any help.
[16] Results concerning the statistical significance of the help variables are robust to use of an ordinal logit model. For simplicity we present the OLS results here and the results from the logit model are available on request.

Longitudinal and Life Course Studies 2015 Volume 6 issue 2 Pp 212 – 227 *ISSN 1757-9597*

Religious involvement over the life course: problems of measurement and classification

David Voas Institute for Social and Economic Research, University of Essex
voas@essex.ac.uk

(Received June 2014 Revised February 2015) http://dx.doi.org/10.14301/llcs.v6i2.311

Abstract

Longitudinal studies have the potential to enhance our understanding of stability and change in religious identity, practice and belief. Good individual-level data would help in developing and testing theories concerning the causes and consequences of religious involvement. Past research has shown, however, that even subtle differences in wording or context can substantially affect responses to questions on religion. The 1970 British Cohort Study offers an important opportunity to test the consistency of self-reported religion and religiosity. In addition, the 2012 sweep asked questions on belief in God and life after death as well as religious affiliation and practice, allowing us to explore the complexity of religious adherence. A close examination of the multiple waves of the BCS70 reveals a large amount of uncertainty in measurement, making it hard to detect whatever genuine change might have occurred. There are indications of considerable unreliability in reported past and present affiliation. It is also difficult to be confident about changes in religious commitment, though a substantial proportion of teenagers who reported that religion was an important part of their lives became relatively unreligious adults. The data on religious belief make it apparent that while some people seem wholly non-religious and a smaller number are actively (and consistently) religious, the majority fall into intermediate categories defined by nominal allegiance, unorthodox belief, or belief in the absence of affiliation or practice. It is clear that multiple survey items covering identity, practice and belief are needed to obtain a reliable picture of religious commitment.

Keywords: religion, religiosity, generations, cohorts, affiliation, identity, belief

Why religion matters in life course research

The founders of the modern discipline of sociology were deeply interested in the relationship between modernization and religion; both Weber and Durkheim wrote extensively on the topic. Social scientists continue to try to explain the relative prevalence of religious involvement in different times and places. Researchers study continuity and change, religious growth and decline, the connection between believing and belonging, non-traditional spirituality, the influence of age, gender and socio-economic characteristics, inter-generational transmission, and much else besides (Beckford & Demerath, 2007; Clarke, 2009; Fenn, 2000; Turner, 2010). Panel data, and particularly birth cohort studies, can provide invaluable clues to whether and why individuals are susceptible to shifts in religious identity, practice and belief over the life course.

A lively debate over the secularization thesis (asserting that modernization of a society erodes the power, prestige and popularity of religion) has continued for decades (Bruce, 2011). Although the inverse relationship between socio-economic development and aggregate participation in religion is clear, opponents of the thesis point to the United States and the persistent religiosity of societies that are now modernizing as counter-examples (Stark &

Finke, 2000). While the debate is frequently depicted in black and white, much of the recent research uses various shades of grey. It is commonly conceded that traditional religious organizations and orthodoxies are struggling, but many scholars claim that religion is being transformed rather than eroded. Belief and practice, in their view, are individualized, possibly privatized, but still with us (Heelas & Woodhead, 2005). For secularization theorists, the issue is whether these new, less formal belief systems have the personal and social significance that characterized institutional religion.

A life course perspective is especially helpful in investigating the mechanisms of secularization. During childhood, religious identity, practice and belief are shaped by parents. How much effort parents invest in the religious socialization of their children is strongly associated with the importance they themselves attach to religion (Bengtson, 2013). As children emerge into adulthood, they may diverge from the family path; in the Western world, the motor of declining affiliation and attendance has been a drift away from religion at this point (Voas & Crockett, 2005). Within any given generation, the average level of religiosity appears to be fairly stable over the life course; contrary to common belief, there is little evidence that people become markedly more religious with age (Crockett & Voas, 2006; Dillon & Wink, 2007). Exceptions are most likely to occur as a result of family formation, when people use their own upbringing as a model for how children should be raised (Stolzenberg, Blair-Loy, & Waite, 1995).

Research on the causes of religious change is not the only reason to be interested in longitudinal data. Social scientists are more often concerned with the impact of religion on other aspects of life. In these studies religion is treated as an explanatory variable. Religion can affect health, drug and alcohol use, crime and deviance, fertility, mortality, marital stability, social and political attitudes, morality and altruism, tolerance, voting, civic participation, trust and social capital, education, economic activity, happiness and any number of other areas (Voas, 2007a). By far the largest literature is in health, where there are journals specifically dedicated to work in this field and thousands of articles have been published (Koenig, King & Carson, 2012). The evidence suggests that the positive effects of religion (e.g. on lifestyle, emotional stability and in social support) are stronger than the negative ones (anxiety, fatalism, etc.), but the balance will depend on the particular case.

It is hard to deny the importance of religion in sustaining collective identities or motivating action, including conflict and even violence. For these and other reasons religion enters into public policy debates. A question on religion was included on the census of population in Great Britain for the first time in 2001, and similar questions have subsequently appeared on major national enquiries such as the Labour Force Survey. It needs to be established, though, whether these questions yield reliable data.

While understanding change over the life course would make an important contribution in all of these areas, we need first to be sure that any apparent changes are real. This study aims to test the reliability of responses on religious involvement. Measurement problems abound: it is not always clear what it means to have a religion or to be religious. There are multiple indicators of religiosity, and responses show high sensitivity to question wording and context.

The 1970 British Cohort Study offers an important opportunity to test the consistency of responses to questions on religion for a single cohort over three decades. The latest sweep (in 2012) asked questions on belief in God and life after death as well as religious affiliation and practice. A close examination of multiple sweeps of the BCS70 underlines the difficulties and limitations of using survey responses on religion (Sullivan, Voas, & Brown, 2012). In what follows I look again at the (un)reliability of the traditional questions and consider how far new questions on religious belief improve our understanding of the cohort members.

Religion versus religiosity

It is common to think of religion itself – Catholic, Anglican, Hindu, Buddhist – as the key variable in this domain. For some people, however, affiliation is purely nominal; others will have a serious personal commitment, seeing faith as important in their lives. Notional affiliation may matter less than degree of religious commitment, or 'religiosity'. (This term is used non-pejoratively to mean the quality of being religious, not the display of excessive or affected piety.) Religiosity is bound up with attitudes, behaviour and values, while religion

per se is arguably more like ethnicity or cultural heritage, something that may or may not be salient.

These two concepts lead to quite separate questions. On the one hand there is the issue of the social significance of being Methodist, Mormon or Muslim, relative to having some other affiliation, or none; on the other, the issue is how far degree of religiosity matters. Change over time may be a matter either of growth or decline in particular denominations or in the commitment shown by those involved. The challenge for social scientists has been to find variables that capture the different dimensions of religious involvement. From a life course perspective, an additional consideration is that affiliation and practice may mean different things at different points, as religion is variously associated with familial culture or autonomous choice.

Table 1 lists the key questions on religion that have been used in the BCS70. Religion of upbringing and current religious identity concern affiliation. Although self-identification with a religion does not necessarily entail commitment, the growth in the number of those who say that they have no religion has ironically turned the simple willingness to accept a denominational label into an indicator of religiosity. The questions shown in the lower part of the table – on religious practice and belief – relate more straightforwardly to religiosity.

Attendance at services is a standard measure of commitment, though it is far from perfect. Some people attend for personal, family or social reasons in the absence of faith or even affiliation. Some groups (e.g. Catholics) attach more importance to it than other denominations. In some religions (e.g. Islam and Orthodox Judaism) the obligation falls on men rather than women. Overstatement of attendance is common (Hadaway, Marler, & Chaves, 1993), to a degree that may vary according to subjective perceptions of the social desirability of religious activity.

Table 1. Questions on religion in BCS70 by topic and survey year

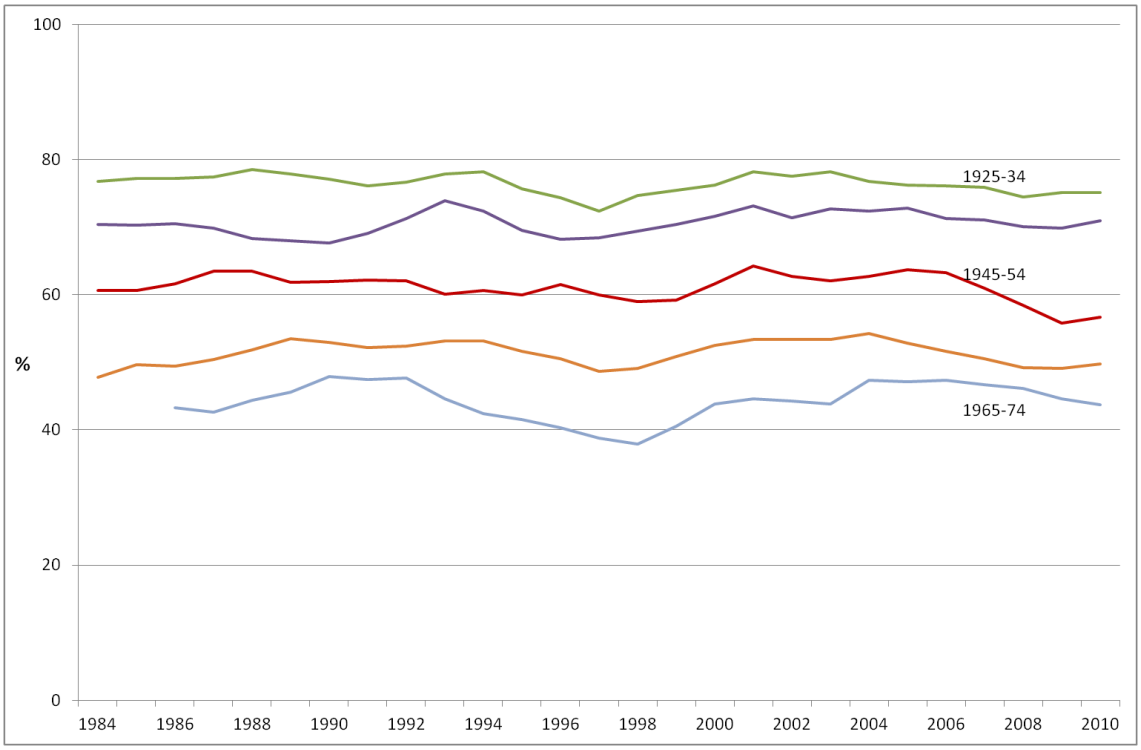
Upbringing / heritage	
1986	What religion were you born into? + tick boxes
2004	Were you raised according to any particular religion? + showcard if 'yes' *
2012	Were you raised according to any particular religion? + tick boxes
Identity / affiliation	
1996	Do you regard yourself as belonging to any particular religion? + write-in
2000	What is your religion, if any? + showcard *
2012	Do you now see yourself as belonging to any particular religion? + tick boxes
Self-perception / behaviour	
1986	Is religion an important part of your life? [three options + NA + DK]
2004	Do you actively practise any religion now? [If 'yes'] How often, if ever, do you attend any kind of religious service or meeting?*
2012	How often, if ever, do you attend any kind of religious service or meeting?
Belief	
2012	Which of these statements comes closest to expressing what you believe about God? [six options]
	Do you believe in life after death? [four options]

** Asked in a face-to-face interview; questions not so marked appeared on self-completion forms.*

Religious change in Great Britain

Religious affiliation in Great Britain has declined in recent decades. As shown in Figure 1, the changes arise entirely from differences between rather than within generations. The lines for each birth cohort are essentially horizontal: there is remarkable stability over the adult life course in religious identification, at least in aggregate. By contrast, affiliation drops from one cohort to the next for years of birth going back to the beginning of the 20th century. Note in particular the line for people born 1965-74, which can be used as a proxy for the 1970 birth cohort.

Figure 1. Belongs to a religion (%) by decade of birth and survey year (BSA 1983-2011)



Source: British Social Attitudes survey, 1983-2011. Three-year moving average. The count in each individual cohort and survey year ranges from 124 to 851.

The same is true for attendance at services. There are substantial differences between the oldest generations (among whom close to 30% claim to attend at least monthly) and the youngest (among whom the figure is around 9%). Secularization is being driven by generational replacement.

These repeated cross-sectional surveys suggest that religious practice as well as affiliation is fairly stable over the adult life course. No cohort seems to have either higher or lower attendance at the beginning than at the end of the three decades covered by the British Social Attitudes surveys. People do not seem to become much more religiously active with age, but nor have they become less active over time.

The aggregate stability may disguise a good deal of individual-level volatility, however. A comparison of waves 1 and 9 (1991 and 1999) of the British Household Panel Survey shows that the frequency distribution of religious affiliation is utterly static, from which it is tempting to conclude that religious identity is a stable attribute. Closer examination at the individual rather than the aggregate level reveals that a remarkable 27% of respondents interviewed in both surveys supplied different religious labels for themselves at the two dates (Voas, 2007b). No doubt some of those panel members really did change allegiance (between denominations or between affiliation and no religion), but it is likely that many are simply uncertain or ambivalent. The line between having

an affiliation and none can be rather fuzzy, and people in this 'liminal' area (sometimes stating a religious identity and sometimes not) may have distinct characteristics (Lim, MacGregor, & Putnam, 2010).

The key issue in looking at individual change is whether we are seeing real shifts in religious identity or religiosity, measurement error, or a kind of indifference that translates into a liminal state between having and not having a religion (Voas, 2009).

Do we know who was raised in a religion?

Cohort surveys make it possible, at least in principle, to see whether someone is changing and in what direction. The BCS70 has gone back to the same people year after year, decade after decade, to find out what is new and what has stayed the same. There are also ways of creating a time series other than through contemporaneous data collection, and the cohort study asks respondents to reconstruct their family, education or work histories. Such retrospective data are also useful with religion; at the simplest level respondents can be asked whether they were raised in one. People could also be asked about their childhood attendance at religious services, the religious affiliation or practice of their parents at that time, and so on, though such data are not available in the BCS70.

The overwhelming majority (93%) of cohort members (CMs) responding at age 16 reported that they had been 'born into' a religion. More than two thirds (69%) of those naming a religion specified the Church of England, which seems rather high. We know, for example, that exactly half of people born in England in 1970 had Anglican christenings (Voas, 2003). Not everyone 'born into' the religion would necessarily have been baptised, but then the BCS70 also includes people from Scotland and Wales, where the Anglican proportion is considerably lower.

The considerable majority of CMs who claimed in 1986 to have been born into a religion did not identify with it ten years later. In 1996, only 39% of CMs responding to a postal survey saw themselves as belonging to any particular religion. The importance of religion at age 16 is a reasonably good predictor of who did or did not retain an affiliation, as shown in Table 2.

Table 2. Belonging to a religion at age 26 by importance of religion at age 16 (%)

	Belongs to a religion, age 26		
Religion important, age 16	**Yes**	**No**	***N***
Yes, very important	86	14	*353*
Yes, quite important	68	32	*924*
Don't know	41	59	*238*
No, not important	31	69	*2,313*
Have no religion	11	89	*395*
	1,778	*2,445*	*4,223*

The sweeps at ages 34 and 42 included essentially identical questions on religious upbringing. The responses were very different: at age 34 only 55% said that they had been raised according to a religion, while at age 42 the figure was 67%. In 2004 the religion questions were asked in a face-to-face interview, whereas in 2012 they appeared on the self-completion form. The earlier version was a two-stage question (first the yes/no filter, then a list of religious groups), while the later question was unitary. Still, the contrast in responses seems astonishing, given that the surveys were only eight years apart and the respondents were in mid-life. The findings shown in Table 3 may shake our confidence in the possibility of studying religion using retrospective data.

Table 3. 'Thinking first of your childhood, were you raised according to any particular religion?' Responses in 2004 and 2012

	2012 response (%)			
2004 response	**Yes**	**No**	**Total**	***N***
Yes	90	10	100	*3,888*
No	40	60	100	*3,070*
				6,958
Respondents who had a religion in 2012				
Yes	97	3	100	*2,585*
No	76	24	100	*929*
				3,514
Respondents who had no religion in 2012				
Yes	76	24	100	*1,263*
No	25	75	100	*2,121*
				3,384

Overall about a quarter of people flip-flop (as predicted by the 'law' that "a quarter of responses to any question on religion are unreliable", Voas, 2007b: 149), and as seen in the lower segments of Table 3 the ones most prone to it (accounting for 43% of all switchers) are those with an affiliation who said in the earlier period that they were not raised in a religion. These respondents are evidently unsure about how to describe their childhood experience, and this lack of clarity may simply suggest that their religious upbringing was weak. An alternative hypothesis is that people tend to reconstruct their histories to make them consistent with their present status.

As mentioned above, 93% of responding CMs in 1986 said that they had been born into a religion. In 2012, two thirds (67%) claimed to have been raised according to a religion. As one might expect, relatively few people (20%) who said at age 16 that they were not born into a religion claimed at age 42 to have been raised in one. The main contrast is the shift from 'born into' to 'not raised according to'. A quarter of the whole sample responding at both dates fall into this group.

Of respondents who said that they were born into the Church of England (C of E) at age 16, slightly fewer than half said that they were raised C of E at age 42. Just short of a third said that they were not raised in a religion, and most of the remainder gave 'Christian' as their religion of upbringing.

The social class of parents at birth of the CM is closely related to upbringing reported at age 42: 62% in classes V and IV were raised in a religion, versus 76% of professionals (I). The same does not apply to 'religion born into' reported at age 16, where it is 92% for classes IV and V, 94% for III, 92% for II, and 90% for I. A similar U-shaped pattern is seen for the importance of religion.

There is a large gender gap in current affiliation (2012): 57% vs 43% for women and men respectively. Although one might suppose that there would be no gender difference in the proportion with a religious upbringing, there is a gap here, too: 69% vs 65%. It seems likely that this difference is largely the result of respondents ironing out the extent of change, but a portion might be genuine (if boys and girls are socialized differently with respect to religion).

For people brought up with a religion, most were raised as Anglican, Catholic or Christian (no denomination). Among men, about 60% of Anglicans and Catholics and 54% of the Christians retain the identity. Among women, the figures are 77% and 69%. Thus women are somewhat more likely than men to report having had a religious upbringing, and among those who do, women are much more likely than men to continue to identify with a religion.

Men are disproportionately likely to report being raised in no religion or as 'Christian', while being less likely than women to report being raised as Anglican or Catholic. The difference is 2 percentage points for both Catholics and Anglicans, though as the base is about 10% and 30% respectively, men seem particularly disinclined to report a Catholic upbringing.

How reliable are statements about religious identity?

In 1996 and 2012 the wording of the question about religion was that used in the British Social Attitudes survey, whereas in 2000 it corresponded to the question that had been adopted for the 2001 census. A considerable amount of attention has been given to the differences that result (Voas & Bruce, 2004). The contrast is especially stark in the BCS70. At age 26, only 39% of responding CMs said that they regarded themselves as belonging to a religion. Four years later, at age 30, nearly twice as many (74%) chose a religious group when asked "what is your religion, if any?". Among people responding to both surveys, the considerable majority (61%) of those who answered negatively in 1996 did name a religion in 2000.

The difficulty of framing an entirely neutral question about religious affiliation is evident (Voas, 2007b; Voas & Bruce, 2004). "What is your religion?" seems to imply that the respondent should have one (though the addition of "if any" clearly helps). Conversely, the term 'belonging' might be understood as requiring active participation rather than passive self-identification. The 2012 question ("Do you see yourself as belonging to any particular religion?") tries to counter-balance the apparent commitment of 'belonging' with the subjectivity of 'see yourself', but the wording could be read as sceptical and hence off-putting: "Do you *really* see yourself as being part of some religious group [despite your lack of involvement]?" How respondents interpret the question is far from clear.

A further puzzle is the contrast between responses at ages 26 and 42 (Table 4). Although the questions appear to be the same, identification with a religion rises from 40% in 1996 to 52% in 2012. This shift might suggest that cohort members had genuinely become more religious, but evidence from elsewhere (e.g. the British Social Attitudes survey, as shown in Figure 1) make it unlikely. Framing effects might be a factor: in 2012 respondents are first asked whether they were raised according to any particular religion, which could incline them to give a consistent answer to the follow-up on current belonging. The most probable explanation concerns a subtle difference in the way the responses were recorded. Both sweeps used self-completion forms, but in the 1996 the question was no/yes, where the respondent was asked to write in which religion if 'yes' was ticked, while in 2012 there was a list of tick boxes with 'no religion' as simply one of the options. Answering 'yes' in 1996 thus involved more work – both in deciding what to call the religion and in writing the name on the form – than in 2012. Such 'satisficing' behaviour by survey respondents has been widely observed, even to the extent of omitting children (Ni Bhrolchain, Beaujouan, & Murphy, 2011).

Table 4. "Do you regard yourself as belonging to any particular religion?" (1996)
"Do you now see yourself as belonging to any particular religion?" (2012)
Responses in 1996 and 2012

	2012 response (%)			
1996 response	**Yes**	**No**	**Total**	***N***
Yes	80	20	100	*2,371*
No	33	67	100	*3,567*
				5,938

How stable is religiosity?

In view of the unreliability of religious affiliation, not to mention the uncertainty over its personal significance, it might seem preferable to focus on religiosity. One approach was that used in 1986, when the teenage CMs were asked about the importance of religion in their lives. Another is to enquire about religious activity. An intermediate concept is that of 'practising a religion', as used in 2004 (Tables 5 & 6). Most (81%) respondents who said that they practised religion attended services at least monthly.

One problem with this approach is that a relatively small minority of the British population is now 'practising': in 2004, only 14% claimed to be, and only 10% said in 2012 that they attend services at least monthly (though 25% do so occasionally). Another problem is that reported religious activity also varies over time. Of people who described themselves as practising in 2004 and who claimed to attend at least monthly, only a bare majority (54%) were still doing so in 2012. While there might be genuine change – for example if some of the attendance at the earlier period was motivated by school choice – there is also likely to be a good deal of churn as people move in and out of churchgoing, and the unreliability of responses will also contribute.

Table 5. Whether practising a religion at age 34 by importance of religion at age 16

	Whether practices any religion, age 34		
Religion important, age 16	**Yes**	**No**	***N***
Yes, very important	55	45	*357*
Yes, quite important	28	72	*936*
Don't know	13	87	*240*
No, not important	8	92	*2,418*
Have no religion	3	97	*377*
			4,328

Table 6. Attendance at services at age 42 by importance of religion at age 16

	Attends at least sometimes, age 42		
Religion important, age 16	**Yes**	**No**	***N***
Yes, very important	67	33	*324*
Yes, quite important	43	57	*843*
Don't know	24	76	*217*
No, not important	18	82	*2,161*
Have no religion	10	90	*351*
			3,896

Are beliefs consistently religious or unreligious?

There is a general assumption among religiously educated people in Western countries that God and life after death are linked. Christianity teaches that souls exist because God exists, and they survive because there is a supreme being who has made it so, as an expression of both love and judgement (Astley, 2010). But what do people actually believe? How closely does folk religion or popular theology coincide with conventional or official doctrine?

Every level of theism (belief in God) can be found among BCS70 members, who in this respect are characteristic of Europeans generally. Likewise belief in life after death is highly varied (Table 7). But belief – or disbelief – in God and in life after death do not always go together. First, the strength of belief is often different. Even among people at the extremes on the theism scale, many people are far from confident about the afterlife. In the BCS70, only two thirds of those who say 'I know that God really exists and I have no doubts about it' think that there is definitely life after death. Only a bare majority of atheists think that there definitely is not. In the other four response categories on theism, only a small minority are definite one way or the other about life after death.

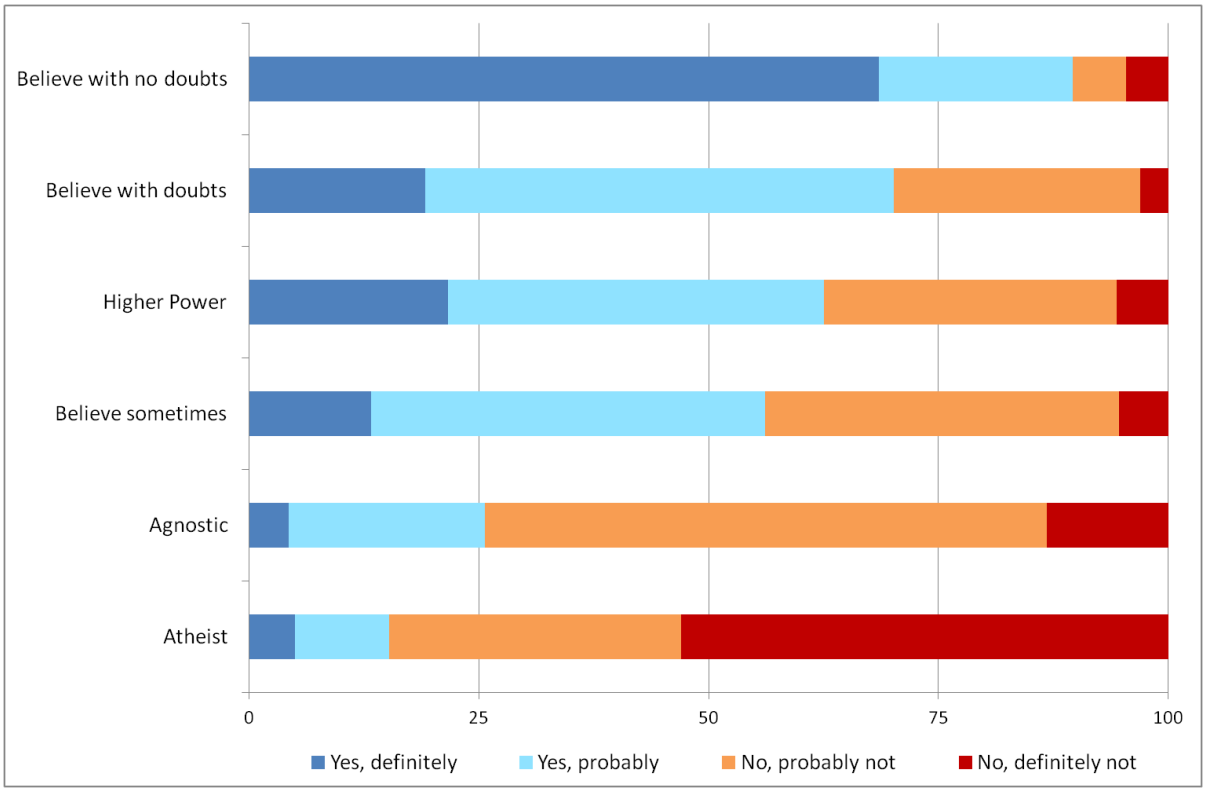
Secondly, the combinations often seem inconsistent by the standards of conventional religion. A quarter of agnostics believe in life after death. Among people who say that notwithstanding occasional doubts they believe in God, nearly a third do not. Even with the atheists and the undoubting theists, an appreciable proportion hold the 'opposite' view about life after death.

Table 7. Belief in God and life after death at age 42 (%)

I don't believe in God.	22
I don't know whether there is a God and I don't believe there is any way to find out.	21
I don't believe in a personal God, but I do believe in a Higher Power of some kind.	14
I find myself believing in God some of the time, but not at others.	12
While I have doubts, I feel that I do believe in God.	19
I know God really exists and I have no doubts about it.	12
Total	100
Believes in life after death	
Yes, definitely	18
Yes, probably	30
No, probably not	35
No, definitely not	17
Total	100

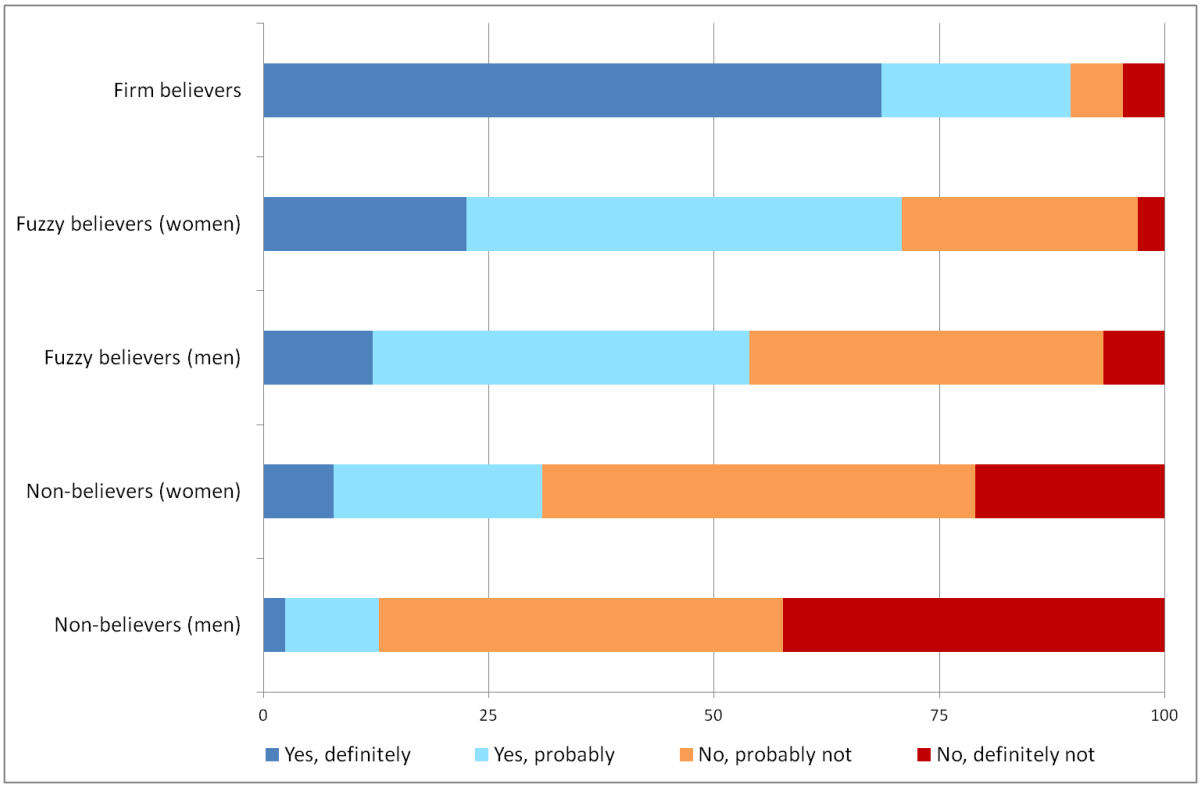
N= 8,579 for belief in God; N = 8,553 for belief in life after death

Figure 2. Belief in God and life after death, 2012



N=8,479

Figure 3. Belief in God and life after death by gender, 2012



N=8,479 Notes. 'Non-believers' are atheists and agnostics; 'Fuzzy believers' believe sometimes, believe in a Higher Power, or believe with doubts; 'Firm believers' have no doubts.

Religious upbringing has a strong influence on belief in God, but not nearly such a strong effect on belief in life after death. Of people raised with a religion (and recorded as such in both 2004 and 2012), 45% believe in God (with or without doubts). Only 29% are atheists or agnostics. Of people without a religious upbringing, two thirds (66%) are atheists or agnostics, and only 12% believe in God.

Life after death divides the cohort members down the middle, with just slightly fewer than half (48%) believing. Only a bare majority (54%) of even those raised in a religion believe in it, compared to 36% among those not raised in a religion.

Gender differences in religious belief are very substantial: 54% of men, but only 34% of women, are atheists or agnostics. The gap is even larger for belief in life after death: 60% of women but only 35% of men. Among believers, women are much more likely to be definite than men, and among non-believers, men are much more likely to be definite than women. Thus even controlling for belief in God, the gender gap in views about life after death is considerable (Figure 3).

Does religious belief coincide with religious identity?

Exactly half the respondents in the 2012 sweep report having a religion, and roughly half (48%) believe in life after death. Two thirds of these groups overlap: 67% of those with no religion do not believe, 63% of those with a religion do believe; 66% of believers in life after death belong to a religion, while 64% of non-believers do not belong. Or to put it another way, 33% neither belong nor believe, 32% belong and believe, and the remaining 35% is split between the 'inconsistent' categories.

Table 8. Belief in God and life after death by religious identity at age 42 (%)

	No religion	Main-line	Roman Catholic	Evan-gelical	Buddist	Jewish	Hindu + Sikh	Muslim	Other
I don't believe in God	40	5	2	0	40	18	1	1	10
I don't know whether there is a God	29	16	7	2	10	20	10	1	4
Believe in a Higher Power but not a personal God	15	14	6	2	45	9	14	1	42
I believe in God some of the time but not at others	8	17	14	2	5	25	12	2	7
While I have doubts, I feel that I do believe in God	7	31	38	23	0	16	26	6	12
I know God really exists and have no doubts	2	16	33	71	0	11	37	88	25
Total	100	100	100	100	100	100	100	100	100
Believes in life after death									
Yes, definitely	10	21	35	68	40	21	29	73	54
Yes, probably	23	38	36	17	25	26	47	12	26
No, probably not	40	33	26	7	30	40	15	6	11
No, definitely not	27	8	4	8	5	14	9	9	8
Total	100	100	100	100	100	100	100	100	100
N	*4,210*	*3,145*	*603*	*133*	*20*	*44*	*86*	*82*	*130*

Notes. 'Mainline' includes Christian (no denomination), Anglican, Methodist, Presbyterian and URC. 'Evangelical' includes Baptists and other Christians.

As one would expect, people who do not see themselves as belonging to a religion tend to be non-believers. That said, a third believe in life after death and close to the same number believe in God or a higher power, if only sometimes. The differences are small between people who identify with the various 'mainline' denominations (to adopt the American term), and hence those Christian groups have been collapsed in Table 8. Many of them express doubts about God and life after death. There is a higher level of belief among Roman Catholics and evangelical Protestants (defined here as Baptists and other Christians, though not all of the latter will in fact be evangelical). The numbers in the non-Christian categories are small, but the very high level of belief in both God and life after death among Muslims is noteworthy.

Belief in God can be dichotomized by contrasting atheists and agnostics with everyone else (including those who say that sometimes they believe and sometimes they don't). On that basis, 43% are non-believers. The overlap between belief in God and in life after death is far from complete: 39% believe in both, 35% are consistent non-believers, and the remaining 26% are split (mostly theists who do not believe in an afterlife). Interestingly, the inconsistent proportion is much the same for people with a religion (27%) and those without (25%). The difference is that the former are mostly theists who do not expect there to be life after death, whereas the latter are split fairly evenly between God and life after death.

A religious typology

Rather than categorising people by whether they identify with a religion or not, or attend services or not, there would seem to be merit in using a more refined classification. The information on belief, and in particular the distinction between people who are sceptical about both God and life after death, those who believe in both, and the unorthodox (who believe in one but not the other), offers the opportunity to propose a new typology.

Table 9. A religious typology

Type	%	Label	Description
1	28	Non-religious	Does not have a religion and believes in neither God nor life after death.
2	7	Nominally religious	Identifies with a religion, but believes in neither God nor life after death.
3	21	Unorthodox non-religious	Does not have a religion or does not attend services. Believes in God or life after death but not both.
4	5	Unorthodox religious	Has a religion and attends services at least occasionally. Believes in God but not life after death (or in few cases, vice versa).
5	10	Non-identifying believers	Does not have a religion, but believes in God and life after death.
6	14	Non-practising religious	Has a religion and believes in God and life after death. Does not attend services.
7	15	Actively religious	Has a religion and believes in God and life after death. Attends services.

Seven religious types are described in Table 9. Note that attendance at services is only used to define categories 4, 6 and 7: the privately religious attend rarely or never, while the religious unorthodox and the actively religious do go at least sometimes. Although non-attendance is not a criterion for classification as non-religious, unreligious unorthodox, or private believers, practically all of the people in these groups are non-attenders (96%, 97% and 93% respectively). Only nominals are somewhat mixed, and even here three quarters do not attend.

An alternative label for type 3, the unorthodox non-religious, might be 'spiritual but not religious'. The 'spiritual' tag is less obviously appropriate than the 'not religious' one, however. These people express supernatural beliefs that do not seem to be conventionally Christian, but whether they have reflected on those beliefs is another matter. The lack of orthodoxy may suggest that the beliefs are weakly held, not that they reflect an unusual but carefully elaborated worldview.

The differences between men and women, and those with and without higher education, are most apparent among the non-religious and the actively religious (Table 10). Nearly half (47%) of men in the sample are non-religious or at best nominally religious, which is twice the proportion of women in those categories (24%). There are no gender differences among the unorthodox. With education, what is most striking is the bimodal distribution of people with degrees. People with few or no qualifications are disproportionately likely to be unorthodox and nonreligious. To put it another way, the key dimensions of religious identity, belief and practice are typically consistent (whether present or absent) among the well educated and inconsistent for others.

Table 10. Religious characteristics at age 42 by gender and education (%)

			Education		
Variable	***Men***	***Women***	***Low***	***Medium***	***High***
Raised according to any particular religion?					
No	35	31	41	31	25
Christian, no denomination	14	14	12	16	15
Church of England/Anglican	29	31	28	31	31
Roman Catholic	10	12	10	10	13
Other Christian	8	8	6	8	11
Non-Christian	3	4	3	3	5
Belongs to any particular religion?					
No	57	43	55	48	46
Christian, no denomination	13	15	11	15	15
Church of England/Anglican	17	23	20	21	20
Roman Catholic	5	9	7	7	8
Other Christian	4	5	3	5	5
Non-Christian	4	5	4	4	5
Attends any kind of religious service or meeting					
Once a week or more	5	7	4	5	9
Once a month or more	4	5	3	5	7
Sometimes but less than once a month	12	17	10	16	19
Never or very rarely	80	71	84	74	65
Belief about God					
I don't believe in God	30	15	23	19	25
I don't know whether there is a God	24	19	23	21	19
Higher Power, not personal God	14	14	13	14	15
I believe in God some of the time	9	14	12	13	9
While I have doubts, I believe in God	15	23	18	20	19
I know God really exists with no doubts	9	15	11	12	14
Whether believes in life after death					
Yes, definitely	12	25	19	19	17
Yes, probably	23	36	32	32	24
No, probably not	39	31	34	35	36
No, definitely not	26	9	15	15	23
Religious type					
Non-religious	38	19	28	25	32
Nominally religious	9	5	7	7	8
Unorthodox non-religious	21	21	24	21	16
Unorthodox religious	5	6	3	6	7
Non-identifying believers	8	12	12	10	6
Non-practising religious	9	18	16	16	10
Actively religious	11	19	10	16	21

Notes. 'Medium' education includes GCSEs through diplomas; 'High' refers to degree-level qualifications.

Discussion / conclusion

At first sight, birth cohort studies seem to offer an ideal opportunity to study the magnitude and direction of religious change over the life course. In the event, close examination of the multiple waves of the BCS70 mainly reveals an enormous amount of uncertainty in measurement, making it hard to detect whatever genuine change might have occurred.

The difficulty is foreshadowed by the responses at age 16 (the first sweep to collect any data on religion). On the one hand, 93% of cohort members said that they had been born into a religion. On the other hand, only 30% said that religion was either very important or quite important to them. The tension between having a religious heritage and not having much personal investment in religion colours all of their subsequent responses from young adulthood into middle age.

The first and most obvious result is a high degree of unreliability about reported past and present affiliation. Between 2004 and 2012, nearly a quarter (23%) of people changed their answers about whether they had been raised in a religion. Between 1996 and 2012, more than a quarter (28%) changed whether they saw themselves as belonging to a religion. Repeated cross-sectional surveys suggest that there is relatively little within-cohort change in religion and religiosity during adulthood, which implies that most of this movement results from the difficulty of pinning down something that respondents themselves are not very sure about.

The real changes that are most evident are those between age 16 and adulthood. Family influence will still be strong for many teenagers, and those reporting that religion was an important part of their lives were presumably describing family values. It is clear that a substantial proportion of those teenagers became relatively unreligious adults: of individuals who said at age 16 that religion was either very or quite important in their lives, barely more than a third (35%) described themselves as practising at age 34 and half said at age 42 that they never attend services. There is some movement in the opposite direction, but not nearly enough to compensate for the losses to religion.

The addition of questions on religious belief in 2012 allows us to classify the respondents by religiosity with a great deal more confidence than previously. The complexity of this topic is immediately apparent; some people seem wholly non-religious and a smaller number are actively (and consistently) religious, but the majority fall into intermediate categories that are defined by nominal allegiance, unorthodox belief, or belief in the absence of affiliation or practice (Voas, 2009). There are differences between men and women on all of the measures studied here, but the gender gaps are particularly striking for belief. The question of why women appear to be so much more religious than men, despite the trends towards both gender equality and secularity, is a large issue that has received a great deal of attention without producing any clear resolution (Trzebiatowska & Bruce, 2012; Voas, McAndrew, & Storm 2013).

If there is a moral, it may be that while single item measures may be useful for looking at religious trends if consistently worded, they are not especially helpful at the individual level. We need multiple items across all three key domains (identity, practice and belief) to obtain a reliable picture of religious commitment. With longitudinal studies, the questions should ideally be introduced at an early stage in order to study religious change over the life course.

Acknowledgements

The author receives funding from the UK Economic and Social Research Council via the ESRC Research Centre on Micro-Social Change. Two anonymous referees provided valuable comments on an earlier version. The author is grateful to Alice Sullivan and Matt Brown, who collaborated on an earlier BCS70 data note that helped to motivate this study.

References

Astley, J. (2010). *SCM Studyguide: Christian doctrine*. London: SMC Press.

Beckford, J.A., & Demerath, N.J. (Eds.). (2007). *The Sage handbook of the sociology of religion*. London: Sage.

Bengtson, V.L. (2013). *Families and faith: How religion is passed down across generations*. New York: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199948659.001.0001

Bruce, S. (2011). *Secularization: In defence of an unfashionable theory*. Oxford: Oxford University Press.

Clarke, P. (2009). *The Oxford handbook of the sociology of religion*. New York: Oxford University Press.

Crockett, A., & Voas, D. (2006). Generations of decline: Religious change in twentieth-century Britain. *Journal for the Scientific Study of Religion, 45*(4), 567-584. http://dx.doi.org/10.1111/j.1468-5906.2006.00328.x

Dillon, M., & Wink, P. (2007). *In the course of a lifetime: Tracing religious beliefs, practice, and change*. Berkeley: University of California Press. http://dx.doi.org/10.1525/california/9780520249004.001.0001

Fenn, R.K. (2000). *The Blackwell companion to the sociology of religion*. Oxford: Wiley-Blackwell.

Hadaway, C. K, Marler, P.L., & Chaves, M. (1993). What the polls don't show: A closer look at U.S. church attendance. *American Sociological Review, 58*(6), 741-752. http://dx.doi.org/10.2307/2095948

Heelas, P., & Woodhead, L. (2005). *The spiritual revolution: Why religion is giving way to spirituality.* Oxford: Blackwell.

Koenig, H., King, D., & Carson, V.B. (2012). *Handbook of religion and health*, second edn. New York: Oxford University Press.

Lim, C., MacGregor, C.A., & Putnam, R.D. (2010). Secular and liminal: Discovering heterogeneity among religious nones. *Journal for the Scientific Study of Religion, 49*(4), 596-618. http://dx.doi.org/10.1111/j.1468-5906.2010.01533.x

Ni Bhrolchain, M., Beaujouan, É., & Murphy, M.J. (2011). Sources of error in reported childlessness in a continuous British household survey. *Population Studies, 65*(3), 305-318. http://dx.doi.org/10.1080/00324728.2011.607901

Stark, R., & Finke, R. (2000). *Acts of faith: Explaining the human side of religion*. Berkeley, CA: University of California Press.

Stolzenberg, R.M., Blair-Loy, M., & Waite, L.J. (1995). Religious participation in early adulthood: Age and family life cycle effects on church membership. *American Sociological Review, 60*(1), 84-103. http://dx.doi.org/10.2307/2096347

Sullivan, A., Voas, D., & Brown, M. (2012). BCS70 data note: The art of asking questions about religion. Centre for Longitudinal Studies, Institute of Education.

Trzebiatowska, M., & Bruce, S. (2012). *Why are women more religious than men?* Oxford: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199608102.001.0001

Turner, B.S. (2010). *The new Blackwell companion to the sociology of religion*.

Oxford: Wiley-Blackwell. http://dx.doi.org/10.1002/9781444320787

Voas, D. (2003). Intermarriage and the demography of secularisation. *British Journal of Sociology, 54*(1), 83-108. http://dx.doi.org/10.1080/0007131032000045914

Voas, D. (2007a). Does religion belong in population studies? *Environment & Planning A, 39*(5), 1166-1180. http://dx.doi.org/10.1080/0007131032000045914

Voas, D. (2007b). Surveys of behaviour, beliefs and affiliation. In J. Beckford & N.J. Demerath (Eds.), *The Sage handbook of the sociology of religion* (pp. 128-150). London: Sage. http://dx.doi.org/10.4135/9781848607965.n8

Voas, D. (2009). The rise and fall of fuzzy fidelity in Europe. *European Sociological Review, 25*(2), 155-168. http://dx.doi.org/10.1093/esr/jcn044

Voas, D., & Bruce, S. (2004). The 2001 census and Christian identification in Britain. *Journal of Contemporary Religion, 19*(1), 23-28. http://dx.doi.org/10.1080/1353790032000165087

Voas, D., & Crockett, A. (2005). Religion in Britain: Neither believing nor belonging. *Sociology, 39*(1), 11-28. http://dx.doi.org/10.1177/0038038505048998

Voas, D., McAndrew, S., & Storm, I. (2013). Modernization and the gender gap in religiosity: Evidence from cross-national European surveys. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 65*, 259-283. http://dx.doi.org/10.1007/s11577-013-0226-5

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.

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

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

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

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

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

References. Please use the APA 6th edition and refer to examples in the full Guidelines.

Authors not complying with these reference guidelines will be asked to make all the necessary alterations themselves, if their paper is accepted for publication.

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

Supplementary material. Supplementary material may be uploaded with the submission, and if the paper is published, will be visible to the reader via a link on the RHS of the screen.

Submissions containing graphs, tables, illustrations or mathematics. All graphs, tables and illustrations should be embedded in the submitted text, and have clear, self-explanatory titles and captions. Mathematical expressions should be created in Word 2003 (or a compatible package) with equation editor 3.0, unless the author has good reason to use other software, in which case please contact 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.

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

INTRODUCTION

PAPERS

LLCS 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

SLLS Society for Longitudinal and Life Course Studies