

## Comparative Data Note

Harmonizing the measurement of social origin, cognitive ability and educational attainment across the National Child Development Study (NCDS), the British Cohort Study (BCS70), the Longitudinal Study of Young People in England (LSYPE), and the Avon Study of Parents and Children (ALSPAC)

Bastian Betthaeuser and Mollie Bourne  
October 2016

*SOCED Project Data Note 2016-10-11*



## Table of Contents

1. Introduction.....	5
2. Parental Social Class.....	7
3. Parental Social Status .....	7
4. Parental Education .....	8
5. Parental Income.....	9
6. Respondent Cognitive Ability.....	11
7. Respondent Educational Attainment .....	13
8. Associations between Social Origin Measures .....	15
9. Social Origins and Educational Attainment.....	16
10. Social Origins and Cognitive Ability .....	26
11. Cognitive Ability and Educational Attainment.....	34

## List of Tables

Table 1. Missingness on Key Variables .....	5
Table 2. Distribution of Parental Social Class (NS-SEC), Reduced Method (%) .....	7
Table 3. Mean and Standard Deviation of Parental CG Scores (normalised, theoretical min/max) .....	8
Table 4. Composite Measure of Parental Education (%).....	8
Table 5. Relative Measure of Parental Education (normalised, 0–1).....	9
Table 6. Summary Statistics for Income Variables.....	9
Table 7. Cognitive Ability Measure and PCA Results.....	11
Table 8. Distribution of Educational Attainment at Age 20 (% , five category measure, up to and including Key Stage 5) .....	13
Table 9. Distribution of Educational Attainment at Age 20 – Academic Only .....	13
Table 10. Distribution of Educational Attainment at Age 38 – NCDS and BCS70.....	14
Table 11. Distribution of Respondents’ Threshold and Transition Measures at Age 20 – LSYPE and ALSPAC.....	14
Table 12. Rank Correlations between Social Origin Dimensions.....	15
Table 13. Cross-tabulation of Social Class by Educational Attainment (Row %) – NCDS.....	16
Table 14. Cross-tabulation of Social Class by Educational Attainment (Row %) – BCS70.....	16
Table 15. Cross-tabulation of Social Class by Educational Attainment (Row %) – LSYPE.....	16
Table 16. Cross-tabulation of Social Class by Educational Attainment (Row %) – ALSPAC.....	16
Table 17. Cross-tabulation of Social Status Quartiles by Educational Attainment (Row %) – NCDS.....	17
Table 18. Cross-tabulation of Social Status Quartiles by Educational Attainment (Row %) – BCS70.....	17
Table 19. Cross-tabulation of Social Status Quartiles by Educational Attainment (Row %) – LSYPE.....	17
Table 20. Cross-tabulation of Social Status Quartiles by Educational Attainment (Row %) – ALSPAC...	17
Table 21. Cross-tabulation of Parental Education by Educational Attainment (Row %) – NCDS.....	18

Table 22. Cross-tabulation of Parental Education by Educational Attainment (Row %) – BCS70.....	18
Table 23. Cross-tabulation of Parental Education by Educational Attainment (Row %) – LSYPE .....	18
Table 24. Cross-tabulation of Parental Education by Educational Attainment (Row %) – ALSPAC .....	18
Table 25. Cross-tabulation of Parental Income Quintiles by Educational Attainment (Row %) – BCS70	19
Table 26. Cross-tabulation of Parental Income Quintiles by Educational Attainment (Row %) – LSYPE .....	19
Table 27. Cross-tabulation of Parental Income Quintiles by Educational Attainment (Row %) – ALSPAC .....	19
Table 28. Average Centre Four Global Log Odds Ratios and Test of Significance in the Differences between Social Origin Measures.....	20
Table 29. Log-Linear Models of Association between Social Origins and Educational Attainment (Age 20) across NCDS, BCS70 and ALSPAC.....	21
Table 30. Log-Linear Models of Association between Social Origins and Educational Attainment (Age 20) across NCDS, BCS70 and LSYPE.....	22
Table 31. Log-Linear Models of Association between Social Origins and Educational Attainment (Academic Only, Age 20) across NCDS, BCS70 and LSYPE.....	23
Table 32. Log-Linear Models of Association between Social Origins and Educational Attainment (Age 38) across NCDS and BCS70 .....	24
Table 33. Average Centre Four Global Log Odds Ratios and Test of Significance in the Differences between Cohorts.....	25
Table 34. Cross-tabulation of Social Class by Cognitive Ability (Row %) – NCDS .....	26
Table 35. Cross-tabulation of Social Class by Cognitive Ability (Row %) – BCS70 .....	26
Table 36. Cross-tabulation of Social Class by Cognitive Ability (Row %) – LSYPE.....	26
Table 37. Cross-tabulation of Social Class by Cognitive Ability (Row %) – ALSPAC.....	26
Table 38. Cross-tabulation of Social Status Quartiles by Cognitive Ability (Row %) – NCDS.....	27
Table 39. Cross-tabulation of Social Status Quartiles by Cognitive Ability (Row %) – BCS70.....	27
Table 40. Cross-tabulation of Social Status Quartiles by Cognitive Ability (Row %) – LSYPE .....	27
Table 41. Cross-tabulation of Social Status Quartiles by Cognitive Ability (Row %) – ALSPAC .....	27
Table 42. Cross-tabulation of Parental Education by Cognitive Ability (Row %) – NCDS.....	28
Table 43. Cross-tabulation of Parental Education by Cognitive Ability (Row %) – BCS70.....	28
Table 44. Cross-tabulation of Parental Education by Cognitive Ability (Row %) – LSYPE.....	28
Table 45. Cross-tabulation of Parental Education by Cognitive Ability (Row %) – ALSPAC.....	28
Table 46. Cross-tabulation of Parental Income Quintiles by Cognitive Ability (Row %) – BCS70.....	29
Table 47. Cross-tabulation of Parental Income Quintiles by Cognitive Ability (Row %) – LSYPE .....	29
Table 48. Cross-tabulation of Parental Income Quintiles by Cognitive Ability (Row %) – ALSPAC .....	29
Table 49. Average Centre Four Global Log Odds Ratios and Test of Significance in the Differences between Social Origin Measures.....	30
Table 50. Log-Linear Models of Association between Social Origins and Cognitive Ability across NCDS, BCS70 and ALSPAC.....	31
Table 51. Log-Linear Models of Association between Social Origins and Cognitive Ability across NCDS, BCS70 and LSYPE.....	32

Table 52. Average Centre Four Global Log Odds Ratios and Test of Significance in the Differences between Cohorts .....	33
Table 53. Cross-tabulation of Cognitive Ability by Educational Attainment (Row %) – NCDS.....	34
Table 54. Cross-tabulation of Cognitive Ability by Educational Attainment (Row %) – BCS70.....	34
Table 55. Cross-tabulation of Cognitive Ability by Educational Attainment (Row %) – LSYPE .....	34
Table 56. Cross-tabulation of Cognitive Ability by Educational Attainment (Row %) – ALSPAC .....	34
Table 57. Log-Linear Models of Association between Cognitive Ability and Educational Attainment (Age 20) across NCDS, BCS70 and ALSPAC.....	35
Table 58. Log-Linear Models of Association between Cognitive Ability and Educational Attainment (Age 20) across NCDS, BCS70 and LSYPE.....	35
Table 59. Log-Linear Models of Association between Cognitive Ability and Educational Attainment (Age 38) across NCDS and BCS70 .....	36
Table 60. Average Centre Four Global Log Odds Ratios and Tests of Significance in the Differences, For All Cohort Pairs – Cognitive Ability and Educational Attainment .....	36

## List of Figures

Figure 1. Frequency Distribution (Raw) and Percentile Distribution (Log-transformed) – BCS70 .....	10	
Figure 2. Frequency Distribution (Raw) and Percentile Distribution (Log-transformed) – LSYPE .....	10	
Figure 3. Frequency Distribution (Raw) and Percentile Distribution (Log-transformed) – ALSPAC.....	10	
Figure 4. Distribution of Cognitive Ability: NCDS	Figure 5. Distribution of Cognitive Ability: BCS70	12
Figure 6. Distribution of Cognitive Ability: LSYPE	Figure 7. Distribution of Cognitive Ability: ALSPAC	12
.....		

## 1. Introduction

This data note has been prepared for the research project on *Social Origins, Cognitive Ability and Educational Attainment: A Birth Cohort and Life-Course Perspective*. The document will present descriptive statistics comparing key variables across four different data sets: National Child Development Study (NCDS), British Cohort Study (BCS70), Longitudinal Study of Young People in England (LSYPE) and Avon Study of Parents and Children (ALSPAC). For detailed information on each of the datasets used please see Bourne and Betthaeuser (2016); Betthaeuser and Bourne (2016a, 2016b, 2016c); and Bourne (2016). Further information on the ALSPAC dataset, the ALSPAC data dictionary and detailed acknowledgements can be found in the Appendix.

The main independent variables are measures of respondents' cognitive ability in childhood and parental class, status, education and income. The key dependent variables are measures of respondents' educational attainment. The key independent and dependent variables are described in sections 2–7. Section 8 examines associations between social origin measures. Sections 9 and 10 present associations between social origins and educational attainment, and social origins and cognitive ability. Section 11 examines the relationship between cognitive ability and educational attainment.

A key aim is to take a consistent approach to the construction of variables in each of the surveys, to enable us to make cross-cohort comparisons. Largely we have been able to achieve this aim, though not without compromise. Namely, the nationally-representative LSYPE survey provides rich information on the majority of our focal variables; however, there are no measures of cognitive ability. Elsewhere we have explored the possibility of using scores on tests at Key Stage 2, but we deem this an inadequate proxy. Consequently, we instead use data from ALSPAC which covers a cohort of a comparable age (those born in the early 1990s). However, the ALSPAC sample is not a random sample of the population and suffers from a considerable amount of missing data. Thus, where possible, we present descriptive statistics for measures constructed in both of these datasets in order to determine the extent to which ALSPAC data can be considered reliable, taking LSYPE as a benchmark.

Each of the panel studies suffer from a degree of sample attrition as well as item missing data. Table 1 summarises the amount of missing data on key variables in each of the datasets.

**Table 1.** Missingness on Key Variables

	NCDS				BCS70				LSYPE				ALSPAC			
	Per variable missing		Cumulative remaining		Per variable missing		Cumulative remaining		Per variable missing		Cumulative remaining		Per variable missing		Cumulative remaining	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Initial sample	--	--	18558	100.00	--	--	18737	100.00	--	--	16122	100.00	--	--	15445	100.00
Educational attainment	3947	21.27	14611	78.73	6600	35.22	12137	64.78	7628	47.31	8494	52.69	3361	21.76	12084	78.24
Cognitive ability	4427	23.85	12450	67.09	7342	39.18	8808	47.01	2145	13.30	7969	49.43	8825	57.14	5651	36.59

Parental social class	7400	39.87	9323	50.24	5604	29.91	7957	42.47	1356	8.41	7668	47.56	1862	12.06	5455	35.32
Parental social status	7400	39.87	9323	50.24	5592	29.84	7865	41.98	1356	8.41	7668	47.56	1862	12.06	5455	35.32
Parental education	7034	37.90	7078	38.14	5276	28.16	7752	41.37	440	2.73	7668	47.56	1717	11.12	5429	35.15

## 2. Parental Social Class

In all four datasets the measure of parental class is based on SOC2000 occupational codes and a five-category measure of employment status. The SOC2000 codes for NCDS and BCS70 were supplied by Tim Morris as part of the ESRC Project ‘An examination of the impact of family socio-economic status on outcomes in late childhood and adolescence’ (ESRC Grant: RES-060- 23-0011), led by Paul Gregg.<sup>1</sup> Table 2 displays the distributions of parental NS-SEC classes across the datasets. These measures have been derived using the reduced method, as described in the ONS NS-SEC User Manual (2005).

When comparing the distribution of social class between datasets it should be noted that the measure for NCDS is based on fathers’ occupation only, as no information on mothers’ occupation is available. For all other datasets both parents’ occupations were considered when generating the social class measure (using the dominance approach). It should further be noted that the datasets vary with regards to the quality and detail of the information provided on parents’ employment status. Particularly the information on employment status given in the NCDS lacks detail.

The distribution of social class changes across the cohorts in the ways that we would expect: that is, a steady increase in the proportions of classes 1 and 2, a contraction of classes 5, 6 and 7, a largely stable proportion making up the intermediate class, and a fluctuating class 4. However it is notable that class 4 in ALSPAC is somewhat smaller than we might expect. Further, there are larger proportions in class 1 coupled with lower proportions in class 7 compared with LSYPE. The higher socioeconomic position of ALSPAC parents compared with those in the national population has previously been documented (Fraser et al, 2012). Aside from the discrepancies described, we consider the measures to be largely consistent.

**Table 2.** Distribution of Parental Social Class (NS-SEC), Reduced Method (%)

	NCDS	BCS70	LSYPE	ALSPAC
1. Higher managerial	8	11	16	21
2. Lower managerial	12	20	29	26
3. Intermediate	9	9	7	13
4. Small employers	13	13	9	5
5. Lower supervisory	20	18	14	12
6. Semi routine	16	14	13	14
7. Routine	22	15	12	9
<i>Total</i>	100	100	100	100

## 3. Parental Social Status

This research project measures social status according to the Chan–Goldthorpe (2004) scale, which assigns scores to individuals based on the occupational structure of close friendship. Parents’ SOC90 occupational codes are required for conversion to the scale.

It should be noted that parents’ occupational information is coded according to different schemas in the different datasets. The dataset supplied by Paul Gregg et al. (2012) provides occupations coded for the NCDS cohort using the three-digit SOC90 codes.<sup>2</sup> The BCS70 data provides occupational codes using the three-digit OPCS Classification of Occupations 1980 (CO80). These are then translated into three-

<sup>1</sup> For further details see: [http://doc.ukdataservice.ac.uk/doc/7023/mrdoc/pdf/ncds\\_bcs\\_occupation\\_coding.pdf](http://doc.ukdataservice.ac.uk/doc/7023/mrdoc/pdf/ncds_bcs_occupation_coding.pdf)

<sup>2</sup> [http://doc.ukdataservice.ac.uk/doc/7023/mrdoc/pdf/ncds\\_bcs\\_occupation\\_coding.pdf](http://doc.ukdataservice.ac.uk/doc/7023/mrdoc/pdf/ncds_bcs_occupation_coding.pdf)

digit SOC90 codes. The ALSPAC data set provides SOC90 codes, and the SOC2000 codes given in LSYPE were converted to SOC90 using conversion tables. It should be noted that the status measure in NCDS is based on fathers' occupation only, as no information on mothers' occupation is available. In all other datasets both parents' occupations were considered (using the dominance approach). Across the datasets, the measures of parental status that we use are therefore broadly comparable.

In order to allow for comparison across cohorts the scale was normalised to a 0 to 1 scale using the minimum and maximum of -0.6 and +0.6 for all datasets. The measures are summarised in Table 3. We observe an over-time increasing trend in parental status scores.

**Table 3.** Mean and Standard Deviation of Parental CG Scores (normalised, theoretical min/max)

	<b>NCDS</b>	<b>BCS70</b>	<b>LSYPE</b>	<b>ALSPAC</b>
Mean	0.24	0.45	0.57	0.61
S.D	0.28	0.29	0.27	0.27

## 4. Parental Education

The composite measure of parental education is generated by combining measures of educational attainment for mothers and fathers. Where the information from one parent is missing the modal value for individuals with partners in a given educational attainment category is used. When comparing the distribution of parental education across cohorts, it should be noted that the available information on mothers' and fathers' educational attainment differs across surveys. The NCDS only provides information on the age at which mothers and fathers left full time education. The other three surveys measure qualifications attained, but differ somewhat with respect to the specific qualification categories used to measure mothers' and fathers' educational attainment. Table 4 presents the distributions.

**Table 4.** Composite Measure of Parental Education (%)

	<b>NCDS</b>	<b>BCS70</b>	<b>LSYPE</b>	<b>ALSPAC</b>
7. Both parents have degree-level qualifications	2	2	7	10
6. One parent has degree-level qualification; other parent has lower qualification	4	11	11	14
5. Both parents have higher secondary or lower tertiary qualification	2	2	4	7
4. One parent has higher secondary or lower tertiary qualification; other parent has lower qualification	8	12	19	20
3. Both parents have secondary or lower qualification	22	20	38	37
2. One parent has secondary or lower qualification; other parent has no qualification	21	20	11	8
1. Neither parent has any qualification	41	33	10	4
<i>Total</i>	100	100	100	100



As would be expected in response to decades of educational expansion, the proportions of children with parents in the highest educational category – those who both have degrees – has steadily increased over time, while the proportions of children whose parents have no qualifications between them has decreased; and in the case of ALSPAC, more dramatically. This is consistent with the slightly higher parental class and status positions of parents in the ALSPAC sample, described above.

To allow for meaningful cross-cohort comparisons, we treat parental education in relative terms (Bukodi and Goldthorpe, 2013). A relative measure of parental education was generated by assigning scores to each category of the composite measure of parental education according to the percentage of parents falling below that level in the cumulative distribution (Bukodi and Goldthorpe, 2013). Table 5 presents the mean and standard deviation of this relative measure across surveys.

**Table 5.** Relative Measure of Parental Education (normalised, 0–1)

	<b>NCDS</b>	<b>BCS70</b>	<b>LSYPE</b>	<b>ALSPAC</b>
Mean	0.38	0.40	0.40	0.44
S.D	0.35	0.33	0.33	0.35

## 5. Parental Income

Measures of parental income are available in all surveys, though often these are not comparable and we expect to use them only in some of our auxiliary analyses. See the individual data notes for detailed descriptions of the different ways in which information on parental income were collected. The variables constructed for BCS70, LSYPE and ALSPAC are presented in this section.

BCS70 and LSYPE provide the most detailed, fine-grained information, and both give the combined weekly income of parents in gross amounts. ALSPAC instead provides information on family take-home income, in far less detail and the question wording is altered across surveys. For these reasons, we consider the parental income information in ALSPAC to be insufficiently reliable and the measure constructed insufficiently comparable. Table 6 provides a summary of the parental income measures constructed.

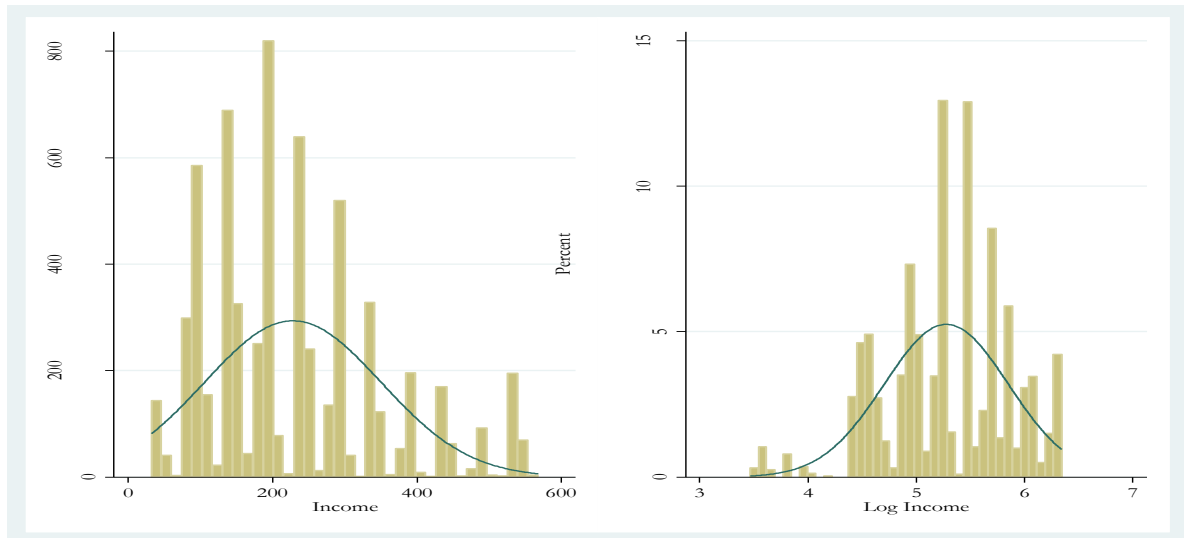
**Table 6.** Summary Statistics for Income Variables

		<b>Mean (SD)</b>	<b>Median</b>	<b>Min</b>	<b>Max</b>	<b>No. categories</b>
BCS70	<i>Gross</i>	228.04 (122.29)	196.30	32.10	567.60	129
LSYPE	<i>Gross</i>	567.59 (540.93)	424.50	4.50	7693.00	91
ALSPAC	<i>Net</i>	455.85 (202.40)	431.00	54.00	800.00	29

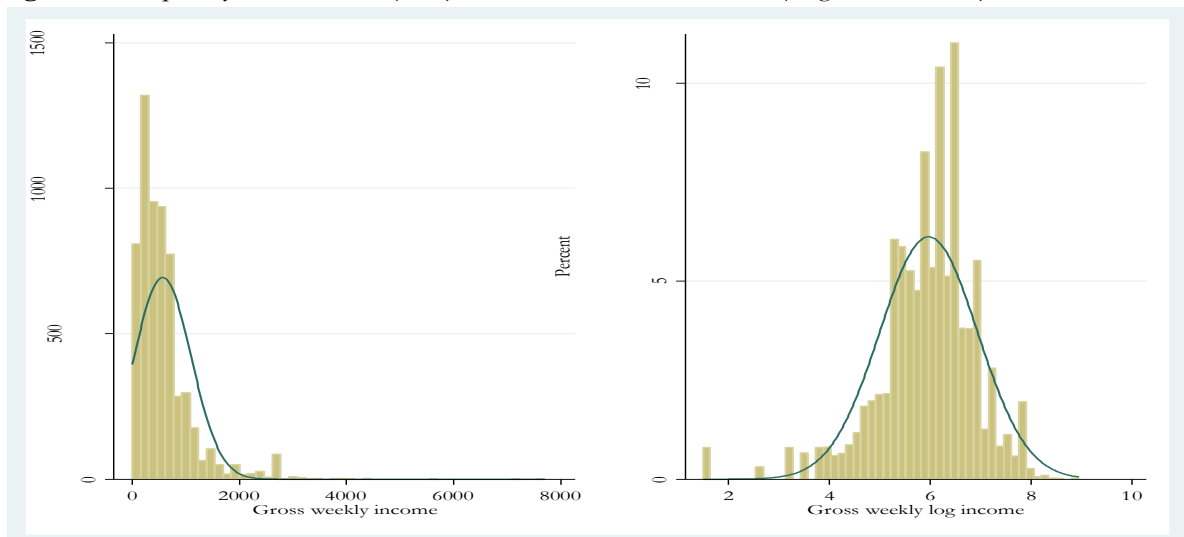
Notes: Prices are unadjusted.

Figures 1–3 give frequency and percentile distributions of raw and log-transformed income measures in each dataset.

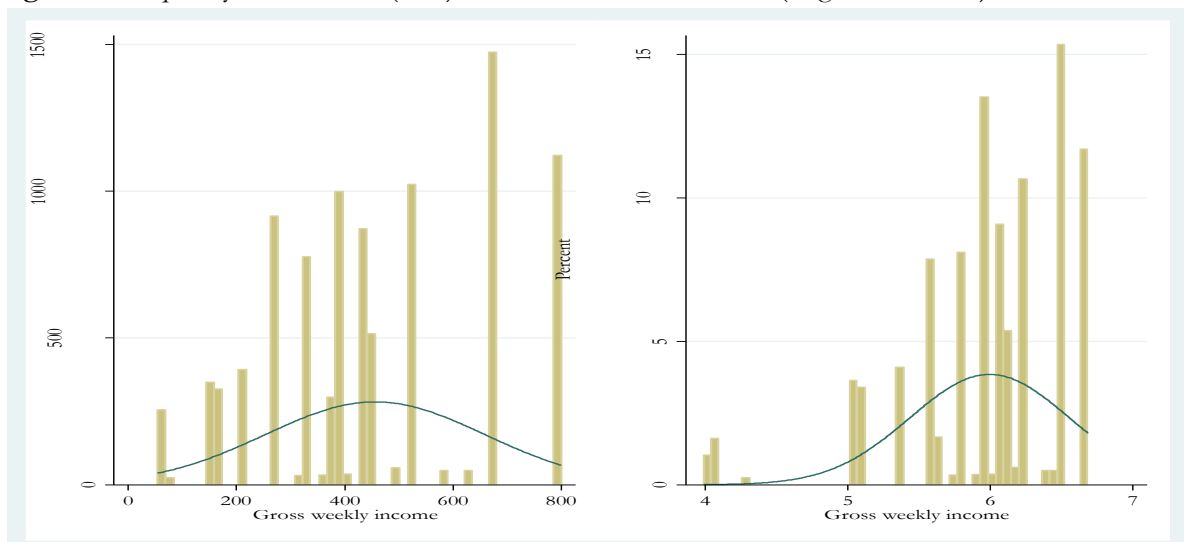
**Figure 1.** Frequency Distribution (Raw) and Percentile Distribution (Log-transformed) – BCS70



**Figure 2.** Frequency Distribution (Raw) and Percentile Distribution (Log-transformed) – LSYPE



**Figure 3.** Frequency Distribution (Raw) and Percentile Distribution (Log-transformed) – ALSPAC



## 6. Respondent Cognitive Ability

Cognitive ability was measured at age 10 and 11 in BCS70 and NCDS, while in ALSPAC it was measured at age 8. The measure of cognitive ability in LSYPE relies on Key Stage 2 performance information (age 11), which is taken as a proxy. A Principal Components Analysis (PCA) was run on the main cognitive ability (or KS2 performance) variables. Scores from the first unrotated component extracted were saved for each case, thus providing a measure of each child's cognitive ability – otherwise termed 'g'. Table 7 presents the Eigenvalue, loading, percentage of variance across the measures used that is accounted for by the first principal component score, and a summary of the generated variable for each survey.

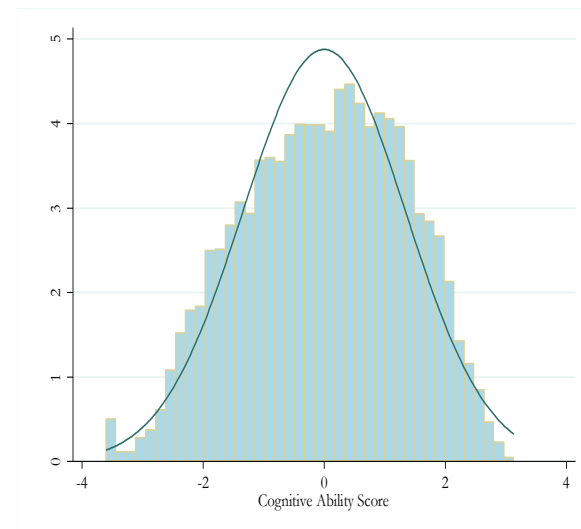
**Table 7.** Cognitive Ability Measure and PCA Results

<i>Survey</i>	<i>Measure</i>	<i>Eigenvalue</i>	<b>PCA</b>				<b>New Variable</b>			
			<i>Loading</i>		<i>% of variance</i>	<i>Min</i>	<i>Max</i>	<i>Mean (SD)</i>		
NCDS		1.81	Verbal	Non-verbal		0.90	-3.61	3.13	0.00 (1.34)	
			0.71	0.71						
BCS70		2.31	Word Def	Word Sim	Recall	Matrices	0.58	-5.64	5.13	0.00 (1.52)
			0.55	0.55	0.39	0.49				
ALSPAC	KS2em	1.75	English	Math	Science	0.88	-3.64	2.55	0.00 (1.32)	
	KS2ems	2.54	0.57	0.58	0.58	0.85	-5.09	3.03	0.00 (1.59)	
	WISC-2	1.51	Verbal	Performance		0.75	-5.07	3.84	0.00 (1.23)	
	WISC-11	3.67	--	--		0.33	-7.58	6.13	0.00 (1.92)	
LSYPE	KS2em	1.75	English	Math	Science	0.87	-3.93	>2.45	0.00 (1.32)	
	KS2ems	2.53	0.57	0.58	0.58	0.84	-5.46	>2.95	0.00(1.59)	

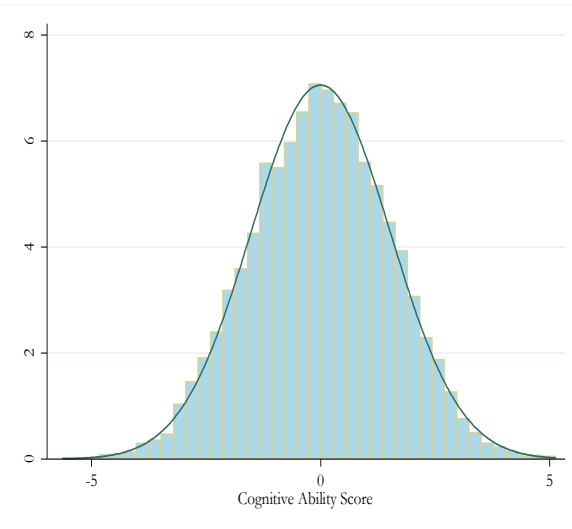
Figures 4–7 give the percentile distributions of each cognitive ability measure across the datasets. While Key Stage 2 scores and scores on tests of cognitive ability tend to be highly correlated, additional analyses carried out elsewhere have led us to conclude that the former cannot be considered a reliable proxy. Consequently, we use LSYPE to carry out auxiliary analyses and robustness checks but do not use these data in our main analyses.

Expectantly, cognitive ability scores tend to be normally distributed within each of the datasets. In our main analyses, we intend to group scores into cohort-specific cognitive ability quintiles in order to account for ‘Flynn effects’ (Flynn 1987) and any potential non-linear effects on educational attainment.

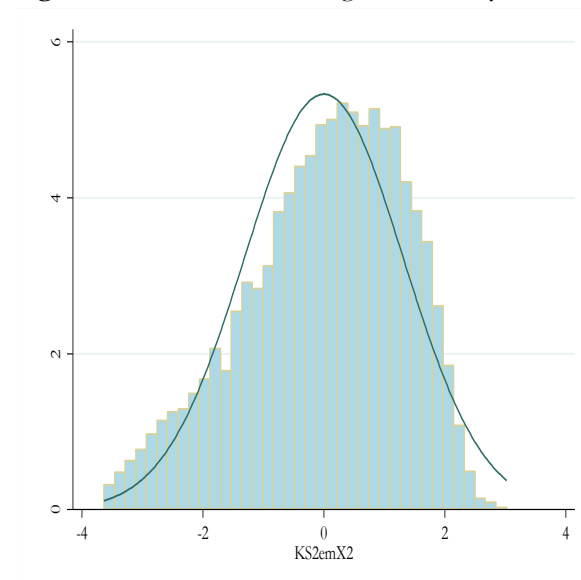
**Figure 4.** Distribution of Cognitive Ability: NCDS



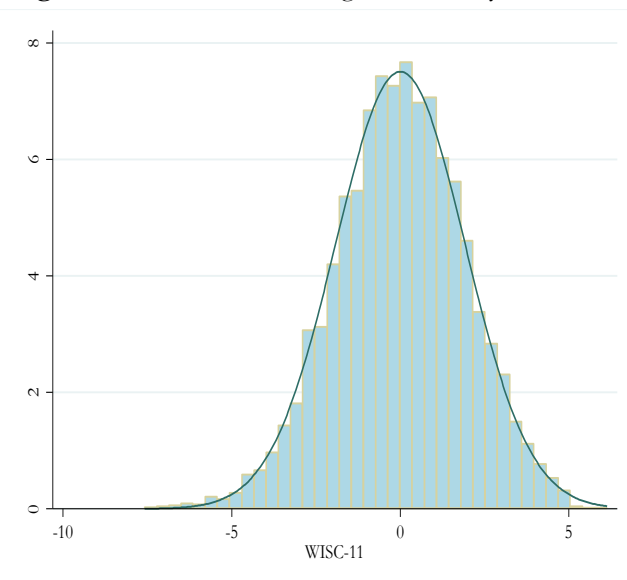
**Figure 5.** Distribution of Cognitive Ability: BCS70



**Figure 6.** Distribution of Cognitive Ability: LSYPE



**Figure 7.** Distribution of Cognitive Ability: ALSPAC



## 7. Respondent Educational Attainment

The measures for respondents' educational attainment were generated by using information on educational qualifications attained by cohort members that was given in each survey of the NCDS and BCS70. A key challenge in generating these measures has been the allocation of individuals for whom no educational qualifications have been recorded. This lack of information on qualifications may either signal that a given cohort member either has not attained any education at all, or it may constitute a case of missing information. This issue has been addressed by coding the information on the educational attainment of individuals as 'missing' where individuals did not participate in the relevant survey wave, and coding all other individuals as not having gained any educational qualification. For LSYPE and ALSPAC, information on educational attainment is determined using administrative data from the National Pupil Database (NPD), which has been linked to the surveys.

Table 8 shows the distribution of educational attainment up to age 20. We observe a general upgrading in educational attainment over time, and distributions between LSYPE and ALSPAC that are very similar.

**Table 8.** Distribution of Educational Attainment at Age 20 (% , five category measure, up to and including Key Stage 5)

	NCDS	BCS70	LSYPE	ALSPAC
5. 2+ A-level passes, NVQ 4	20	20	39	38
4. 5+ O-level passes or 1 A-level pass, NVQ 3	19	26	28	22
3. 1–4 O-level passes, NVQ 2	21	22	16	21
2. Below O-level, NVQ 1	11	8	15	15
1. No qualifications	29	24	2	4
<i>Total</i>	100	100	100	100

Two additional measures of educational attainment have been constructed. Using the NCDS, BCS70 and LSYPE datasets, we derive a measure of respondents' highest educational attainment at age 20, as above, but include only the achievement of academic qualifications. It has not been possible to create a comparable 'academic-only' measure with ALSPAC data. See Table 9. Using only NCDS and BCS70, whose samples are older, we also derive a measure of respondents' highest educational attainment at age 38, considering both academic and vocational qualifications. Since by this later age respondents have been able to acquire higher qualifications, the variable has two additional categories. The distributions of these measures are given in Table 10.

**Table 9.** Distribution of Educational Attainment at Age 20 – Academic Only

	Age 20 – Academic only		
	NCDS	BCS70	LSYPE
5. 2+ A-level passes	16	18	39
4. 5+ O-level passes or 1 A-level pass	15	21	23
3. 1–4 O-level passes	23	14	20
2. Below O-level	12	12	15
1. No qualifications	34	35	2
<i>Total</i>	100	100	100

**Table 10.** Distribution of Educational Attainment at Age 38 – NCDS and BCS70

	<b>Age 38</b>	
	NCDS	BCS70
7. Degree, NVQ 5 or 6, higher degree	13	19
6. Tertiary level sub-degree qualification, NVQ 4	14	14
5. 2+ A-level passes	4	3
4. 5+ O-level passes or 1 A-level pass, NVQ 3	17	21
3. 1–4 O-level passes, NVQ 2	22	21
2. Below O-level, NVQ 1	12	8
1. No qualifications	18	14
<i>Total</i>	100	100

Using LSYPE and ALSPAC, we are able to derive measures examining the proportions of respondents who made the transition to further study after the completion of compulsory education. Table 11 presents the proportions that made the transition to academic study at Key Stage 5 (A-level) and the proportions that made the transition to either academic or vocational study at Key Stage 5.

**Table 11.** Distribution of Respondents' Threshold and Transition Measures at Age 20 – LSYPE and ALSPAC

	<b>LSYPE</b>		<b>ALSPAC</b>	
	%	N	%	N
<b>Transition to A-level (academic only)</b>				
Did transition	54	4919	45	5454
Did not transition	46	2749	55	6612
<b>Transition to Key Stage 5 (academic and vocational)</b>				
Did transition	64	5692	61	7337
Did not transition	36	1976	39	4729

## 8. Associations between Social Origin Measures

Table 12 shows the rank correlations between the different social origin measures. Consistent for each dataset, the highest correlations exist between parental class and parental status, though these are always low enough to avoid posing problems of multicollinearity for our multivariate analyses.

**Table 12.** Rank Correlations between Social Origin Dimensions

		Parental class	Parental education	Parental status	Parental income
NCDS	Parental class	1.00**			
	Parental education	0.37**	1.00**		
	Parental status	0.60**	0.30**	1.00**	
	Parental income	N/A	N/A	N/A	N/A
BCS70	Parental class	1.00**			
	Parental education	0.58**	1.00**		
	Parental status	0.63**	0.54**	1.00**	
	Parental income	0.48**	0.49**	0.41**	1.00**
LSYPE	Parental class	1.00**			
	Parental education	0.57**	1.00**		
	Parental status	0.71**	0.60**	1.00**	
	Parental income	0.43**	0.38**	0.40**	1.00**
ALSPAC	Parental class	1.00**			
	Parental education	0.49**	1.00**		
	Parental status	0.70**	0.52**	1.00**	
	Parental income	0.37**	0.40**	0.39**	

Notes: \*\* p<0.01

## 9. Social Origins and Educational Attainment

Tables 13–16 present cross-classifications of **parental social class** and cohort members' educational attainment across each dataset. The proportion of cohort members attaining the highest qualification level is markedly higher among younger generations and for all classes alike. Class differences in the proportions attaining the highest qualification level grow slightly over time, while differences in the proportions with no qualifications shrink by around two-thirds between the 1958 and 1990s cohorts.

**Table 13.** Cross-tabulation of Social Class by Educational Attainment (Row %) – NCDS

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
1. Higher managerial	47	24	12	4	13	100
2. Lower managerial	37	23	18	6	16	100
3. Intermediate	29	24	21	7	19	100
4. Small employers	18	22	23	11	27	100
5. Lower supervisory	15	21	25	12	27	100
6. Semi routine	13	17	26	11	32	100
7. Routine	9	15	23	15	37	100
<i>Total</i>	21	20	22	11	27	100

**Table 14.** Cross-tabulation of Social Class by Educational Attainment (Row %) – BCS70

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
1. Higher managerial	41	28	14	4	12	100
2. Lower managerial	30	30	18	5	17	100
3. Intermediate	24	28	23	7	18	100
4. Small employers	19	28	22	8	24	100
5. Lower supervisory	13	26	26	10	26	100
6. Semi routine	11	23	24	11	31	100
7. Routine	8	19	29	11	33	100
<i>Total</i>	21	26	22	8	23	100

**Table 15.** Cross-tabulation of Social Class by Educational Attainment (Row %) – LSYPE

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
1. Higher managerial	65	22	8	--	--	100
2. Lower managerial	48	30	13	8	1	100
3. Intermediate	39	31	16	--	--	100
4. Small employers	37	29	17	--	--	100
5. Lower supervisory	23	32	22	21	2	100
6. Semi routine	23	29	21	24	3	100
7. Routine	19	25	20	29	7	100
<i>Total</i>	39	28	16	15	2	100

**Table 16.** Cross-tabulation of Social Class by Educational Attainment (Row %) – ALSPAC

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
1. Higher managerial	66	18	10	4	1	100
2. Lower managerial	49	23	17	8	3	100
3. Intermediate	39	25	21	12	3	100
4. Small employers	21	23	28	24	5	100
5. Lower supervisory	27	26	25	18	4	100
6. Semi routine	21	23	30	22	5	100
7. Routine	19	22	26	25	9	100
<i>Total</i>	40	23	21	13	4	100



Tables 17–20 present cross-classifications of **parental social status** and cohort members' educational attainment across each dataset. There is relatively little change in the associations between parental status and respondents' educational attainment between the 1958 and 1970 cohorts, while in the 1990s cohorts the status gradient is more marked.

**Table 17.** Cross-tabulation of Social Status Quartiles by Educational Attainment (Row %) – NCDS

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
4. Top	41	23	16	5	15	100
3. 3 <sup>rd</sup>	18	22	25	11	24	100
2. 2 <sup>nd</sup>	15	17	24	12	31	100
1. Bottom	13	18	23	13	33	100
<i>Total</i>	21	20	22	11	27	100

**Table 18.** Cross-tabulation of Social Status Quartiles by Educational Attainment (Row %) – BCS70

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
4. Top	37	30	15	4	15	100
3. 3 <sup>rd</sup>	21	28	23	8	20	100
2. 2 <sup>nd</sup>	13	25	26	10	27	100
1. Bottom	12	23	24	11	31	100
<i>Total</i>	21	26	22	8	23	100

**Table 19.** Cross-tabulation of Social Status Quartiles by Educational Attainment (Row %) – LSYPE

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
4. Top	64	22	9	--	--	100
3. 3 <sup>rd</sup>	47	31	12	--	--	100
2. 2 <sup>nd</sup>	31	32	19	16	2	100
1. Bottom	20	27	21	27	4	100
<i>Total</i>	39	28	16	15	2	100

Notes: -- Values omitted to preserve anonymity

**Table 20.** Cross-tabulation of Social Status Quartiles by Educational Attainment (Row %) – ALSPAC

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
4. Top	68	17	10	4	1	100
3. 3 <sup>rd</sup>	49	25	16	8	2	100
2. 2 <sup>nd</sup>	36	27	23	12	3	100
1. Bottom	18	21	29	25	7	100
<i>Total</i>	40	23	21	14	4	100

Tables 21–24 give cross-classifications of **parental education** and cohort members' educational attainment across each dataset. For the 1958 cohort, the proportions of respondents achieving the highest qualification level differ rather dramatically between those whose parents both have degrees and those whose parents have no qualifications. Unexpectedly, these proportions then fall for all categories of parental education, before picking up again among the 1990s cohorts. Here we also see some quite distinct differences between the LSYPE and ALSPAC samples.

**Table 21.** Cross-tabulation of Parental Education by Educational Attainment (Row %) – NCDS

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
7. Both have degrees	69	12	7	2	10	100
6.	53	24	10	2	12	100
5.	52	23	11	2	12	100
4.	34	26	19	5	16	100
3.	19	20	24	11	26	100
2.	19	20	24	12	25	100
1. No qualifications	11	18	24	14	33	100
<i>Total</i>	20	20	22	11	27	100

**Table 22.** Cross-tabulation of Parental Education by Educational Attainment (Row %) – BCS70

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
7. Both have degrees	57	23	5	3	13	100
6.	46	29	10	3	12	100
5.	37	33	11	3	16	100
4.	26	32	21	5	17	100
3.	19	29	24	8	20	100
2.	13	25	26	9	27	100
1. No qualifications	10	21	25	12	32	100
<i>Total</i>	21	26	22	8	24	100

**Table 23.** Cross-tabulation of Parental Education by Educational Attainment (Row %) – LSYPE

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
7. Both have degrees	77	15	5	--	--	100
6.	68	23	6	--	--	100
5.	63	25	9	--	--	100
4.	44	30	15	9	1	100
3.	28	33	19	18	2	100
2.	23	29	21	23	3	100
1. No qualifications	20	23	19	32	7	100
<i>Total</i>	39	28	16	15	2	100

Notes: -- Values omitted to preserve anonymity

**Table 24.** Cross-tabulation of Parental Education by Educational Attainment (Row %) – ALSPAC

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	<i>Total</i>
7. Both have degrees	83	12	3	1	1	100
6.	67	19	10	3	2	100
5.	61	21	12	5	1	100
4.	40	26	21	11	2	100
3.	26	26	26	18	5	100
2.	16	20	31	26	7	100
1. No qualifications	11	15	28	32	14	100
<i>Total</i>	40	22	21	14	4	100

Tables 25–27 give cross-classifications of **parental income** and respondents' educational attainment for BCS70, LSYPE and ALSPAC. As would be expected, we see an increase in the proportions of those achieving the highest qualification level (as well as a decrease in those with no qualifications) over time, for all income groups. Again some differences between the LSYPE and ALSPAC samples can be noted.

**Table 25.** Cross-tabulation of Parental Income Quintiles by Educational Attainment (Row %) – BCS70

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	Total
5. Top	40	29	15	4	13	100
4. 4 <sup>th</sup>	26	32	18	6	18	100
3. 3 <sup>rd</sup>	20	28	23	8	22	100
2. 2 <sup>nd</sup>	15	25	25	10	24	100
1. Bottom	10	22	26	10	32	100
<i>Total</i>	22	27	22	8	22	100

**Table 26.** Cross-tabulation of Parental Income Quintiles by Educational Attainment (Row %) – LSYPE

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	Total
5. Top	64	23	8	--	--	100
4. 4 <sup>th</sup>	42	32	15	--	--	100
3. 3 <sup>rd</sup>	34	29	18	16	3	100
2. 2 <sup>nd</sup>	27	29	21	20	3	100
1. Bottom	29	26	17	23	4	100
<i>Total</i>	39	28	16	15	2	100

Notes: -- Values omitted to preserve anonymity

**Table 27.** Cross-tabulation of Parental Income Quintiles by Educational Attainment (Row %) – ALSPAC

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	Total
5. Top	74	15	8	2	1	100
4. 4 <sup>th</sup>	60	20	13	6	1	100
3. 3 <sup>rd</sup>	47	25	19	8	2	100
2. 2 <sup>nd</sup>	36	27	21	13	3	100
1. Bottom	27	24	26	19	5	100
<i>Total</i>	46	23	18	11	2	100

5x5 contingency tables were then produced for all social origin-by-educational attainment pairs. Social class was recoded to five ordinal categories; for parental status and income, quintiles were used; and for parental education, categories 2–3 and categories 6–7 were collapsed. In an attempt to determine whether the association between educational attainment and social origin differs depending on which measure is taken for the latter, global log odds ratios were calculated for each of the tables, using the method proposed by Cox and Jackson (2009). Differences between the averages of the centre four global log odds ratios were calculated for each pair of tables, and were then tested for significance. The results of each of the tests are given in Table 28.

As indicated by the size of the average centre four global log odds ratios, the association between educational attainment and parental education is stronger than with any other measure of social origin, across all cohorts alike. Reflecting much of what has been shown above, the strength of association for each origin measure fluctuates slightly between the NCDS and BCS70 cohorts, then increases more substantially for the 1990 cohorts. The strength of associations is similar between ALSPAC and LSYPE for parental class and status; the association between parental education and educational attainment is

stronger among the LSYPE sample. For both the BCS70 and LSYPE samples, for whom the measures are comparable, income has the weakest association with educational attainment.

**Table 28.** Average Centre Four Global Log Odds Ratios and Test of Significance in the Differences between Social Origin Measures

		Ave. centre four		Difference	SE	Lower CI	Upper CI
		I.	II.				
NCDS	I.Class – II.Status	0.99	0.72	0.26**	0.05	-0.37	-0.16
	I.Class – II.Education	0.99	1.47	0.49**	0.08	0.33	0.64
	I.Status – II.Education	0.72	1.47	0.75**	0.08	0.60	0.90
BCS70	I.Class – II.Status	0.87	0.85	-0.01	0.05	-0.12	0.09
	I.Class – II.Education	0.87	1.18	0.31**	0.06	0.19	0.44
	I.Class – II.Income	0.87	0.82	-0.05	0.06	-0.18	0.08
	I.Status – II.Education	0.85	1.18	0.33**	0.06	0.20	0.45
	I.Status – II.Income	0.85	0.82	-0.03	0.06	-0.16	0.09
	I.Education – II.Income	1.18	0.82	-0.36**	0.07	-0.50	-0.22
ALSPAC	I.Class – II.Status	1.14	1.38	0.24**	0.06	0.13	0.36
	I.Class – II.Education	1.14	1.57	0.43**	0.07	0.30	0.56
	I.Class – II.Income	1.14	1.15	0.01	0.07	-0.13	0.15
	I.Status – II.Education	1.38	1.57	0.19**	0.07	0.05	0.32
	I.Status – II.Income	1.38	1.15	-0.23**	0.07	-0.38	-0.09
	I.Education – II.Income	1.57	1.15	-0.42**	0.08	-0.57	-0.27
LSYPE	I.Class – II.Status	1.19	1.33	0.14	0.08	-0.01	0.29
	I.Class – II.Education	1.19	1.72	0.53**	0.09	0.35	0.72
	I.Class – II.Income	1.19	1.03	-0.16**	0.08	-0.31	-0.01
	I.Status – II.Education	1.33	1.72	0.39**	0.10	0.20	0.58
	I.Status – II.Income	1.33	1.03	-0.30**	0.08	-0.46	-0.14
	I.Education – II.Income	1.72	1.03	-0.69**	0.10	-0.89	-0.50

## 9.1 Changes in the Associations between Social Origins and Educational Attainment across Cohorts

A series of log-linear models were run in order to test whether the associations between each component of social origins and educational attainment change in any systematic way across the cohorts. These analyses are of a three-way data array of cohort by origin by education. Categories for social origin variables are the same as are given in the tables above. The first model estimated is the log-linear model of conditional independence, which includes interactions between cohorts and origins, and cohorts and education. This model is substantively implausible, but serves as a useful baseline (Bukodi and Goldthorpe, 2015). The second model estimated is the log-linear model of constant association between origins and education across cohorts. This model recognises that an association exists between origins and education (net of marginal effects) by including a further interaction between origins and education, but it requires that the log odds ratios defining the origins–education association do not change across cohorts (Bukodi and Goldthorpe, 2015). The third model estimated is the log-multiplicative model – known as the unidiff model – which additionally estimates a set of coefficients (referred to below as unidiff parameters) that indicate the relative strength of the origins–education association that is specific to each cohort. This model allows us to test the possibility that from one cohort to another the log odds ratios defining the origins–education association uniformly increase or decrease by some common multiplicative factor (Bukodi and Goldthorpe, 2015). The results of these analyses are presented in Table 29.

**Table 29.** Log-Linear Models of Association between Social Origins and Educational Attainment (Age 20) across NCDS, BCS70 and ALSPAC

	<b>G<sup>2</sup></b>	<b>% total association explained</b>	<b>df.</b>	<b>p</b>	<b>DI</b>	<b>N</b>
<b>Class–Educational attainment</b>						
(1) Conditional independence	3746.0		72	0.00	13.4	31109
(2) Constant association	111.6	97.02%	48	0.00	2.0	31109
(3) Unidiff for OE	105.4	97.19%	46	0.00	2.0	31109
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>0.93 (BCS70)</i>	<i>1.04 (ALSPAC)</i>			
(3) vs. (2)	6.2	5.56%	2	0.05		
<b>Status–Educational attainment</b>						
(1) Conditional independence	3744.7		36	0.00	13.3	31109
(2) Constant association	348.5	90.69%	24	0.00	3.9	31109
(3) Unidiff for OE	198.6	94.70%	22	0.00	3.2	31109
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>1.08 (BCS70)</i>	<i>1.67 (ALSPAC)</i>			
(3) vs. (2)	149.9	43.01%	2	0.00		
<b>Education–Educational attainment</b>						
(1) Conditional independence	5042.4		72	0.00	14.1	31990
(2) Constant association	183.0	96.37%	48	0.00	2.3	31990
(3) Unidiff for OE	127.1	97.48%	46	0.00	2.0	31990
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>1.00 (BCS70)</i>	<i>1.33 (ALSPAC)</i>			
(3) vs. (2)	55.9	30.55%	2	0.00		

For each of the three social origin measures considered, the conditional independence models, as expected, give a poor fit to the data, misclassifying around 13–14% of all individual cases. The constant association model, again for each social origin measure, also does not give an acceptable fit to the data but the lack of fit is much reduced. The unidiff models still do not fit the data well, but in each case suggest a statistically significant improvement on the constant association model, indicating that some of the change in the origins–education association is of a uniform kind. For parental class, the unidiff parameters reported do not suggest any systematic trend across cohorts. For parental status, the association with educational attainment appears to strengthen only slightly between the 1958 and 1970 cohorts, and then quite substantially between the 1970 and 1990 cohorts. For parental education, the unidiff parameters indicate no change between the 1958 and 1970 cohorts, and then an increase in the association with education for the 1990 cohort.

Sets of identical log-linear models were then run using the NCDS, BCS70 and LSYPE datasets in order to examine whether the strength of associations identified differ when using the nationally-representative 1990s cohort. The results of these analyses are given in Table 30. The over-time trends we observe are broadly similar to those shown above using ALSPAC. While the unidiff parameters confirm the strengthening associations for parental status and parental education among the 1990 cohort, this is less pronounced among the LSYPE sample. For parental income, we find some strengthening in the association with educational attainment between the 1970 and 1990 cohorts.

**Table 30.** Log-Linear Models of Association between Social Origins and Educational Attainment (Age 20) across NCDS, BCS70 and LSYPE

	<b>G<sup>2</sup></b>	<b>% total association explained</b>	<b>df.</b>	<b>p</b>	<b>DI</b>	<b>N</b>
<b>Class–Educational attainment</b>						
(1) Conditional independence	316493.3		72	0.00	12.9	2778762
(2) Constant association	7668.0	97.58%	48	0.00	1.8	2778762
(3) Unidiff for OE	7304.6	97.69%	46	0.00	1.8	2778762
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>0.95 (BCS70)</i>	<i>1.04 (LSYPE)</i>			
(3) vs. (2)	363.4	4.74%	2	0.01		
<b>Status–Educational attainment</b>						
(1) Conditional independence	285836.6		36	0.00	12.3	2778762
(2) Constant association	18987.1	93.36%	24	0.00	2.9	2778762
(3) Unidiff for OE	12769.3	95.53%	22	0.00	2.5	2778762
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>1.06 (BCS70)</i>	<i>1.47 (LSYPE)</i>			
(3) vs. (2)	6217.8	32.75%	2	0.01		
<b>Education–Educational attainment</b>						
(1) Conditional independence	393903.0		72	0.00		2853162
(2) Constant association	11663.7	97.04%	48	0.00		2853162
(3) Unidiff for OE	10504.4	97.33%	46	0.00		2853162
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>0.96 (BCS70)</i>	<i>1.12 (LSYPE)</i>			
(3) vs. (2)	1159.3	9.94%	2	0.01		
<b>Income–Educational attainment</b>						
(1) Conditional independence	110485.6		32	0.00	11.5	1160003
(2) Constant association	6174.8	94.41%	16	0.00	2.1	1160003
(3) Unidiff	5795.6	94.75%	15	0.00	1.9	1160003
<i>Unidiff parameters</i>		<i>1.00 (BCS70)</i>	<i>1.16 (LSYPE)</i>			
(2) vs. (3)	379.2	6.14%	1	0.01		

Notes: N is weighted. Weight=1 for NCDS and BCS70 and =wave 7 weight for LSYPE. Weights have been multiplied by 100 and rounded, to permit use of appropriate Stata commands which prepare the data for log-linear modelling. N is very high for this reason.

Log-linear models were then run using the alternative measures of educational attainment described in Section 7. Table 31 uses NCDS, BCS70 and LSYPE data to examine the changing strength of associations between social origin measures and academic-only educational attainment.<sup>3</sup> In terms of model fit, we find the same patterns as in Table 30; the unidiff models are the best fitting, suggesting some change that is of a uniform kind. In the case of parental class, we again see no strong systematic change across cohorts; however, we do find that the 1990 cohort, compared with the 1958 cohort, now has a slightly weaker association. Differences in the unidiff parameters are not large (0.09) and we do not know if their comparison intervals overlap.

In terms of parental status, we see, similar to Table 30, a weakening association between the 1958 and 1970 cohorts, and then a strengthening association for the 1990 cohort; the unidiff parameter for LSYPE is, however, not as large as we find when educational attainment is measured as the achievement of both academic and vocational qualifications. In the case of parental education, we find a slight weakening in the association between the 1958 and 1970 cohorts, not dissimilar to what is shown in Table 30, and then a strengthening association, so that the 1958 and 1990 cohorts have unidiff parameters which are almost equal. For parental income, we again find a strengthening association between the 1970 and 1990 cohorts.

<sup>3</sup> ALSPAC could not be used for this exercise.

Taken together, the associations between our three main social origin variables and educational attainment suggest less over-time change across the three cohorts than is the case when we consider the achievement of both academic and vocational qualifications together.

**Table 31.** Log-Linear Models of Association between Social Origins and Educational Attainment (Academic Only, Age 20) across NCDS, BCS70 and LSYPE

	<b>G<sup>2</sup></b>	<b>% total association explained</b>	<b>df.</b>	<b>p</b>	<b>DI</b>	<b>N</b>
<b>Class–Educational attainment</b>						
(1) Conditional independence	345610.9		72	0.00	13.4	2778579
(2) Constant association	11066.4	96.80%	48	0.00	2.1	2778579
(3) Unidiff for OE	10358.6	97.00%	46	0.00	2.1	2778579
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>0.89 (BCS70)</i>	<i>0.91 (LSYPE)</i>			
(3) vs. (2)	707.8	6.40%	2	0.01		
<b>Status–Educational attainment</b>						
(1) Conditional independence	306907.1		36	0.00	12.7	2778579
(2) Constant association	16381.7	94.66%	24	0.00	2.6	2778579
(3) Unidiff for OE	12835.1	95.82%	22	0.00	2.4	2778579
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>0.98 (BCS70)</i>	<i>1.30 (LSYPE)</i>			
(3) vs. (2)	3546.6	21.65%	2	0.01		
<b>Education–Educational attainment</b>						
(1) Conditional independence	430975.3		72	0.00	13.9	2852979
(2) Constant association	12255.8	97.16%	48	0.00	2.3	2852979
(3) Unidiff for OE	11453.6	97.34%	46	0.00	2.2	2852979
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>0.90 (BCS70)</i>	<i>0.99 (LSYPE)</i>			
(3) vs. (2)	802.2	6.55%	2	0.01		
<b>Income–Educational attainment</b>						
(1) Conditional independence	110910.4		32	0.00	11.6	1159820
(2) Constant association	6313.8	94.31%	16	0.00	2.2	1159820
(3) Unidiff	5747.9	94.82%	15	0.00	1.8	1159820
<i>Unidiff parameters</i>		<i>1.00 (BCS70)</i>	<i>1.21 (LSYPE)</i>			
(2) vs. (3)	565.9	8.96%	1	0.01		

Notes: N is weighted. Weight=1 for NCDS and BCS70 and =wave 7 weight for LSYPE. Weights have been multiplied by 100 and rounded, to permit use of appropriate Stata commands which prepare the data for log-linear modelling. N is very high for this reason.

Identical log-linear models were run using only the NCDS and BCS70 data to examine changes in the strength of associations over time between social origins and educational attainment (both academic and vocational), when measured at age 38. The results of these analyses are given in Table 32. When educational attainment at age 38 is considered, the unidiff models do not significantly improve upon the constant association models for any of the social origin measures considered; the results indicate that the strength of the associations did not change between the 1958 and 1970 cohorts.

**Table 32.** Log-Linear Models of Association between Social Origins and Educational Attainment (Age 38) across NCDS and BCS70

	<b>G<sup>2</sup></b>	<b>% total association explained</b>	<b>df.</b>	<b>p</b>	<b>DI</b>	<b>N</b>
<b>Class–Educational attainment</b>						
(1) Conditional independence	1973.7		72	0.00	12.3	18328
(2) Constant association	51.0	97.42%	36	0.05	2.0	18328
(3) Unidiff for OE	49.9	97.47%	35	0.05	2.0	18328
<i>Unidiff parameters</i>	1.00 (NCDS)	0.95 (BCS70)				
(3) vs. (2)	1.1	2.16%	1	0.29		
<b>Status–Educational attainment</b>						
(1) Conditional independence	1697.6		36	0.00	11.4	18328
(2) Constant association	29.3	98.27%	18	0.05	1.6	18328
(3) Unidiff for OE	27.7	98.37%	17	0.05	1.5	18328
<i>Unidiff parameters</i>	1.00 (NCDS)	1.07 (BCS70)				
(3) vs. (2)	1.6	5.46%	1	0.21		
<b>Education–Educational attainment</b>						
(1) Conditional independence	2723.4		72	0.00	12.7	18972
(2) Constant association	71.0	97.39%	36	0.00	2.1	18972
(3) Unidiff for OE	70.9	97.40%	35	0.00	2.0	18972
<i>Unidiff parameters</i>	1.00 (NCDS)	1.02 (BCS70)				
(3) vs. (2)	0.1	0.14%	1	0.75		

Finally, we use the Cox–Jackson (Cox and Jackson 2009) method to take an alternative approach to examining the changing strength of associations between social origins and educational attainment over time. For this exercise, social origins – by – educational attainment associations are considered for each pair of cohorts in turn. Table 33 presents the results.

In the case of parental class, the patterns described above are repeated here; a slight weakening in the association with educational attainment between the 1958 and 1970 cohorts, and then a strengthening association for the 1990 cohorts. Using this method, we can further infer that differences in the strength of associations between cohorts are statistically significant. In the case of parental status, we again find a slightly stronger association in the 1970 cohort than in the 1958 cohort, and then a marked increase in the strength of the association among the 1990 cohorts.

For parental education, contrary to the largely stable association between the 1958 and 1970 cohorts implied by the log-linear models presented above, we here find that the association weakens across this period. Also in contrast to what is shown above, the association between parental education and educational attainment is shown to be stronger among the LSYPE sample than the ALSPAC sample. The association between parental income and educational attainment is again shown to be stronger for the 1990 cohorts.



**Table 33.** Average Centre Four Global Log Odds Ratios and Test of Significance in the Differences between Cohorts

		Ave. centre four		Difference	SE	Lower CI	Upper CI
		I.	II.				
CLASS	I.NCDS – II.BCS70	0.99	0.87	-0.12**	0.06	-0.23	-0.01
	I.BCS70 – II.LSYPE	0.87	1.19	0.32**	0.06	0.20	0.45
	I.NCDS – II.LSYPE	0.99	1.19	0.20**	0.06	0.08	0.33
	I.BCS70 – II.ALSPAC	0.87	1.14	0.27**	0.06	0.16	0.38
	I.NCDS – II.ALSPAC	0.99	1.14	0.15**	0.06	0.04	0.27
STATUS	I.NCDS – II.BCS70	0.72	0.85	0.13**	0.05	0.03	0.23
	I.BCS70 – II.LSYPE	0.85	1.33	0.48**	0.07	0.35	0.61
	I.NCDS – II.LSYPE	0.72	1.33	0.61**	0.07	0.48	0.74
	I.BCS70 – II.ALSPAC	0.85	1.38	0.53**	0.06	0.42	0.65
	I.NCDS – II.ALSPAC	0.72	1.38	0.66**	0.06	0.55	0.77
EDUCATION	I.NCDS – II.BCS70	1.47	1.18	-0.29**	0.08	-0.46	-0.13
	I.BCS70 – II.LSYPE	1.18	1.72	0.55**	0.09	0.36	0.73
	I.NCDS – II.LSYPE	1.47	1.72	0.25**	0.10	0.05	0.46
	I.BCS70 – II.ALSPAC	1.18	1.57	0.39**	0.07	0.25	0.54
	I.NCDS – II.ALSPAC	1.47	1.57	0.10	0.09	-0.07	0.27
INCOME	I.BCS70 – II.LSYPE	0.82	1.03	0.21**	0.08	0.06	0.36
	I.BCS70 – II.ALSPAC	0.82	1.15	0.33**	0.08	0.18	0.48

## 10. Social Origins and Cognitive Ability

This section examines the relationships between social origins and cognitive ability. Tables 34–37 present cross-classifications between **social class** and cognitive ability quintiles in each of the datasets.

**Table 34.** Cross-tabulation of Social Class by Cognitive Ability (Row %) – NCDS

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
1. Higher managerial	39	27	18	11	5	100
2. Lower managerial	35	26	20	12	8	100
3. Intermediate	28	22	22	16	12	100
4. Small employers	19	20	24	22	17	100
5. Lower supervisory	18	21	20	21	20	100
6. Semi routine	15	19	21	23	23	100
7. Routine	11	15	20	24	29	100
<i>Total</i>	21	20	21	20	19	100

**Table 35.** Cross-tabulation of Social Class by Cognitive Ability (Row %) – BCS70

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
1. Higher managerial	38	26	17	11	7	100
2. Lower managerial	32	24	18	16	9	100
3. Intermediate	23	23	22	18	15	100
4. Small employers	20	21	23	19	17	100
5. Lower supervisory	16	18	21	23	22	100
6. Semi routine	11	17	20	24	29	100
7. Routine	7	16	21	24	32	100
<i>Total</i>	20	20	20	20	19	100

**Table 36.** Cross-tabulation of Social Class by Cognitive Ability (Row %) – LSYPE

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
1. Higher managerial	34	25	19	12	10	100
2. Lower managerial	22	22	21	20	15	100
3. Intermediate	20	22	17	20	22	100
4. Small employers	14	18	20	25	24	100
5. Lower supervisory	10	14	20	27	28	100
6. Semi routine	9	13	17	25	36	100
7. Routine	7	12	20	22	38	100
<i>Total</i>	18	19	20	21	23	100

**Table 37.** Cross-tabulation of Social Class by Cognitive Ability (Row %) – ALSPAC

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
1. Higher managerial	31	25	20	15	9	100
2. Lower managerial	24	22	20	18	16	100
3. Intermediate	19	19	21	21	21	100
4. Small employers	11	21	23	22	22	100
5. Lower supervisory	9	17	21	26	27	100
6. Semi routine	9	12	21	27	31	100
7. Routine	9	13	17	24	37	100
<i>Total</i>	20	20	20	20	20	100

Tables 38–41 present cross-classifications of **parental social status** and cognitive ability across each dataset.

**Table 38.** Cross-tabulation of Social Status Quartiles by Cognitive Ability (Row %) – NCDS

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
4. Top	36	26	20	13	6	100
3. 3 <sup>rd</sup>	21	20	22	20	17	100
2. 2 <sup>nd</sup>	15	19	21	21	23	100
1. Bottom	15	18	20	23	24	100
<i>Total</i>	21	20	21	20	19	100

**Table 39.** Cross-tabulation of Social Status Quartiles by Cognitive Ability (Row %) – BCS70

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
4. Top	37	25	17	13	7	100
3. 3 <sup>rd</sup>	21	22	21	20	16	100
2. 2 <sup>nd</sup>	15	18	21	23	23	100
1. Bottom	10	17	21	24	29	100
<i>Total</i>	20	20	20	20	19	100

**Table 40.** Cross-tabulation of Social Status Quartiles by Cognitive Ability (Row %) – LSYPE

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
4. Top	32	26	19	14	9	100
3. 3 <sup>rd</sup>	22	22	21	19	17	100
2. 2 <sup>nd</sup>	14	17	20	24	25	100
1. Bottom	7	12	18	26	36	100
<i>Total</i>	18	19	20	21	23	100

**Table 41.** Cross-tabulation of Social Status Quartiles by Cognitive Ability (Row %) – ALSPAC

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
4. Top	33	24	20	14	9	100
3. 3 <sup>rd</sup>	23	22	21	17	16	100
2. 2 <sup>nd</sup>	14	19	20	26	22	100
1. Bottom	8	13	19	25	35	100
<i>Total</i>	20	20	20	20	20	100

Tables 42–45 give cross-classifications of **parental education** and respondents' cognitive ability across each dataset.

**Table 42.** Cross-tabulation of Parental Education by Cognitive Ability (Row %) – NCDS

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
7. Both have degrees	50	26	13	5	6	100
6.	42	25	20	9	4	100
5.	42	26	19	9	5	100
4.	35	24	21	12	8	100
3.	21	21	21	20	17	100
2.	19	21	21	20	19	100
1. No qualifications	13	18	21	24	24	100
<i>Total</i>	20	20	21	20	19	100

**Table 43.** Cross-tabulation of Parental Education by Cognitive Ability (Row %) – BCS70

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
7. Both have degrees	64	22	8	5	1	100
6.	44	26	17	9	4	100
5.	41	30	13	10	6	100
4.	26	25	21	16	11	100
3.	22	23	21	20	15	100
2.	15	20	22	21	22	100
1. No qualifications	9	14	20	25	31	100
<i>Total</i>	20	20	20	20	19	100

**Table 44.** Cross-tabulation of Parental Education by Cognitive Ability (Row %) – LSYPE

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
7. Both have degrees	46	25	13	9	8	100
6.	34	26	18	16	6	100
5.	29	34	19	11	8	100
4.	19	23	22	19	17	100
3.	12	17	21	25	26	100
2.	8	13	18	28	33	100
1. No qualifications	7	9	17	22	46	100
<i>Total</i>	18	19	20	21	23	100

**Table 45.** Cross-tabulation of Parental Education by Cognitive Ability (Row %) – ALSPAC

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
7. Both have degrees	47	23	15	10	4	100
6.	29	22	22	15	11	100
5.	23	26	19	17	14	100
4.	14	22	23	21	20	100
3.	11	17	21	25	27	100
2.	4	12	17	30	38	100
1. No qualifications	4	7	15	19	55	100
<i>Total</i>	20	20	20	20	20	100

Tables 46–48 give cross-classifications of **parental income** and respondents' cognitive ability for BCS70, LSYPE and ALSPAC.

**Table 46.** Cross-tabulation of Parental Income Quintiles by Cognitive Ability (Row %) – BCS70

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
5. Top	38	26	18	11	7	100
4. 4 <sup>th</sup>	27	26	18	17	11	100
3. 3 <sup>rd</sup>	22	23	21	19	16	100
2. 2 <sup>nd</sup>	18	18	21	20	23	100
1. Bottom	12	16	18	26	28	100
<i>Total</i>	23	21	20	19	17	100

**Table 47.** Cross-tabulation of Parental Income Quintiles by Cognitive Ability (Row %) – LSYPE

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
5. Top	32	25	20	13	10	100
4. 4 <sup>th</sup>	20	22	22	20	16	100
3. 3 <sup>rd</sup>	16	17	20	23	24	100
2. 2 <sup>nd</sup>	13	15	17	23	32	100
1. Bottom	12	14	19	23	32	100
<i>Total</i>	18	19	20	21	23	100

**Table 48.** Cross-tabulation of Parental Income Quintiles by Cognitive Ability (Row %) – ALSPAC

	5. Top	4. 4th	3. 3rd	2. 2nd	1. Bottom	<i>Total</i>
5. Top	36	24	18	14	7	100
4. 4 <sup>th</sup>	25	23	20	18	14	100
3. 3 <sup>rd</sup>	19	21	21	22	18	100
2. 2 <sup>nd</sup>	14	19	20	24	23	100
1. Bottom	13	15	21	22	29	100
<i>Total</i>	21	21	20	20	18	100

5x5 contingency tables were produced for all social origin-by-cognitive ability pairs. Differences between the averages of the centre four global log odds ratios were calculated for each pair of tables, and were then tested for significance. The results of each of the tests are given in Table 49.

Similar to what is found for educational attainment, the average centre four global log odds ratios suggest the overall association between cognitive ability and parental education is stronger than the associations between cognitive ability and any other measure of social origin. Similarly, the association with parental income is consistently the weakest. Differences in associations tend, on the whole, to be statistically significant, providing some justification for treating each measure of social origin independently.

**Table 49.** Average Centre Four Global Log Odds Ratios and Test of Significance in the Differences between Social Origin Measures

		Ave. centre four		Difference	SE	Lower CI	Upper CI
		I.	II.				
NCDS	I.Class – II.Status	1.00	0.71	0.30**	0.05	-0.40	-0.19
	I.Class – II.Education	1.00	1.34	0.34**	0.08	0.19	0.49
	I.Status – II.Education	0.71	1.34	0.63**	0.08	0.49	0.78
BCS70	I.Class – II.Status	1.00	1.00	0.00	0.05	-0.10	0.10
	I.Class – II.Education	1.00	1.48	0.44**	0.06	0.35	0.60
	I.Class – II.Income	1.00	0.96	-0.05	0.07	-0.17	0.08
	I.Status – II.Education	1.00	1.48	0.48**	0.06	0.35	0.60
	I.Status – II.Income	1.00	0.96	-0.05	0.06	-0.17	0.08
	I.Education – II.Income	1.48	0.96	-0.52**	0.07	-0.67	-0.40
ALSPAC	I.Class – II.Status	0.99	1.06	0.07	0.07	-0.06	0.21
	I.Class – II.Education	0.99	1.16	0.17**	0.07	0.03	0.30
	I.Class – II.Income	0.99	0.78	-0.21**	0.07	-0.35	-0.08
	I.Status – II.Education	1.06	1.16	0.10	0.07	-0.04	0.23
	I.Status – II.Income	1.06	0.78	-0.29**	0.07	-0.42	-0.15
	I.Education – II.Income	1.16	0.78	-0.38**	0.07	-0.51	-0.25
LSYPE	I.Class – II.Status	0.98	1.13	0.15**	0.06	0.03	0.27
	I.Class – II.Education	0.98	1.35	0.36**	0.07	0.24	0.49
	I.Class – II.Income	0.98	0.82	-0.16**	0.07	-0.29	-0.03
	I.Status – II.Education	1.13	1.35	0.22**	0.07	0.09	0.34
	I.Status – II.Income	1.13	0.82	-0.31**	0.07	-0.45	-0.19
	I.Education – II.Income	1.35	0.82	-0.53**	0.07	-0.66	-0.39

### 10.1 Changes in the Associations between Social Origins and Cognitive Ability across Cohorts

A series of log-linear models were run to examine whether the strength of associations between social origins and cognitive ability change across cohorts. The results of these analyses are given in Table 50. None of the models provide a good fit to the data, though for each social origin measure considered the unidiff model is found to significantly improve upon the constant association models. The association between class and cognitive ability scarcely changes between the 1958 and 1970 cohorts, and then declines a little for the 1990 cohort. The associations between status and cognitive ability and parental education and cognitive ability, on the other hand, both strengthen between the 1958 and 1970 cohorts, and are then stable between the 1970 and 1990 cohorts.

**Table 50.** Log-Linear Models of Association between Social Origins and Cognitive Ability across NCDS, BCS70 and ALSPAC

	<b>G<sup>2</sup></b>	<b>% total association explained</b>	<b>df.</b>	<b>p</b>	<b>DI</b>	<b>N</b>
<b>Class– Cognitive ability</b>						
(1) Conditional independence	2751.0		72	0.00	11.7	26904
(2) Constant association	75.5	97.26%	48	0.01	1.9	26904
(3) Unidiff for OE	68.2	97.52%	46	0.02	1.8	26904
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>1.05 (BCS70)</i>	<i>0.91 (ALSPAC)</i>			
(3) vs. (2)	7.3	9.67%	2	0.03		
<b>Status– Cognitive ability</b>						
(1) Conditional independence	2502.6		36	0.00	11.0	26916
(2) Constant association	128.4	94.87%	24	0.00	2.5	26916
(3) Unidiff for OE	108.2	95.68%	22	0.00	2.4	26916
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>1.24 (BCS70)</i>	<i>1.24 (ALSPAC)</i>			
(3) vs. (2)	20.2	15.73%	2	0.00		
<b>Education– Cognitive ability</b>						
(1) Conditional independence	3512.1		72	0.00		26908
(2) Constant association	135.6	96.14%	48	0.00		26908
(3) Unidiff for OE	95.9	97.27%	46	0.00		26908
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>1.33 (BCS70)</i>	<i>1.34 (ALSPAC)</i>			
(3) vs. (2)	39.7	29.28%	2	0.00		

This exercise was then repeated using NCDS, BCS70 and LSYPE data. Note, however, that we do not consider the Key Stage 2 scores in LSYPE as an adequate proxy for cognitive ability. The results are produced in Table 51 below. In the case of parental class, the results of Table 50 are essentially unchanged. In the case of parental status, we now find an increase in the strength of the association between the 1970 and 1990 cohorts. In the case of parental income, we find a weakening association between the 1970 and 1990 cohorts.

**Table 51.** Log-Linear Models of Association between Social Origins and Cognitive Ability across NCDS, BCS70 and LSYPE

	<b>G<sup>2</sup></b>	<b>% total association explained</b>	<b>df.</b>	<b>p</b>	<b>DI</b>	<b>N</b>
<b>Class–Cognitive ability</b>						
(1) Conditional independence	295658.4		72	0.00	12.0	2801655
(2) Constant association	8869.6	97.00%	48	0.00	2.1	2801655
(3) Unidiff for OE	8487.5	97.13%	46	0.00	2.1	2801655
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>1.05 (BCS70)</i>	<i>0.96 (LSYPE)</i>			
(3) vs. (2)	382.1	4.31%	2	0.01		
<b>Status–Cognitive ability</b>						
(1) Conditional independence	265830.1		36	0.00	11.2	2802855
(2) Constant association	12449.0	95.32%	24	0.00	2.4	2802855
(3) Unidiff for OE	9706.9	96.35%	22	0.00	2.2	2802855
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>1.24 (BCS70)</i>	<i>1.31 (LSYPE)</i>			
(3) vs. (2)	2742.1	22.03%	2	0.01		
<b>Education–Cognitive ability</b>						
(1) Conditional independence	--	--	--	--	--	--
(2) Constant association						
(3) Unidiff for OE						
<i>Unidiff parameters</i>						
(3) vs. (2)						
<b>Income–Cognitive ability</b>						
(1) Conditional independence	95711.3		32	0.00	11.1	1117225
(2) Constant association	3566.4	96.27%	16	0.00	2.2	1117225
(3) Unidiff for OE	3035.8	96.83%	15	0.00	2.2	1117225
<i>Unidiff parameters</i>		<i>1.00 (BCS70)</i>	<i>0.85 (LSYPE)</i>			
(3) vs. (2)	530.6	14.88%	1	0.01		

Notes: N is weighted. Weight=1 for NCDS and BCS70 and =wave 7 weight for LSYPE. Weights have been multiplied by 100 and rounded, to permit use of appropriate Stata commands which prepare the data for log-linear modelling. N is very high for this reason. Model for parental education failed to converge.

The Cox–Jackson (Cox and Jackson 2009) method was then used to examine the changing strength of associations between social origins and cognitive ability over time by considering each pair of cohorts in turn. The results are given in Table 52. For parental class, the strength of association appears essentially stable over time, regardless of which sample is used to represent the 1990 cohort. For parental status, we find an increase in the strength of association with cognitive ability between the 1958 and 1970 cohorts, as shown with the log-linear models presented above, then a slight strengthening again among the ALSPAC sample, and a more marked increase among the LSYPE sample. In the case of parental education, the Cox method implies a different over-time trend to that presented in Table 50 above; that is, we now see an increase in the strength of the association between the 1958 and 1970 cohorts, and then a weakening association for the 1990 cohorts (which is most marked for the ALSPAC sample). For parental income, we again find a weakening association with cognitive ability.



**Table 52.** Average Centre Four Global Log Odds Ratios and Test of Significance in the Differences between Cohorts

		Ave. centre four		Difference	SE	Lower CI	Upper CI
		I.	II.				
CLASS	I.NCDS – II.BCS70	1.00	1.00	-0.00	0.06	-0.11	0.10
	I.BCS70 – II.LSYPE	1.00	0.98	-0.02	0.06	-0.13	0.10
	I.NCDS – II.LSYPE	1.00	0.98	-0.02	0.06	-0.14	0.10
	I.BCS70 – II.ALSPAC	1.00	0.99	-0.01	0.06	-0.13	0.11
	I.NCDS – II.ALSPAC	1.00	0.99	-0.01	0.06	-0.14	0.11
STATUS	I.NCDS – II.BCS70	0.71	1.00	0.29**	0.05	0.19	0.39
	I.BCS70 – II.LSYPE	1.00	1.13	0.13**	0.06	0.02	0.24
	I.NCDS – II.LSYPE	0.71	1.13	0.42**	0.06	0.31	0.53
	I.BCS70 – II.ALSPAC	1.00	1.06	0.06	0.06	-0.05	0.18
	I.NCDS – II.ALSPAC	0.71	1.06	0.35**	0.06	0.24	0.47
EDUCATION	I.NCDS – II.BCS70	1.34	1.48	0.14	0.08	-0.03	0.30
	I.BCS70 – II.LSYPE	1.48	1.35	-0.13	0.07	-0.27	0.01
	I.NCDS – II.LSYPE	1.34	1.35	0.00	0.08	-0.16	0.16
	I.BCS70 – II.ALSPAC	1.48	1.16	-0.32**	0.07	-0.46	-0.18
	I.NCDS – II.ALSPAC	1.34	1.16	-0.18**	0.08	-0.34	-0.03
INCOME	I.BCS70 – II.LSYPE	0.96	0.82	-0.14	0.07	-0.27	0.00
	I.BCS70 – II.ALSPAC	0.96	0.78	-0.18**	0.07	-0.32	-0.04

## 11. Cognitive Ability and Educational Attainment

This section examines the relationship between respondents' early-life cognitive ability and their later-life educational attainment. Tables 53–56 cross-classify these measures in each dataset. There is little change in the proportions achieving the highest educational qualifications for each quintile of cognitive ability between the 1958 and 1970 cohorts; however, for the 1990s cohort these proportions almost double.

**Table 53.** Cross-tabulation of Cognitive Ability by Educational Attainment (Row %) – NCDS

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	Total
5. Top	46	25	16	2	11	100
4. 4 <sup>th</sup>	26	26	23	7	17	100
3. 3 <sup>rd</sup>	15	21	28	12	23	100
2. 2 <sup>nd</sup>	7	16	26	18	34	100
1. Bottom	2	6	16	15	60	100
<i>Total</i>	20	19	22	11	29	100

**Table 54.** Cross-tabulation of Cognitive Ability by Educational Attainment (Row %) – BCS70

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	Total
5. Top	44	31	10	2	12	100
4. 4 <sup>th</sup>	23	32	20	5	19	100
3. 3 <sup>rd</sup>	15	30	27	8	20	100
2. 2 <sup>nd</sup>	9	21	30	12	28	100
1. Bottom	5	14	26	15	40	100
<i>Total</i>	20	26	22	8	23	100

**Table 55.** Cross-tabulation of Cognitive Ability by Educational Attainment (Row %) – LSYPE

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	Total
5. Top	85	14	--	--	--	100
4. 4 <sup>th</sup>	61	30	--	--	--	100
3. 3 <sup>rd</sup>	37	42	15	--	--	100
2. 2 <sup>nd</sup>	21	36	26	16	1	100
1. Bottom	7	21	26	41	5	100
<i>Total</i>	40	29	16	14	2	100

Notes: -- Values omitted to preserve anonymity

**Table 56.** Cross-tabulation of Cognitive Ability by Educational Attainment (Row %) – ALSPAC

	5. 2+ A- levels	4.	3.	2.	1. No qualifications	Total
5. Top	85	13	2	0	0	100
4. 4 <sup>th</sup>	66	26	7	1	0	100
3. 3 <sup>rd</sup>	52	31	13	3	1	100
2. 2 <sup>nd</sup>	37	26	28	8	2	100
1. Bottom	15	18	33	29	5	100
<i>Total</i>	50	23	17	8	2	100

Log-linear models were run in order to examine whether the association between cognitive ability and educational attainment changes across cohorts. The results of these analyses are given in Table 51. The baseline model of conditional independence gives a very poor fit to the data, misclassifying 20.8% of cases. The lack of fit is much reduced with the constant association model and even further reduced for the unidiff model, though still neither of these provides a good fit to the data by conventional standards.

The unidiff parameters suggest that the strength of association is weaker for the 1970 cohort than for the 1958 cohort, but is stronger for the 1990 cohort than for the 1958 cohort.

**Table 57.** Log-Linear Models of Association between Cognitive Ability and Educational Attainment (Age 20) across NCDS, BCS70 and ALSPAC

	<b>G<sup>2</sup></b>	<b>% total association explained</b>	<b>df.</b>	<b>p</b>	<b>DI</b>	<b>N</b>
<b>Cognitive ability–Educational attainment</b>						
(1) Conditional independence	7625.8		48	0.00	20.8	26909
(2) Constant association	370.8	95.14%	32	0.00	3.9	26909
(3) Unidiff for OE	168.2	97.79%	30	0.00	2.9	26909
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>0.75 (BCS70)</i>	<i>1.35 (ALSPAC)</i>			
(3) vs. (2)	202.6	54.64%	2	0.00		

Table 58 presents the results of similar models using NCDS, BCS70 and LSYPE. We find that the conditional independence model misclassifies 22.2% of all individual cases. The constant association model improves upon the conditional independence model, but still has a poor fit by conventional standards. The unidiff model improves upon the constant association model. The beta parameters indicate a decreasing association between cognitive ability and educational attainment between the 1958 and 1970 cohorts, and then a substantially stronger association for the 1990 cohort.

**Table 58.** Log-Linear Models of Association between Cognitive Ability and Educational Attainment (Age 20) across NCDS, BCS70 and LSYPE

	<b>G<sup>2</sup></b>	<b>% total association explained</b>	<b>df.</b>	<b>p</b>	<b>DI</b>	<b>N</b>
<b>Cognitive ability–Educational attainment</b>						
(1) Conditional independence	934978.5		48	0.00	22.2	2876055
(2) Constant association	58155.1	93.78%	32	0.00	4.9	2876055
(3) Unidiff for OE	14786.4	98.42%	30	0.00	2.4	2876055
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>0.75 (BCS70)</i>	<i>1.63 (LSYPE)</i>			
(3) vs. (2)	43368.7	74.57%	2	0.01		

Notes: N is weighted. Weight=1 for NCDS and BCS70 and =wave 7 weight for LSYPE. Weights have been multiplied by 100 and rounded, to permit use of appropriate Stata commands which prepare the data for log-linear modelling. N is very high for this reason.

Table 59 presents the results of these models using the educational attainment at age 38 measure. Again, we observe that the unidiff model is a better fit to the data than the constant association model, and that there is some weakening in the association between these two cohorts.

**Table 59.** Log-Linear Models of Association between Cognitive Ability and Educational Attainment (Age 38) across NCDS and BCS70

	G <sup>2</sup>	% total association explained	df.	p	DI	N
<b>Cognitive ability–Educational attainment</b>						
(1) Conditional independence	5106.0		48	0.00	19.3	18999
(2) Constant association	130.4	97.45%	24	0.00	2.8	18999
(3) Unidiff for OE	73.3	98.56%	23	0.00	2.2	18999
<i>Unidiff parameters</i>	<i>1.00 (NCDS)</i>	<i>0.78 (BCS70)</i>				
(3) vs. (2)	57.1	43.79%	1	0.00		

Finally, the 5x5 contingency tables of cognitive ability quintiles by educational attainment were tested in cohort-pairs for the significance of differences in their average centre four global log odds ratios. The results are given in Table 60. As above, we find a weakening in the association between cognitive ability and educational attainment between the 1958 and 1970 cohorts, and then a substantial strengthening for the 1990 cohorts. This is most pronounced for the LSYPE dataset, which uses Key Stage 2 instead of actual cognitive ability test scores, the former of which we have elsewhere shown to be a better predictor of educational attainment.

**Table 60.** Average Centre Four Global Log Odds Ratios and Tests of Significance in the Differences, For All Cohort Pairs – Cognitive Ability and Educational Attainment

		Ave centre four		Difference	SE	Lower CI	Upper CI
		I.	II.				
COGNITIVE ABILITY	I.NCDS–II.BCS70	1.73	1.24	-0.49**	0.05	-0.60	-0.39
	I.BCS70–II.LSYPE	1.24	2.87	1.64**	0.11	1.43	1.85
	I.NCDS–II.LSYPE	1.73	2.87	1.14**	0.11	0.94	1.35
	I.BCS70–II.ALSPAC	1.24	2.61	1.38**	0.12	1.14	1.62
	I.NCDS–II.ALSPAC	1.73	2.61	0.89**	0.12	0.65	1.12

## Appendix

### **ALSPAC Study description**

ALSPAC recruited 14,541 pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992. 14,541 is the initial number of pregnancies for which the mother enrolled in the ALSPAC study and had either returned at least one questionnaire or attended a “Children in Focus” clinic by 19/07/99. Of these initial pregnancies, there was a total of 14,676 fetuses, resulting in 14,062 live births and 13,988 children who were alive at 1 year of age.

When the oldest children were approximately 7 years of age, an attempt was made to bolster the initial sample with eligible cases who had failed to join the study originally. As a result, when considering variables collected from the age of seven onwards (and potentially abstracted from obstetric notes) there are data available for more than the 14,541 pregnancies mentioned above.

The number of new pregnancies not in the initial sample (known as Phase I enrolment) that are currently represented on the built files and reflecting enrolment status at the age of 18 is 706 (452 and 254 recruited during Phases II and III respectively), resulting in an additional 713 children being enrolled. The phases of enrolment are described in more detail in the ALSPAC cohort profile paper which should be used for referencing purposes.

The total sample size for analyses using any data collected after the age of seven is therefore 15,247 pregnancies, resulting in 15,458 fetuses. Of this total sample of 15,458 fetuses, 14,775 were live births and 14,701 were alive at 1 year of age.

A 10% sample of the ALSPAC cohort, known as the Children in Focus (CiF) group, attended clinics at the University of Bristol at various time intervals between 4 to 61 months of age. The CiF group were chosen at random from the last 6 months of ALSPAC births (1432 families attended at least one clinic). Excluded were those mothers who had moved out of the area or were lost to follow-up, and those partaking in another study of infant development in Avon.

### **ALSPAC data dictionary**

The ALSPAC website contains details of all the data that is available through a fully searchable data dictionary: <http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary>.

### **Ethical approval**

Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.

### **Acknowledgements:**

This specific research for this paper was funded by the Nuffield Foundation. We are extremely grateful to all the families who took part in the NCDS, BCS70, LSYPE and ALSPAC studies and all teams and individuals involved in the development of these datasets, which include interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, nurses and midwives. The UK Medical Research Council and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors, who take responsibility for the contents of this paper.

### **References**

- Betthaeuser, B. and M. Bourne (2016a) "Harmonizing the measurement of social origin, cognitive ability and educational attainment in four British population surveys: NCDS, BCS70, LSYPE and ALSPAC.", *SOCED Project Data Note 2016-10-11*
- Betthaeuser, B. and M. Bourne (2016b) "Measuring social origin, cognitive ability and educational attainment in the British Birth Cohort Study (BCS70).", *SOCED Project Data Note 2016-10-04*.
- Betthaeuser, B. and M. Bourne (2016c) "Measuring social origin, cognitive ability and educational attainment in the National Child Development Study (NCDS).", *SOCED Project Data Note 2016-08-09*
- Bourne, M. and Betthaeuser, B. (2016) "Measuring social origin, cognitive ability and educational attainment in the Longitudinal Study of Young People in England (LSYPE).", *SOCED Project Data Note 2016-06-03*
- Bourne, M. (2016) "Measuring cognitive ability in the National Child Development Study (NCDS), the British Cohort Study (BCS70), the Longitudinal Study of Young People in England (LSYPE), and the Avon Study of Parents and Children (ALSPAC).", *SOCED Project Data Note 2016-07-24*
- Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J., Macleod, J., Molloy, L., Ness, A., Ring, S., Nelson, S. M. and Lawlor, D. A. (2012) 'Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC Mothers Cohort', *International Journal of Epidemiology* 42(1) 97-110

