Measuring social origin, cognitive ability and educational attainment in the 1958 National Child Development Study (NCDS)

Bastian Betthaeuser and Mollie Bourne
August 2016

SOCED Project Data Note 2016-08-09

Contents

List of Tables

List of Figures

1   Introduction

This data note was prepared for the research project on *Social Origins, Cognitive Ability and Educational Attainment: A Birth Cohort and Life-Course Perspective (SOCED)*. The document will present descriptive statistics for the key variables in the 1958 National Child Development Study (NCDS) data that are of relevance to the project. The main independent variables are measures of respondents' cognitive ability in childhood and the parental education, class, status and income. The key dependent variables are measures of respondents' highest qualification completed and measures indicating whether respondent's have crossed certain educational qualification thresholds, especially whether they have completed A-levels and whether they have completed a higher education qualification. Furthermore, measures of respondents' performance at Key Stage 4 (GCSEs) and Key Stage 5 will be used to identify the relative importance of primary and secondary effects of parental background on respondent's educational attainment.

Section 2 will provide a brief introduction to the NCDS dataset, including information on the size of the sample for analysis; a description is provided of the amount of missing data on each variable and the resulting final number of valid cases. Sections 3–6 will examine each of the dimensions of social origin in turn and describe their distributions. Section 7 will describe the distribution of respondents' cognitive ability in the sample, while Section 8 focuses on respondents' educational attainment. Sections 9 through 12 will use cross-tabulations and correlations to present the associations between the above measures.

2   The 1958 National Child Development Study (NCDS)

The 1958 National Child Development Study (NCDS) is a longitudinal study which began collecting data on about 17,500 babies born in England, Scotland, and Wales in the week of March 3rd to 9th in 1958.[1] Since the birth survey in 1958, there have been nine 'sweeps' of all cohort members at ages 7, 11, 16, 23, 33, 42, 46, 50 and 55 (at age 45 9,000 cohort members also participated in a special bio-medical survey). In the first three sweeps (at ages 7, 11 and 16), the target sample was augmented to include immigrants born in the same week.

At each sweep, different sources and methods were used to gather information on the cohort members. When CMs were children information was obtained through interviews with parents and teachers, while CMs carried out some simple tests. After the age of 16, CMs provided most of the information and at age 33, data from partners was also collected. In the latest 2013-2014 sweep, data collection involved a sequential mixed mode approach whereby cohort members were invited to complete a web survey and those not responding were invited to participate via telephone. More than 9,100 cohort members took part in total, with 66 per cent doing so online.

The questionnaires used for each NCDS sweep do not include a large number of repeated questions. This is because the NCDS is a cohort study rather than a panel study and focuses on the progress of the cohort members (CMs) over their whole life course. For each sweep, the information collected is relevant to the development of the cohort at that time point - the information collected when the CMs are seven-years old is therefore different to that collected when the CMs are aged 33. Some questions are repeated, however. These generally relate to the CM's education, employment status, family relationships, physical attributes, health status and housing.

---

[1] This overview of the NCDS data draws on information taken from the NCDS user guides and the CLS website (http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=724&sitesectiontitle=Welcome+to+the+1958+National+Child+Development+Study)

The data collections for sweep 5 contain employment histories in which occupations are coded according to various standard schemas, including SOC, KOS, CODOT, the EGP social class schema and the Registrar General's social class schema. There are also standard educational assessment scores such as the Southgate Group Reading Score and the Bristol Social Adjustment Guide included in the data files for sweeps 1 to 3. Data for a number of standard health indicators have also been collected over the course of the NCDS; among them, the Malaise Inventory and the General Health Questionnaire (GHQ) by Rutter et al.

There are over 16,000 NCDS variables in the nine full sweeps. The Online Data Dictionary is a searchable database containing a 'record' for every variable in all the main sweeps since birth. It is structured to reflect the data collection by subject within questionnaire / instrument and is also searchable using free text search. The exact question wording corresponding to each variable is shown, and one can get frequency counts for each value.

Table 1 shoes the development of the sample size across the different waves of the study. Table 2 shows the missingness on key variables for the study. As shown in Table 2, a substantial number of cases is missing information on the parental origin dimensions.

Table 1. NCDS Sample Size across Waves

Wave	Year	Age	N
Wave 0	1958	0	17,415
Wave 1	1965	7	15,425
Wave 2	1969	11	15,337
Wave 3	1974	16	14,654
Wave 4	1981	23	12,537
Wave 5	1991	33	11,469
Wave 6	1999-2000	41-42	11,419
Wave 7	2004-2005	46-47	9,534
Wave 8	2008-2009	50	9,790
Wave 9	2013-2014	55	9,137

Notes: These numbers are taken from the CLS website. The NCDS data file containing information from the first four sweeps has a sample size of 18558 cases. The higher number (compared to the Wave 1 sample) may be due to a boost sample that was added to the original sample and increases the number of combined cases in Waves 0, 1, 2 and 3.

Table 2. Missingness on Key Variables

	Per Variable: Missing		Cumulative: Missing		Cumulative: Remaining	
	N	**%**	**N**	**%**	**N**	**%**
NCDS sample included in Qualifications Data Set	——	100.00%	——	100.00%	14,637	100.00%
Parental class (Reduced method)	4,650	31.77%	4,650	31.77%	9,987	68.23%
Parental status (Chan/Goldthorpe)	4,650	31.77%	0	31.77%	9,987	68.23%
Parental education (Composite Measure)	4,124	28.18%	2,434	48.40%	7,553	51.60%
Cognitive Ability	2,175	14.86%	469	51.60%	7,084	48.40%
Highest qualification (At age 38)	2,784	19.02%	667	56.16%	6,417	43.84%

Notes: Variables listed in order of proportion of missing cases.

3   Parental Social Class

For the construction of the parental class variable we use the information on occupational group codes in the NCDS supplied in the data file prepared 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. As part of this project the questionnaire response text strings in the NCDS data were processed using the CASCOT (Computer Assisted Structured Coding Tool) programme to assign Standard Occupational Classification 2000 codes (SOC2000) and Standard Occupational Classification 90 (SOC90) codes to entries. Please see the Data Note prepared by Tim Morris on this procedure for further details.[2] Only information on fathers' occupation in 1969 (i.e. age 11 of cohort member) is available; no information is given on mothers' occupation. Fathers' occupation is therefore used as a proxy for parental occupation.

For the accurate coding of the NS-SEC measure for social class both information on occupation and on employment status is needed. Information on fathers' employment status is taken from the original NCDS data, Wave 2 (variables n2385 and n1175). These variables indicate fathers' socio-economic group. Information from n2385 ("Father, male heads socio-economic grp, GRO 1970") was used for a first round of coding and any observation with missing information on this variable were coded using variable n1175 ("Father, male head's socio-economic grp, GRO 1966"). The two variables have identical categories. A number of the categories in these socio-economic group variables do not clearly distinguish between employers and managers or between individuals with and without supervisory functions. The variables also do not give information on the number of employees in a given organization. Two alternative measures of employment status were generated. The first measure codes the seventeen categories of the socio-economic group variable into the five category employment status variable shown in Table 3. Note that categories that could not clearly be allocated to any employment status were coded as missing. Consequently, individuals in these categories are allocated to NS-SEC according to their occupational group only. The second measure of employment status is a dummy variable that differentiates self-employed from non-self-employed (see Table 4). Individuals for whom information on the source variables n2385 and n1175 was missing were coded as missing for both generated variables on employment status.

---

[2] http://doc.ukdataservice.ac.uk/doc/7023/mrdoc/pdf/ncds_bcs_occupation_coding.pdf

Four alternative measures of parental class have been generated. The first measure is based on the SOC2000 occupational codes. In a first step, all individuals are allocated to NS-SEC classes using the simplified method of allocating individuals to social classes based on their occupation only.[3] In a second step, individuals for whom information on employment status is available were reallocated based on the five-category employment status variable described above (following the reduced method of allocating individuals to NS-SEC classes). Table 5 shows the distribution of this first measure of parental class using the simplified method, the reduced method and the distribution of parental class found in Bukodi and Goldthorpe (2013, p. 1039). It is important to note that when comparing the distribution of parental class across the newly generated measure and that by Bukodi and Goldthorpe (2013), one sees a marked difference in the percentage of parents found in Class 2 (Lower managerial positions), Class 3 (Intermediate positions) and in Class 4 (Small employers). The second measure of parental class follows the same procedure of allocating individuals to social classes as the first measure, but uses SOC90 occupational group codes instead of SOC2000 occupational group codes. The distribution of this second measure of parental class is shown in Table 6. Similarly, to the NS-SEC distribution based on SOC2000, there is a marked difference across the newly generated measure and that by Bukodi and Goldthorpe (2013), with regards to Classes 2, 3, 4. Additionally, there is a large difference in the percentage of parents found in Class 6.

The third measure of parental class takes an alternative approach to allocating individuals to social classes. First, all individuals are allocated to NS-SEC classes *as if they were employees*, using their SOC2000 codes. In a second step, the dummy variable indicating self-employment is used to reallocate individuals who are coded as self-employed. In a third step, individuals for whom no information is available on whether they are self-employed or not are reallocated using the simplified method of allocating individuals to NS-SEC classes. The distribution of this third measure of parental class is shown in Table 7. The fourth measure of parental class follows the same procedure of allocating individuals to social classes as the third measure, but uses SOC90 occupational group codes instead of SOC2000 occupational group codes. The distribution of this fourth measure of parental class is shown in Table 8. As expected, the relative size of NS-SEC Class 4 tends to be smaller for the third and fourth measure of parental class, when compared to the size of NS-SEC Class 4 for the first and second measure of parental class. However, the distribution of these measures three and four still differ substantially from that shown in Bukodi and Goldthorpe (2013).

Table 9 shows the distributions of the four newly generated measures of parental class as well as the distribution in Bukodi and Goldthorpe (2013) and the distribution of the measure of parental NS-SEC supplied in the data set provided by Paul Gregg. It can be seen that the distribution of the measure of parental NS-SEC supplied in the data set provided by Paul Gregg is nearly identical to the version of the first newly generated measure of parental class when using the simplified method of allocating individuals to NS-SEC classes. This can also be seen in Table 10 which cross-tabulates these two measures.

The versions of the newly generated measures one and two using the reduced method of allocating individuals to NS-SEC classes are cross-tabulated in Table 11. Surprisingly, one finds considerable differences in the relative distribution of the parental class measures based on SOC2000 and SOC90. To illustrate some of the differences consider the extreme case of the 126 individuals whose parents are allocated into Class 1 based on SOC2000, but Class 6 based on SOC90. The SOC2000 occupation of all these individuals is "Production, works and maintenance managers" and the SOC90 is "Other plant and machine operatives, n.e.c.". Given that SOC2000 consists of four digit codes and thus provides more

---

[3] For further information on the simplified, reduced and full method of allocating individuals to NS-SEC classes based on their occupational information see Office for National Statistics (2005) *The National Statistics Socio-Economic Classification: User Manual.* Palgrave: Newport (available online).

detailed occupational groups than SOC90 (which consist of three digit codes), the NS-SEC measure generated based on the SOC2000 occupational codes is likely to be more reliable. As in the example given above, four-digit SOC2000 occupational codes also appear to provide more information on parents' employment status (within the occupational group code), compared to the three-digit SOC90 codes. The difference in the two measures of parental class may thus originate from the poor information on parents' employment status in the NCDS. This would be consistent with the pattern that the parental class measure based on SOC2000 tends to allocate individuals to higher classes than is the case for the measure based on SOC90 (see Table 11). A similar pattern emerges when the third and fourth measure of parental class are cross-tabulated (see Table 12).

Table 3. Coding of five-category employment status variable

Socio-economic group variable in NCDS data	Allocated employment status variable for generating NS-SEC	Allocation rationale
[-1] NA	[.] Missing	N/A
[1] Emp,mana,large	[.] Missing	Not clear whether employers or managers
[2] Emp,manag,small	[.] Missing	Not clear whether employers or managers
[3] Prof-self-emp	[2] Self-employed	Clearly identified as self-employed
[4] Prof. employees	[.] Missing	Not clear whether supervisors or not
[5] Intermed non-man	[.] Missing	Not clear whether supervisors or not
[6] Junior non-man	[5] Other employees	Clearly identified as non-supervisory
[7] Personal service	[.] Missing	Not clear whether supervisors or not
[8] Foremen-manual	[4] Supervisors	Clearly identified as supervisors
[9] Skilled manual	[.] Missing	Not clear whether supervisors or not
[10] Semi skld manual	[.] Missing	Not clear whether supervisors or not
[11] Unskilled manual	[5] Other employees	Unlikely not to be supervisors
[12] Work own account	[2] Self-employed	Clearly identified as own-account
[13] Farm emp,manag	[.] Missing	Not clear whether employers or not
[14] Farm-own account	[2] Self-employed	Clearly identified as own-account
[15] Agric worker	[5] Other employees	Unlikely not to be supervisors
[16] Armed forces	[.] Missing	Not clear whether supervisors or not
[17] Inadequate info	[.] Missing	N/A

Notes: Categories [1] Employers and [3] Managers were not assigned, due to ambiguous information in the source variables. Accordingly these will be assigned based on occupational code only.

Source variable 1: n2385 (ncds0123) 3P Father,male heads socio-economic grp (GRO 1970)

Source variable 2: n1175 (ncds0123) 2P Father,male head's socio-economic grp (GRO 1966) [used when source variable 1 is missing]

Coding of five category employment status variable: [.] Missing, [1] Employers, [2] Self-employed, [3] Managers, [4] Supervisors, [5] Other employees

Table 4. Coding of two-category employment status variable

Socio-economic group variable in NCDS data	Allocated employment status variable for generating NS-SEC
[-1] NA	[.] Missing
[1[Emp,mana,large	[0] Not self-employed
[2] Emp,manag,small	[0] Not self-employed
[3] Prof-self-emp	[0] Not self-employed
[4] Prof. employees	[0] Not self-employed
[5] Intermed non-man	[0] Not self-employed
[6] Junior non-man	[0] Not self-employed
[7] Personal service	[0] Not self-employed
[8] Foremen-manual	[0] Not self-employed
[9] Skilled manual	[0] Not self-employed
[10] Semi skld manual	[0] Not self-employed
[11] Unskilled manual	[0] Not self-employed
[12] Work own account	[1] Self-employed
[13] Farm emp,manag	[1] Self-employed
[14] Farm-own account	[1] Self-employed
[15] Agric worker	[0] Not self-employed
[16] Armed forces	[0] Not self-employed
[17] Inadequate info	[.] Missing

Notes: Source variable 1: n2385 (ncds0123) 3P Father,male heads socio-economic grp (GRO 1970)

Source variable 2: n1175 (ncds0123) 2P Father,male head's socio-economic grp (GRO 1966), used when source variable 1 is missing

Table 5. Distribution of First Measure of Parental Social Class (based on SOC2000 codes provided by Gregg et al., 2012, and simplified and reduced allocation approach)

	Simplified Method		Reduced Method		Bukodi & Goldthorpe (2013)
	%	**N**	**%**	**N**	**%**
1. Higher managerial	8.04	803	7.98	797	6.6
2. Lower managerial	12.02	1,200	12.44	1,242	18
3. Intermediate	9.54	953	9.04	903	14.9
4. Small employers	12.26	1,224	13	1,298	5.3
5. Lower supervisory	16.34	1,632	20.36	2,033	23.2
6. Semi routine	17.66	1,764	15.46	1,544	12.3
7. Routine	24.14	2,411	21.73	2,170	19.7
Missing	—	4,650	—	4,650	—
Total	100	14,637	100	14,637	100

Note: The simplified method allocates individuals to NS-SEC based on their occupational code only. The reduced method allocates individuals based on their occupational code and information on their employment status.

Table 6. Distribution of Second Measure of Parental Social Class (based on SOC90 codes provided by Gregg et al., 2012, and simplified and reduced allocation approach)

	Simplified Method		Reduced Method		Bukodi & Goldthorpe (2013)
	%	**N**	**%**	**N**	**%**
1. Higher managerial	4.35	434	4.37	436	6.6
2. Lower managerial	8.03	802	8.47	846	18
3. Intermediate	10.34	1033	9.59	958	14.9
4. Small employers	8.06	805	9.88	987	5.3
5. Lower supervisory	16.83	1,681	20.65	2,062	23.2
6. Semi routine	24.28	2,425	21.39	2,136	12.3
7. Routine	28.11	2,807	25.65	2,562	19.7
Missing	—	4,650	—	4,650	—
Total	100	14,637	100	14,637	100

Note: The simplified method allocates individuals to NS-SEC based on their occupational code only. The reduced method allocates individuals based on their occupational code and information on their employment status.

Table 7. Distribution of Third Measure of Parental Social Class (based on SOC2000 codes provided by Gregg et al., 2012, and alternative allocation method)

	Alternative Method		Bukodi & Goldthorpe (2013)
	%	**N**	**%**
1. Higher managerial	7.98	797	6.6
2. Lower managerial	13.31	1,329	18
3. Intermediate	9.51	950	14.9
4. Small employers	8.05	804	5.3
5. Lower supervisory	15.14	1,512	23.2
6. Semi routine	18.05	1,803	12.3
7. Routine	27.96	2,792	19.7
Missing	—	4,650	—
Total	100	14,637	100

Table 8. Distribution of Fourth Measure of Parental Social Class (based on SOC90 codes provided by Gregg et al., 2012, and alternative allocation method)

	Alternative Method		Bukodi & Goldthorpe (2013)
	%	**N**	**%**
1. Higher managerial	4.35	434	6.6
2. Lower managerial	8.21	820	18
3. Intermediate	10.15	1014	14.9
4. Small employers	6.75	674	5.3
5. Lower supervisory	14.61	1,459	23.2
6. Semi routine	23.81	2,378	12.3
7. Routine	32.12	3,208	19.7
Missing	—	4,650	—
Total	100	14,637	100

Table 9. Distributions of Alternative Measures of Parental NS-SEC (%)

	Bukodi & Goldthorpe (2013)	Paul Gregg Measure	1st Measure		2nd Measure		3rd Measure	4th Measure
			simplified	*reduced*	*simplified*	*reduced*	*alternative*	*alternative*
1. Higher managerial	6.6	8.02	8.04	7.98	4.35	4.37	7.98	4.35
2. Lower managerial	18	12.02	12.02	12.44	8.03	8.47	13.31	8.21
3. Intermediate	14.9	9.54	9.54	9.04	10.34	9.59	9.51	10.15
4. Small employers	5.3	12.26	12.26	13	8.06	9.88	8.05	6.75
5. Lower supervisory	23.2	16.34	16.34	20.36	16.83	20.65	15.14	14.61
6. Semi routine	12.3	17.67	17.66	15.46	24.28	21.39	18.05	23.81
7. Routine	19.7	24.15	24.14	21.73	28.11	25.65	27.96	32.12
Total	100	100	100	100	100	100	100	100

Note: The simplified method allocates individuals to NS-SEC based on their occupational code only. The reduced method allocates individuals based on their occupational code and information on their employment status.

Table 10. Cross-tabulation of the Simplified Version of First Measure of Parental NS-SEC with the Measure of Parental NS-SEC given in the Gregg et al. (2012) data set

		Paul Gregg Measure							Total	
		1. Higher managerial	2. Lower managerial	3. Interme-diate	4. Small employers	5. Lower supervisory	6. Semi routine	7. Routine		
New Measure (simplified, SOC2000)	1. Higher managerial	**801**	0	0	0	0	0	0	801	*Frequencies*
		100	0	0	0	0	0	0	100	*Row %*
	2. Lower managerial	0	**1200**	0	0	0	0	0	1200	*Frequencies*
		0	**100**	0	0	0	0	0	100	*Row %*
	3. Intermediate	0	0	**953**	0	0	0	0	953	*Frequencies*
		0	0	**100**	0	0	0	0	100	*Row %*
	4. Small employers	0	0	0	**1224**	0	0	0	1224	*Frequencies*
		0	0	0	**100**	0	0	0	100	*Row %*
	5. Lower supervisory	0	0	0	0	**1632**	0	0	1632	*Frequencies*
		0	0	0	0	**100**	0	0	100	*Row %*
	6. Semi routine	0	0	0	0	0	**1764**	0	1764	*Frequencies*
		0	0	0	0	0	**100**	0	100	*Row %*
	7. Routine	0	0	0	0	0	0	**2411**	2411	*Frequencies*
		0	0	0	0	0	0	**100**	100	*Row %*
	Total	801	1200	953	1224	1632	1764	2411	9985	*Frequencies*
		8	12	10	12	16	18	24	100	*Row %*

Table 11. Cross-tabulation of First and Second Measure of Parental NS-SEC (reduced method, occupational data from Gregg et al., 2012)

		NS-SEC based on SOC90								
		1. Higher managerial	2. Lower managerial	3. Interme-diate	4. Small employers	5. Lower supervisory	6. Semi routine	7. Routine	*Total*	
NS-SEC based on SOC2000	1. Higher managerial	**396**	114	132	14	15	126	0	797	*Frequencies*
		50	14	17	2	2	16	0	100	*Row %*
	2. Lower managerial	40	**654**	201	19	28	197	103	1242	*Frequencies*
		3	**53**	16	2	2	16	8	100	*Row %*
	3. Intermediate	0	8	**600**	39	99	139	18	903	*Frequencies*
		0	1	**66**	4	11	15	2	100	*Row %*
	4. Small employers	0	64	9	**867**	233	122	3	1298	*Frequencies*
		0	5	1	**67**	18	9	0	100	*Row %*
	5. Lower supervisory	0	6	16	17	**1633**	150	211	2033	*Frequencies*
		0	0	1	1	**80**	7	10	100	*Row %*
	6. Semi routine	0	0	0	20	1	**1320**	203	1544	*Frequencies*
		0	0	0	1	0	**85**	13	100	*Row %*
	7. Routine	0	0	0	11	53	82	**2024**	2170	*Frequencies*
		0	0	0	1	2	4	**93**	100	*Row %*
	Total	436	846	958	987	2062	2136	2562	9987	*Frequencies*
		4	8	10	10	21	21	26	100	*Row %*

Table 12. Cross-tabulation of Third and Fourth Measure of Parental NS-SEC (reduced method, occupational data from Gregg et al., 2012)

		NS-SEC based on SOC90								
		1. Higher managerial	2. Lower managerial	3. Interme-diate	4. Small employers	5. Lower supervisory	6. Semi routine	7. Routine	*Total*	
NS-SEC based on SOC2000	1. Higher managerial	**394**	115	146	0	8	134	0	797	*Frequencies*
		49	14	18	0	1	17	0	100	*Row %*
	2. Lower managerial	40	**674**	206	12	3	218	176	1329	*Frequencies*
		3	**51**	16	1	0	16	13	100	*Row %*
	3. Intermediate	0	8	**636**	14	108	166	18	950	*Frequencies*
		0	1	**67**	1	11	17	2	100	*Row %*
	4. Small employers	0	23	6	**626**	86	61	2	804	*Frequencies*
		0	3	1	**78**	11	8	0	100	*Row %*
	5. Lower supervisory	0	0	20	12	**1052**	172	256	1512	*Frequencies*
		0	0	1	1	**70**	11	17	100	*Row %*
	6. Semi routine	0	0	0	8	1	**1524**	270	1803	*Frequencies*
		0	0	0	0	0	**85**	15	100	*Row %*
	7. Routine	0	0	0	2	201	103	**2486**	2792	*Frequencies*
		0	0	0	0	7	4	**89**	100	*Row %*
	Total	434	820	1014	674	1459	2378	3208	9987	*Frequencies*
		4	8	10	7	15	24	32	100	*Row %*

4   Parental Status

For the construction of the parental status variables, the information on SOC90 and SOC2000 codes indicating fathers' occupation in 1969 (at age 11 of cohort member) is used. This measure was taken from the data file prepared in the project led by Paul Gregg.[4] Based on the SOC codes, two measures of parental status have been constructed. The first is based on the Chan–Goldthorpe (CG) status order (Chan and Goldthorpe, 2004), which is derived from the SOC90 codes. Since no information is available in the NCDS on mothers' occupation, social status scores could only be generated for fathers. Fathers' social status scores are then taken to indicate parents' social status.

Figure 1 illustrates the distribution of Parental CG status scores which is based on fathers' CG status scores. The measure has been normalized to a 0 to 1 scale using the minimum and maximum of -0.6 and +0.6. The mean normalised Parental CG score is 0.24 and the standard deviation is 0.28. Table 13 presents the quartiles and frequencies taken from the original interval-scale variables.

Figure 1. Boxplot Distribution of Parental CG Status Scores (normalized, theoretical max/min)

Parental Chan/Goldthorpe 2004, normalised (theoretical min/max)
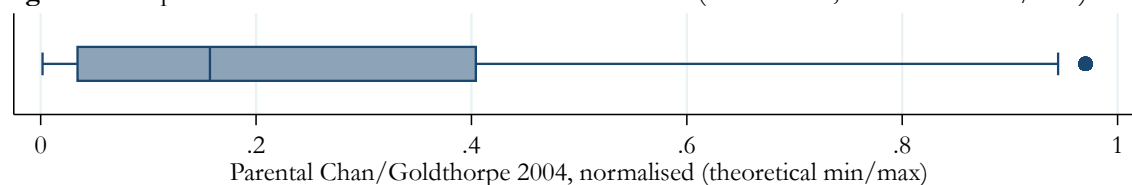
Table 13. Distribution of Parental CG Social Status Scores (Quartiles)

	%	N
1st Quartile	34.60	3,456
2nd Quartile	22.14	2,211
3rd Quartile	21.96	2,193
4th Quartile	21.30	2,127
Missing	—	4,650
Total	100	14,637

The second parental status variable is based on the CAMSIS scale. This provides two separate scales for men and women, assuming that social hierarchies vary according to gender. Fathers' CAMSIS status scores were derived from their SOC2000 codes using the conversion table. The measure has been normalized to a 0 to 1 scale using the minimum and maximum of 0 and 100 of the CAMSIS scale. Figure 2 illustrates the distribution of Parental CAMSIS scores which is based on fathers' CAMSIS scores. The mean Parental CAMSIS score is 0.42 and the standard deviation is 0.13. Table 14 presents the quartiles and frequencies taken from the original interval-scale variables.

Figure 2. Boxplot distribution of Parental CAMSIS Status Scores (normalized, theoretical max/min)

Parental CAMSIS 2004, normalised (theoretical min/max)

---
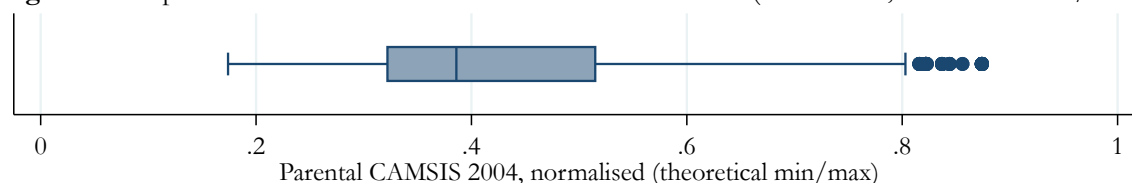
[4] http://doc.ukdataservice.ac.uk/doc/7023/mrdoc/pdf/ncds_bcs_occupation_coding.pdf

Table 14. Distribution of Parental CAMSIS Social Status Scores (Quartiles)

	%	N
1st Quartile	26.15	2,612
2nd Quartile	24.26	2,423
3rd Quartile	25.98	2,595
4th Quartile	23.60	2,357
Missing	—	4,650
Total	100	14,637

Table 15 presents Pearson's correlation between the CG and CAMSIS scale variables described above.

Table 15. Pearson's Correlations: Parental Status Scores (CG and CAMSIS)

	Parents' Status (CG)	Parents' Status (CAM)
Parents' Status (CG)	1	
Parents' Status (CAM)	0.82***	1

*** $p < 0.001$

5 Parental Education

The only information on parental education available in the NCDS indicates the age at which fathers and mothers left full time education (variables n2396 and n2397). This information was used to generate a four category variable (age 12-15 = no qualification, age 15-17 = lower secondary, age 17-19 upper secondary or lower tertiary, age 19+ = degree). The distribution of this four category measure of fathers' and mothers' education is given in Table 16 below. The distribution of the parental education variable using the dominance approach is given in Figure 3 below.

Table 16. Fathers' and mothers' Highest Educational Qualifications

	Fathers		Mothers	
	%	N	%	N
1. No qualifications	58.90	5,977	48.13	5,019
2. Lower secondary	29.15	2,958	41.61	4,339
3. Higher secondary or lower tertiary	7.00	710	6.85	714
4. Degree	4.96	503	3.42	357
Missing	—	4,489	—	4,208
Total	100.00	14,637	100.00	14,637

Figure 3. Distribution of Parental Education (Dominance Approach)
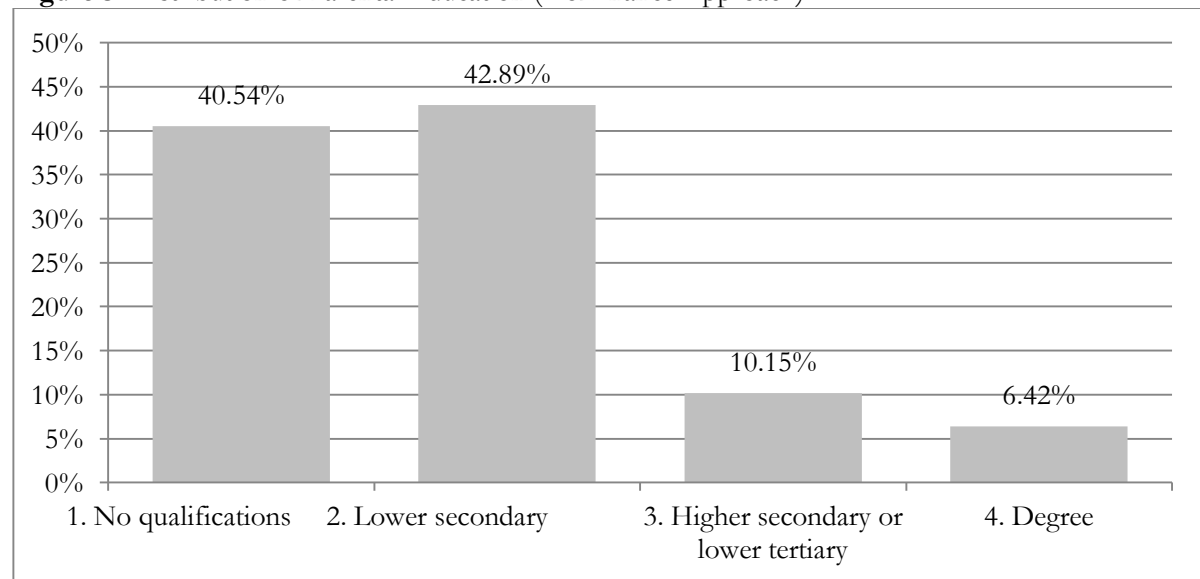
In addition, a combined parental education variable was created, reflecting the approach taken by Bukodi and Goldthorpe (2013). Information from both parents is incorporated into this measure and reflected in the following categories:

1. Neither parent has any qualification
2. One parent has secondary or lower qualification; other parent has no qualification
3. Both parents have secondary or lower qualification
4. One parent has higher secondary or lower tertiary qualification; other parent has lower qualification
5. Both parents have higher secondary or lower tertiary qualification
6. One parent has degree-level qualification; other parent has lower qualification
7. Both parents have degree-level qualifications

This method requires a plausible assumption to be made regarding missing values if educational information is only available for one parent. For the generation of the composite parental education variable it was assumed that the value of the missing parent equals the modal value for partners corresponding to the valid category. Table 17 shows the distributions for the generated composite parental education measure in comparison with the measure generated by Bukodi and Goldthorpe (2013) with the NCDS data. As can be seen, there are some differences particularly with regards to categories 2 and 3.

Table 17. Parental Education: Composite measure

	Generated variable	Bukodi & Goldthorpe 2013
1. Neither parent has any qualification	40.87%	40.90%
2. One parent has secondary or lower qualification; other parent has no qualification	21.17%	24.80%
3. Both parents have secondary or lower qualification	21.58%	18.00%
4. One parent has higher secondary or lower tertiary qualification; other parent has lower qualification	8.09%	8.10%
5. Both parents have higher secondary or lower tertiary qualification	1.97%	2.00%
6. One parent has degree-level qualification; other parent has lower qualification	4.29%	4.40%
7. Both parents have degree-level qualifications	2.02%	1.90%

To allow for meaningful cross-cohort comparison of educational attainment, it is useful to treat education in relative terms. To this end scores can be assigned 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 18 presents the distribution of this relative scale, split into two variables: the first has seven levels and the second has four levels. For the remainder of this data note, the composite measure of parental education is used when parents' educational attainment is referred to.

Table 18. Parental Education: Scores of relative scale (% falling below given educational level)

7 levels	4 levels	Generated variable	Bukodi & Goldthorpe 2013
1	1	0.0	0.0
2	2	40.9	41.5
3		62.1	64.6
4	3	83.6	84.5
5		91.7	92.4
6	4	93.7	94.3
7		98.0	98.4
Mean level (7 levels)		0.375	0.363
Standard deviation (7 levels)		0.346	0.340

The mean educational level is calculated by normalising the relative scale variable so that values are within the range 0–1. It should be compared to those found in Bukodi and Goldthorpe (2013) using the earlier cohorts and to the values found for the respondents of the LSYPE survey (born in 1989/1990), listed below.

1946 Cohort: Mean 0.306, Standard deviation 0.371
1970 Cohort: Mean 0.389, Standard deviation 0.326
1989/1990 Cohort (LSYPE): Mean 0.400, Standard deviation 0.332

6   Cognitive Ability[5]

At age 11, respondents took part in a general cognitive ability test, which yielded scores for verbal and non-verbal ability. Scores on each item for each test were summed, and the NCDS provides the summed scores in two interval-scale variables. 3,223 cases were missing information and 1,204 were coded 'NA' on both measures. Those coded 'NA' were set two missing and a Principal Components Analysis (PCA) was subsequently run on the verbal and non-verbal 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'. The first principal component score accounted for 90% of the total variance across the two measures. Table 19 presents the factor loadings. 4,227 of the original 18,558 cases are missing information on this new measure.

Table 19. PCA to Derive Cognitive Ability Measure

PCA			New Variable			
Eigenvalue	Loading	% of variance	Min	Max	Mean (SD)	N
	Verbal Non-verbal					
1.81	0.71 0.71	90%	-3.61	3.13	0.00 (1.34)	14,131

Figure 4 presents the percentile distribution of the cognitive ability variable. Table 20 presents quintiles.

Figure 4. Distribution of Cognitive Ability
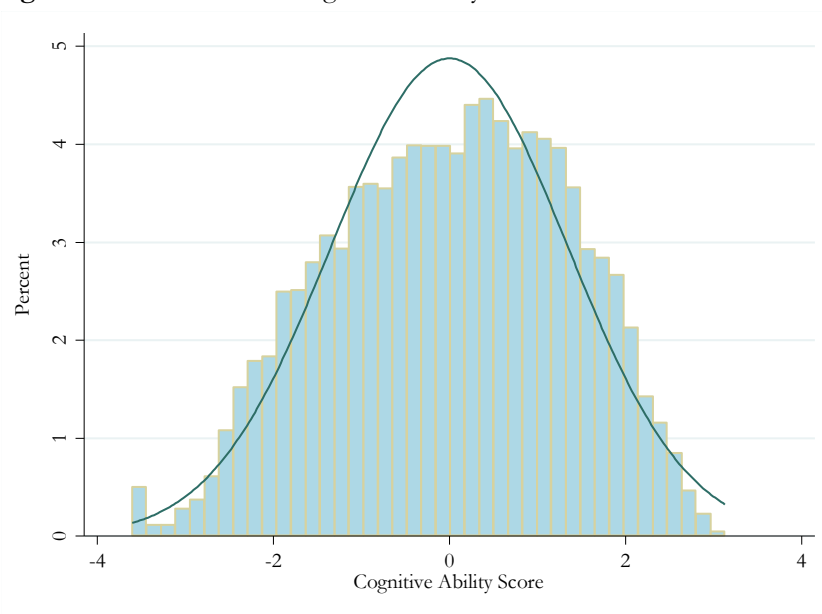
Table 20. Quintiles of Cognitive Ability

	%	**N**
1st Quintile (Lowest)	20	2828
2nd Quintile	20	2834
3rd Quintile	20	2831
4th Quintile	20	2815
5th Quintile (Highest)	20	2823
Missing	--	4427
Total	100	18558

7   Respondent Educational Attainment

This section provides an overview of the key aspects of respondent's educational attainment that will be used in the analysis for the research project. A key challenge in generating the educational attainment 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 educational 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 others as not having gained any educational qualification. In order to allow comparison with the other data notes compiled for the project, Table 21 and Table 22 below show the distribution of respondent educational attainment attained by age 20 and age 38, respectively.

Table 21. Highest educational qualification at age 20

	%	**N**
[1] No qualifications	29.51	4,312
[2] Below O-level, NVQ 1 [Sub-secondary]	10.78	1,575
[3] 1–4 O-level passes, NVQ 2 [Secondary—low performance]	21.46	3,135
[4] 5+ O-level passes or 1 A-level pass, NVQ 3 [Secondary—high perf.]	18.57	2,713
[5] 2+ A-level passes [Higher secondary]	15.17	2216
[6] Tertiary sub-degree qualification, NVQ 4 [Lower tertiary]	4.51	659
[7] Degree, NVQ5 or 6, higher degree [Higher tertiary]	0.01	1
Missing	—	26
Total	100.00	14,637

Table 22. Highest educational qualification at age 38

	%	N
[1] No qualifications	18.40	2,181
[2] Below O-level, NVQ 1 [Sub-secondary]	11.59	1,374
[3] 1–4 O-level passes, NVQ 2 [Secondary—low performance]	21.87	2,592
[4] 5+ O-level passes or 1 A-level pass, NVQ 3 [Secondary—high perf.]	16.65	1,974
[5] 2+ A-level passes [Higher secondary]	4.04	479
[6] Tertiary sub-degree qualification, NVQ 4 [Lower tertiary]	14.22	1,686
[7] Degree, NVQ5 or 6, higher degree [Higher tertiary]	13.22	1,567
Missing	—	2,784
Total	100.00	14,637

In addition to the measure of educational attainment shown in Tables 21 and 22 a series of measures were generated indicating whether cohort members have crossed a Key Stage 5 threshold. Table 23 and Table 24 show the proportion of individuals who have crossed the threshold by age 20 and 38 respectively. Here the threshold is operationalized as having attained two or more A-levels or any higher academic qualification by the respective age. Tables 25 and 26 show the proportion of individuals who have crossed the threshold by age 20 and 38 respectively, using an operationalisation that also takes into account vocational qualifications. Table 27 cross-tabulates the level of educational qualification attained at age 20 and the measure for crossing the KS5 threshold by age 20.

Table 23. Passing of Threshold at Key Stage 5 by age 20 (academic only)

	%	N
No	83.76	12,238
Yes	16.24	2,373
Missing		26
Total	100.00	14,637

Table 24. Passing of Threshold at Key Stage 5 by age 38 (academic only)

	%	N
No	77.64	9,203
Yes	22.36	2,650
Missing		2,784
Total	100.00	14,637

Table 25. Passing of Threshold at Key Stage 5 by age 20 (academic & vocational)

	%	N
No	80.32	11,735
Yes	19.68	2,876
Missing		26
Total	100.00	14,637

Table 26. Passing of Threshold at Key Stage 5 by age 38 (academic & vocational)

	%	N
No	68.51	8,121
Yes	31.49	3,732
Missing		2,784
Total	100.00	14,637

Table 27. Crosstab of highest educational qualification at age 20 and passing of KS5 Threshold (academic and vocational)

	Transition to KS 5			
	No	Yes	Total	
1. No qualifications	4312	0	4312	N
	37	0	30	% (column)
2. Below O-level, NVQ 1 [Sub-secondary]	1575	0	1575	N
	13	0	11	% (column)
3. 1–4 O-level passes, NVQ2 [Secondary—low performance]	3135	0	3135	N
	27	0	21	% (column)
4. 5+ O-level passes or 1 A-level pass, NVQ 3 [Secondary—high perf.]	2713	0	2713	N
	23	0	19	% (column)
5. 2+ A-level passes [Higher secondary]	0	2216	2216	N
	0	77	15	% (row)
6. Tertiary sub-degree qualification, NVQ 4 [Lower tertiary]	0	659	659	N
	0	23	5	% (column)
7. Degree, NVQ 5 or 6, higher degree [Higher tertiary]	0	1	1	N
	0	0	0	% (column)
Total	11735	2876	14611	N
	100	100	100	% (column)

An additional measure was generated indicating whether cohort members have crossed the Higher Education threshold. Table 28 shows the proportion of individuals who have crossed the threshold by age 38. Here the threshold is operationalized so that individuals with sub-degrees are not counted as having made the transition. Table 29 shows the distribution when the measure includes sub-degrees. Tables 30 and 31 show the proportion of individuals who have crossed the threshold, again including and excluding individuals with sub-degrees respectively, but also counting the vocational qualifications at the (sub-)degree level. Table 32 cross-tabulates the level of educational qualification attained at age 38 and the measure for crossing the Higher Education threshold by age 38 (excluding sub-degree).

Table 28. Passing of Higher Education Threshold by age 38 (academic only), excluding sub-degree

	%	N
No	87.37	10,356
Yes	12.63	1,497
Missing		2,784
Total	100.00	14,637

Table 29. Passing of Higher Education Threshold by age 38 (academic only), including sub-degree

	%	N
No	83.84	9,938
Yes	16.16	1,915
Missing		2,784
Total	100.00	14,637

Table 30. Passing of Higher Education Threshold by age 38 (academic & vocational), excluding sub-degree

	%	N
No	86.78	10,286
Yes	13.22	1,567
Missing		2,784
Total	100.00	14,637

Table 31. Passing of Higher Education Threshold by age 38 (academic & vocational), including sub-degree

	%	N
No	72.56	8,600
Yes	27.44	3,253
Missing		2,784
Total	100.00	14,637

Table 32. Crosstab of highest educational qualification at age 38 and transition to HE (academic & vocational), excluding sub-degree

	Transition to HE			
	No	Yes	Total	
1. No qualifications	2181	0	2181	N
	21	0	18	*% (column)*
2. Below O-level, NVQ 1 [Sub-secondary]	1374	0	1374	N
	13	0	12	*% (column)*
3. 1–4 O-level passes, NVQ2 [Secondary—low performance]	2592	0	2592	N
	25	0	22	*% (column)*
4. 5+ O-level passes or 1 A-level pass, NVQ 3 [Secondary—high perf.]	1974	0	1974	N
	19	0	17	*% (column)*
5. 2+ A-level passes [Higher secondary]	479	0	479	N
	5	0	4	*% (row)*
6. Tertiary sub-degree qualification, NVQ 4 [Lower tertiary]	1686	0	1686	N
	16	0	14	*% (column)*
7. Degree, NVQ 5 or 6, higher degree [Higher tertiary]	0	1567	1567	N
	0	100	13	*% (column)*
Total	10286	1567	11853	N
	100	100	100	*% (column)*

The NCDS data provide pre-coded variables indicating the number of GCSE/CSE/O-Level qualifications grade A*-C by 1974 (at age 16 of cohort members) and by 1978 (at age 20 of cohort

members). Variables indicating the number of A-Level / SCE 'Higher' passes is provided for qualifications completed by 1976 (at age 18 of cohort members) and by 1978 (at age 20 of cohort members). The cut-offs at age 16 and 18 capture information on what might be considered the 'conventional' time of examination for the respective qualification and does not include any retakes in November or December of those years[6]. Measures for the different age cut-offs are reported in Tables 33 – 36 below. Figures 5 shows the distribution of the number of GCSE/CSE/O-Level qualifications passed grade A*-C by age 16 and Figure 6 shows the distribution of the A-Levels passed by age 18.

Table 33. Performance at lower secondary level — Number of GCSE/CSE/O-Level qualifications at grades A* – C (by age 16)

	%	N
0	51.42	6,501
1	11.04	1,396
2	6.76	855
3	5.29	669
4	5.18	655
5	4.37	552
6	4.30	544
7	4.48	567
8	3.73	472
9+	3.42	433
Missing	—	1,993
Total	100.00	14,637

Table 34. Performance at lower secondary level — Number of GCSE/CSE/O-Level qualifications at grades A* – C (by age 20)

	%	N
0	47.56	6,014
1	11.03	1,395
2	6.76	855
3	5.25	664
4	4.63	586
5	4.59	580
6	4.37	553
7	4.63	585
8	4.17	527
9+	7.00	885
Missing	—	1,993
Total	100.00	14,637

---

[6] See NCDS document 'public_examination_results_1978'.

Table 35. Performance at upper secondary level — Number of A levels passed (by age 18)

	%	N
0	84.45	10,678
1	3.26	412
2	4.06	513
3	6.07	767
4	1.09	138
5	0.58	73
6	0.36	45
7	0.09	12
8	0.04	5
9	0.01	1
Missing	—	1,993
Total	100.00	14,637

Table 36. Performance at upper secondary level (for respondents who made transition to KS 5) — Number of A levels passed (by age 20)

	%	N
0	82.96	10,489
1	3.52	445
2	4.62	584
3	6.59	833
4	1.13	143
5	0.64	81
6	0.38	48
7	0.10	13
8	0.05	6
9	0.02	2
Missing	—	1,993
Total	100.00	14,637

Figure 5. Performance lower sec. level: Number of GCSE/CSE/O-Level qualifications passed at grades A* – C by age 16
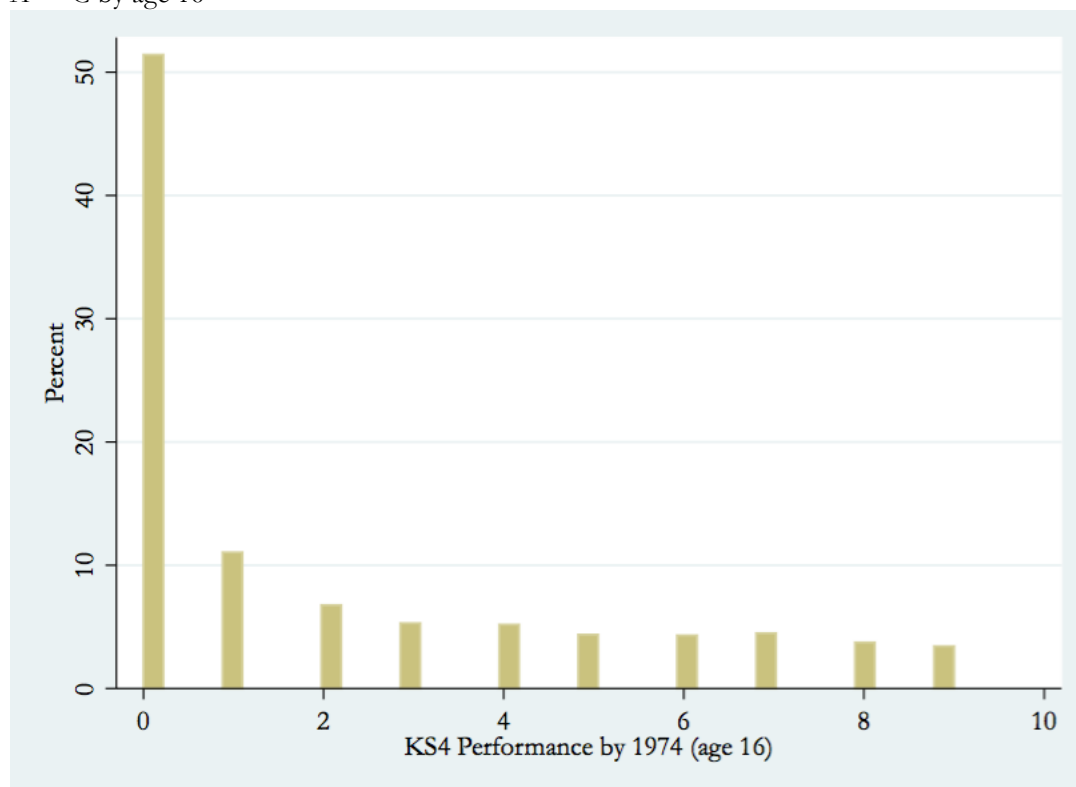
Figure 6. Performance upper sec. level: Number of A-Levels passed by age 18
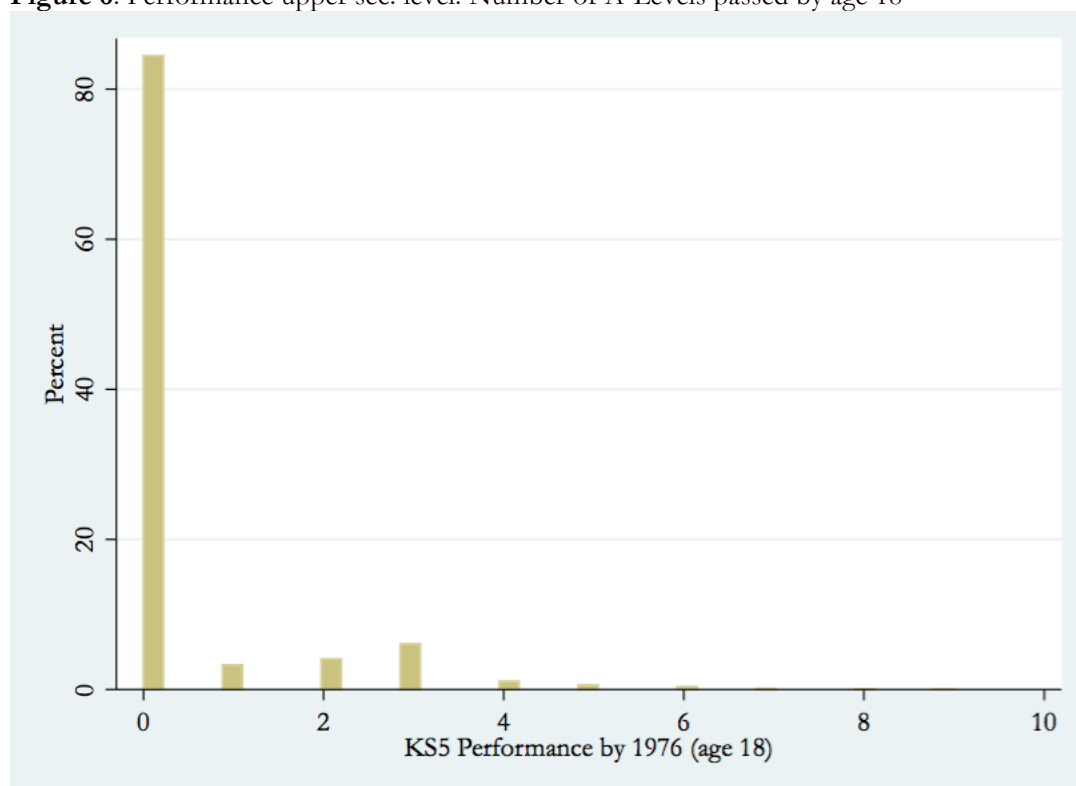
8   Associations between Social Origin Measures

This section presents the results from a series of bivariate analyses between the social origin measures. First, the parental CG status score variable (original scale) is considered in terms of its distribution across parental social class. Table 37 reports the mean Chan-Goldthorpe social status scored for each of the seven NS-SEC classes. Figure 7 shows the distribution of the CG social status measure for each of the seven NS-SEC classes. Interestingly, there appears to be a jump in the level of CG social status between NS-SEC 4 (Small employers and own account workers) and NS-SEC 3 (Intermediate occupations).

Table 37. Mean Parental CG Status Scores by Parental Class

NS-SEC (reduced method)	Mean	Std. Dev.	Freq.
1. Higher managerial	0.19	0.38	797
2. Lower managerial	0.11	0.33	1242
3. Intermediate	-0.09	0.25	903
4. Small employers	-0.41	0.21	1298
5. Lower supervisory	-0.49	0.10	2033
6. Semi routine	-0.46	0.16	1544
7. Routine	-0.51	0.11	2170
Total	-0.31	0.33	9987

Figure 7. Boxplot Distribution of Parental CG Status Scores by Parental Class
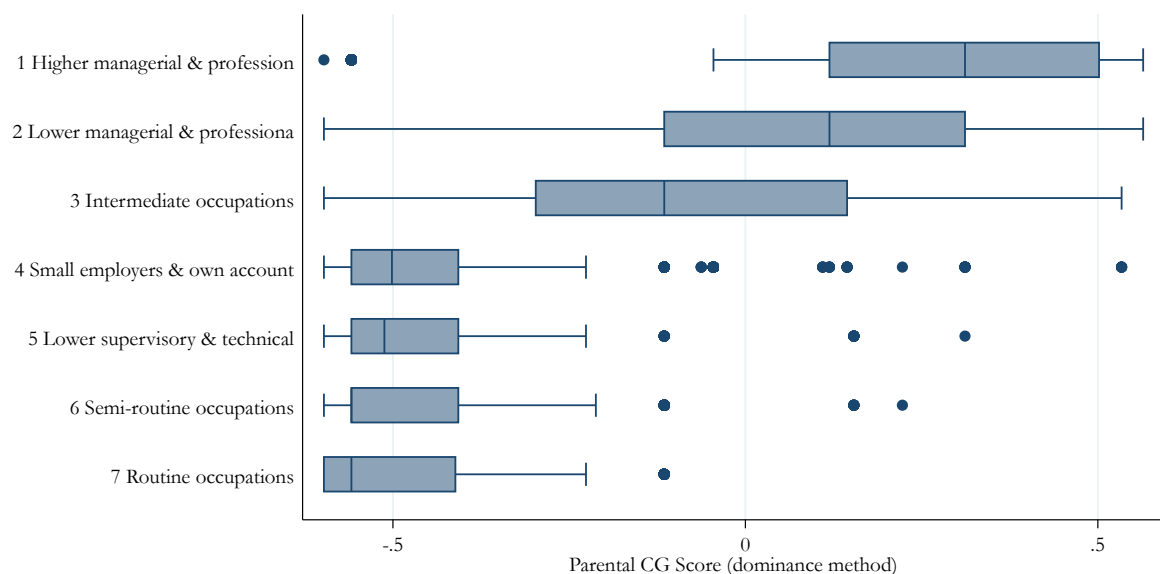
Table 38 presents the results of a one-way analysis of variance, showing that the differences between groups are statistically significant (F=2475.14, p<0.01) and that 59% of the total variance in parental status is explained by parental class. For the purpose of comparison, note that in the LSYPE data 52% of the total variance in parental status is explained by parental class (see LSYPE Data Note).

Table 38. One-way ANOVA: Parental Status by Parental Class

Between groups	Within groups	F	p	%
666.21	447.7	2475.14	0.00	59

Note: % denotes the percentage of total variance in status explained by social class

Table 39 uses the 7-level relative scale parental education variable (normalised with range 0–1) to present the mean level of parental education for each social-class group.

Table 39. Mean Parental Education (7-level Relative Scale, Normalised) by Parental Class

NS-SEC (reduced method)	Mean	Std. Dev.	Freq.
1. Higher managerial	0.68	0.33	586
2. Lower managerial	0.58	0.35	944
3. Intermediate	0.46	0.34	690
4. Small employers	0.37	0.33	963
5. Lower supervisory	0.30	0.31	1601
6. Semi routine	0.29	0.31	1155
7. Routine	0.25	0.29	1614
Total	0.38	0.34	7553

Table 40 gives the results of a one-way analysis of variance. Differences in the mean educational level between social-class groups are statistically significant (F= 237.47, p<0.01) and 16% of the total variance in parental education is explained by parental class. For the purpose of comparison, note that in the BCS70 and LSYPE data about 33% of the total variance in parental education is explained by parental class (see LSYPE Data Note). This difference may be related to the less specific information on parental education given in the NCDS data (only information on the age that fathers and mothers left education is given, see section on Parental Education above for further details). It may also be due to the fact that in the NCDS parental class was coded using information from fathers only (information on mothers' occupation is not available in the NCDS), while parental education was coded using information from both fathers and mothers.

Table 40. One-way ANOVA: Parental Education by Parental Class

Between groups	Within groups	F	p	%
142.61	755.27	237.47	0.000	16

Note: % denotes the percentage of total variance in status explained by social class

A cross-tabulation between parental class and education is given in Table 56 in the Appendix, which uses the composite parental education variable (i.e., not relativized/not normalized). Figure 8 below provides a cruder summary of the relationship by comparing the proportion of cases for which: (i) neither parent has any qualifications, and (ii) at least one parent has a degree-level qualification, by social-class. It is important to note that the operationalization of category (ii) differs from the definition used in the LSYPE Data Note. In the latter it is defined as 'both parents have a degree-level qualifications'. This definition was note appropriate for the purpose of showing the distribution in NCDS, since this extreme category (category 7) had too few observations (see Table in the Appendix). Thus category 6 and 7 of the composite measure of parental education were combined.

Figure 8. Proportion of Cases: (i) Neither Parent has Any Qualifications, and (ii) At least one parent has a Degree-Level Qualifications, by Parental Class
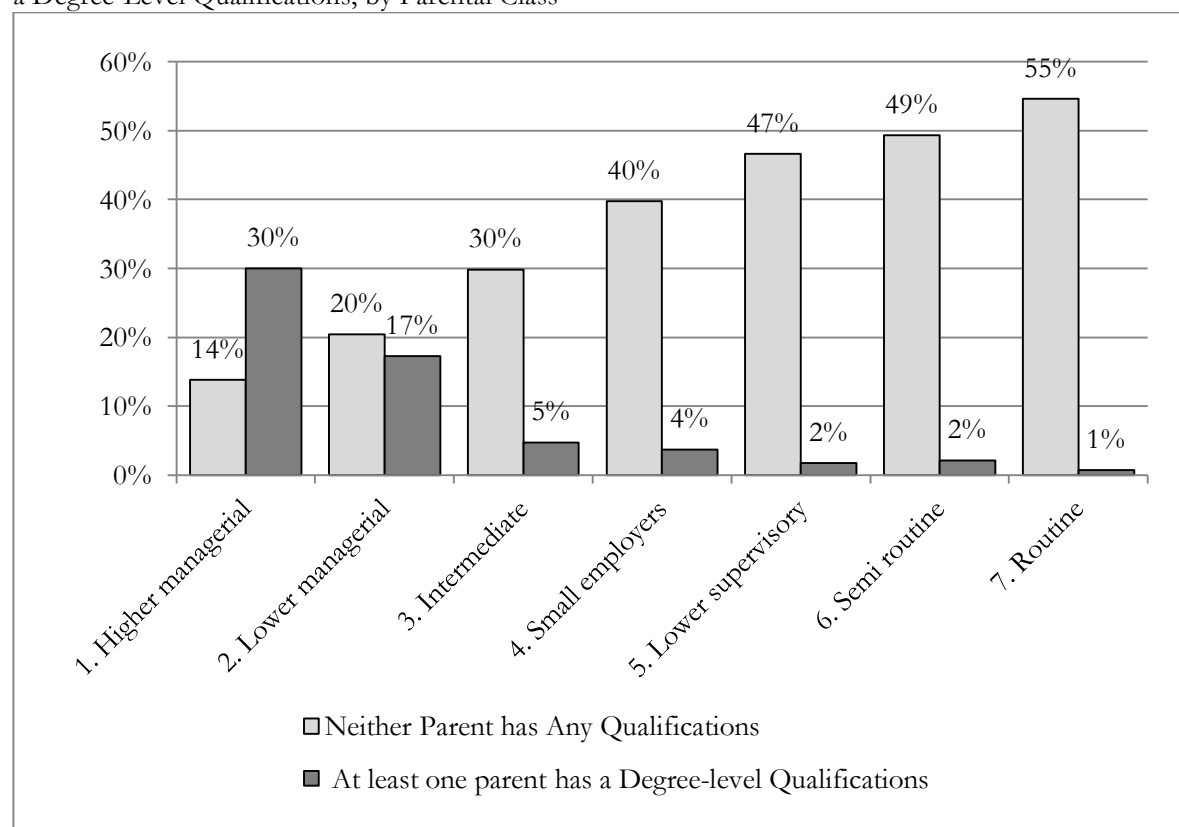
Figure 9 shows how the parental CG status variable is distributed within each of the seven groups of the composite parental education variable. As expected, those with higher status scores tend to have higher levels of education.

Figure 9. Boxplot Distribution of Parental CG Status Scores (Original Scale) by Parental Education (Composite measure)
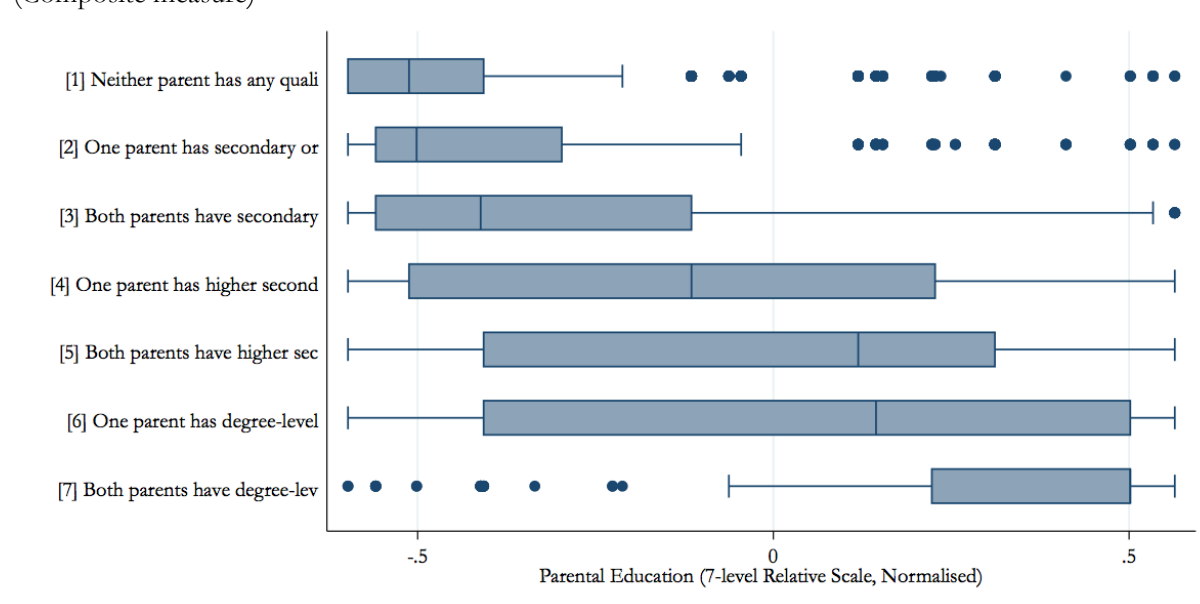
Table 41 uses the parental CG status quartiles variable to examine the relationship between parental status and parental education.

Table 41. Cross-tabulation of CG Parental Status Quartiles and Parental Education (Composite measure)

		1. No qualifications	2	3	4	5	6	7. Both have degrees	Total	
		Parental Education								
Parental Status	1st Quartile	1,316	614	496	138	21	47	9	2,641	N
		49.83	23.25	18.78	5.23	0.8	1.78	0.34	100	Row %
		43.01	36.99	30.98	22.62	14.09	14.92	5.7	34.97	Column %
	2nd Quartile	764	391	367	75	14	29	11	1,651	N
		46.27	23.68	22.23	4.54	0.85	1.76	0.67	100	Row %
		24.97	23.55	22.92	12.3	9.4	9.21	6.96	21.86	Column %
	3rd Quartile	683	407	380	113	22	39	9	1,653	N
		41.32	24.62	22.99	6.84	1.33	2.36	0.54	100	Row %
		22.32	24.52	23.74	18.52	14.77	12.38	5.7	21.89	Column %
	4th Quartile	297	248	358	284	92	200	129	1,608	N
		18.47	15.42	22.26	17.66	5.72	12.44	8.02	100	Row %
		9.71	14.94	22.36	46.56	61.74	63.49	81.65	21.29	Column %
	Total	3,060	1,660	1,601	610	149	315	158	7,553	N
		40.51	21.98	21.2	8.08	1.97	4.17	2.09	100	Row %
		100	100	100	100	100	100	100	100	Column %

Finally, correlations between the social origin variables are presented below. Table 42 gives Spearman rank correlations between categorical measures. Table 43 gives Pearson correlations between alternative social origin measures, all of which have been normalised to take values in the range 0–1. All correlations are statistically significant (p<0.01). Parental class and parental status are strongly correlated. Correlations between parental class and parental education are moderate. Correlations between parental education and parental status are also moderate.

Table 42. Spearman Rank Correlations: Parental Class, CG Status Quartiles and Education (Composite Measure)

	Class	**Education**	**Status**
Class	1.00**		
Education	0.36**	1.00**	
Status	0.60**	0.30**	1.00**

Notes: Values have been reversed so that lower values indicate a lower status/class/education category

** p<0.01

Table 43. Pearson Correlations: Parental Class, CG Status (Original Scale) and Education (Composite Measure and Relative Scale)

	Class	**Education†**	**Education‡**	**Status**
Class	1.00**			
Education†	0.38**	1.00**		
Education‡	0.42**	0.94**	1.00**	
Status	0.71**	0.38**	0.45**	1.00**

Notes: Values have been reversed so that lower values indicate a lower status/class/education category

All variables have been normalised between 0–1

† Uses 7-level relative scale

‡ Uses Composite Measure

** p<0.01

9 Social Origin and Cognitive Ability

This section examines how cognitive ability, measured at age 10, is associated with each measure of social origin described above. Figure 10 and Table 44 consider how cognitive ability at age 10 is distributed across social class categories.

Figure 10. Boxplot Distribution of Cognitive Ability by Parental Class
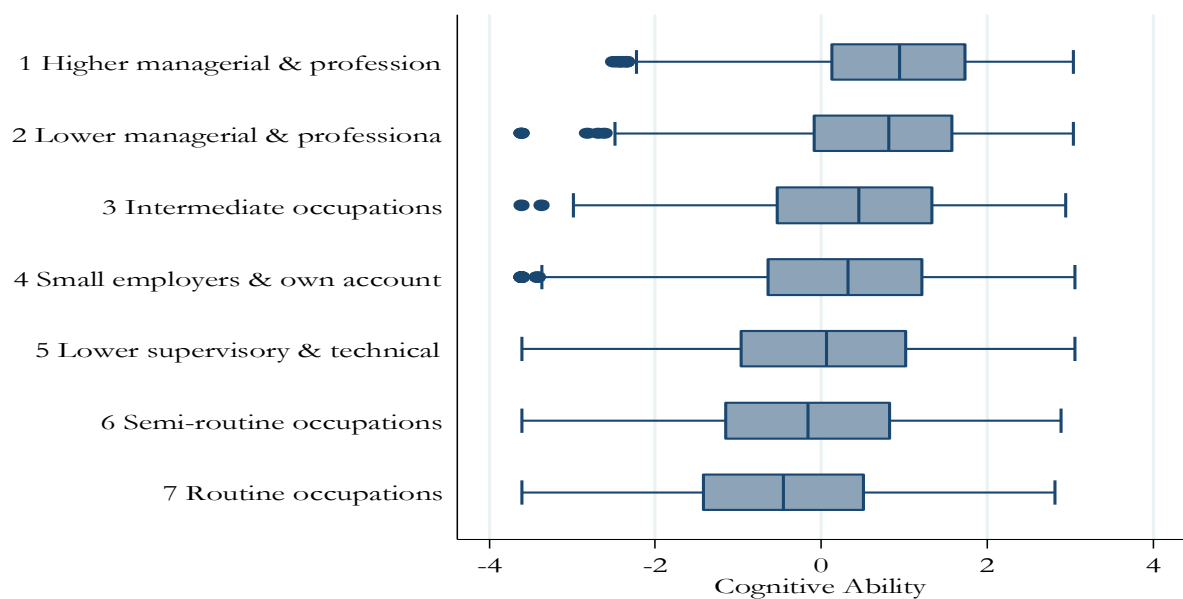
Table 44. Mean Cognitive Ability by Parental Class

NS-SEC (reduced method)	Mean	Std. Dev.	Freq.
1. Higher managerial	0.84	1.14	586
2. Lower managerial	0.67	1.19	966
3. Intermediate	0.37	1.26	719
4. Small employers	0.24	1.29	1961
5. Lower supervisory	-0.00	1.30	1796
6. Semi routine	-0.18	1.27	1357
7. Routine	-0.08	1.36	3131

Table 45 and Figure 11 report how cognitive ability at age 10 is distributed across the quintiles of the parental CG status measure. Table 46 provides a cross-tabulation of parental CG status score quintiles by cognitive ability quintiles. The two measures, in their original interval forms, are correlated at 0.26 ($p < 0.01$).

Table 45. Mean Cognitive Ability by Parental CG Status Quintiles

Chan–Goldthorpe Status	*Mean*	*Std. Dev.*	*Freq.*
1st Quintile (Lowest)	-0.32	1.33	1860
2nd Quintile	-0.05	1.30	1949
3rd Quintile	-0.01	1.29	2410
4th Quintile	0.16	1.29	1281
5th Quintile (Highest)	0.74	1.16	1831

Figure 11. Boxplot Distribution of Cognitive Ability by Parental CG Status Quintiles
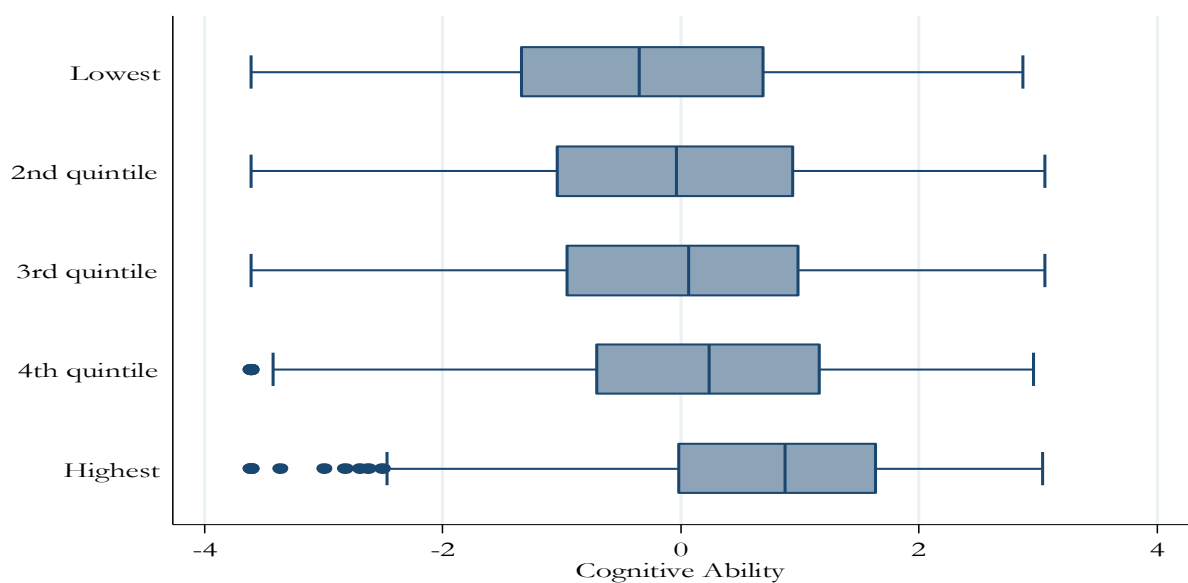
Table 46. Cross-Tabulation of Parental CG Status Quintiles by Cognitive Ability Quintiles

		Cognitive Ability					
		Lowest	2nd Quintile	3rd Quintile	4th Quintile	Highest	Total
Parental CG Status	Lowest	28	24	19	16	13	100
		29	24	19	16	13	20
	2nd Quintile	22	22	20	20	17	100
		24	23	20	21	17	21
	3rd Quintile	20	21	23	19	17	100
		28	27	28	25	22	26
	4th Quintile	17	18	22	21	21	100
		12	13	15	14	14	14
	Highest	6	13	19	25	36	100
		7	13	18	25	35	20
	Total	19	20	21	20	21	100
		100	100	100	100	100	100

Table 47 uses the composite parental education measure to examine the distribution of cognitive ability scores at age 10 across categories.

Table 47. Mean Cognitive Ability by Parental Education

	Mean	SD	Min	Max	N
1. Neither parent has any qualification	-0.23	1.28	-3.61	3.06	3779
2.	0.06	1.29	-3.61	2.96	1991
3.	0.14	1.30	-3.61	2.95	1973
4.	0.68	1.15	-2.92	3.04	742
5.	0.92	1.08	-2.72	2.96	176
6.	0.95	1.09	-2.43	3.06	388
7. Both parents have degrees	1.12	1.15	-2.40	2.95	175

Finally, 5x5 contingency tables were produced for all social origin-by-cognitive ability quintile pairs. For social class, the variable was recoded to five ordinal categories (classes 3–5 were collapsed); for parental CG status, quintiles were used; and for parental education, categories 2–3 and categories 6–7 were collapsed). These tables are not reproduced here, but each shows the general trend that those from more advantaged social backgrounds – however defined – tend to have higher scores of cognitive ability.

In an attempt to determine whether the association between cognitive ability and social origin differs according to how one choses to operationalise 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 these tests are given in Table 48 below. The differences in the overall associations, measured via global log odds, between cognitive ability and each of the social origin variables are statistically significant.

Table 48. Average Centre Four Global Log Odds Ratios and a Test of Significance in the Differences

	Average Centre Four		Difference	SE Difference	Lower CI	Upper CI
	i	ii				
i.Class–ii.Status	0.96	0.70	-0.26	0.06	-0.37	-0.15
i.Class–ii.Education	0.96	1.34	0.38	0.08	0.22	0.54
i.Status–ii.Education	0.70	1.34	0.64	0.08	0.48	0.79

10   Social Origin and Educational Attainment

This section examines the relationship between each of the social origin measures and the main outcome variable of respondents' educational attainment at age 38. Tables 49–52 present cross-tabulations between parental class, education, and status by respondents' educational attainment.

Table 49. Cross-tabulation of Parental Class by Cohort Member's Educational Attainment at age 38

| Parental NS-SEC | Cohort Member's Educational Attainment at Age 38 | | | | | | | Total | |
	1	2	3	4	5	6	7		
Higher managerial	49	26	74	117	55	137	233	691	N
	7	4	11	17	8	20	34	100	Row %
	4	3	4	8	16	11	21	8	Column %
Lower managerial	106	76	167	172	69	198	265	1053	N
	10	7	16	16	7	19	25	100	Row %
	8	8	9	12	19	16	24	13	Column %
Intermediate	81	52	141	149	49	144	147	763	N
	11	7	18	20	6	19	19	100	Row %
	6	6	8	10	14	12	13	9	Column %
Small employers	179	111	245	218	40	150	108	1051	N
	17	11	23	21	4	14	10	100	Row %
	13	12	14	15	11	12	10	13	Column %
Lower supervisory	265	198	412	321	61	250	168	1675	N
	16	12	25	19	4	15	10	100	Row %
	20	22	23	22	17	20	15	20	Column %
Semi routine	280	164	330	209	45	160	101	1289	N
	22	13	26	16	3	12	8	100	Row %
	21	18	18	14	13	13	9	16	Column %
Routine	399	279	428	278	35	182	105	1706	N
	23	16	25	16	2	11	6	100	Row %
	29	31	24	19	10	15	9	21	Column %
Total	1359	906	1797	1464	354	1221	1127	8228	N
	17	11	22	18	4	15	14	100	Row %
	100	100	100	100	100	100	100	100	Column %

Note: Cohort members educational attainment categories are as follows:

1. No qualifications

2. Below O-level, NVQ 1 [Sub-secondary]

3. 1–4 O-level passes, NVQ2 [Secondary—low performance]

4. 5+ O-level passes or 1 A-level pass, NVQ 3 [Secondary—high perf.]

5. 2+ A-level passes [Higher secondary]

6. Tertiary sub-degree qualification, NVQ 4 [Lower tertiary]

7. Degree, NVQ 5 or 6, higher degree [Higher tertiary]

Table 50. Cross-tabulation of Parental Education (Composite Measure) by Cohort Members'
Educational Attainment at age 38

| Parental Education | *Cohort Member's Educational Attainment at Age 38* | | | | | | | Total | |
	1	2	3	4	5	6	7		
	753	501	844	611	104	418	231	3462	N
1. No qualifications	22	14	24	18	3	12	7	100	Row %
	52	51	44	40	28	33	20	40	Column %
	275	213	448	329	71	294	206	1836	N
2	15	12	24	18	4	16	11	100	Row %
	19	22	23	22	19	23	18	21	Column %
	305	226	440	343	70	262	249	1895	N
3	16	12	23	18	4	14	13	100	Row %
	21	23	23	23	19	21	22	22	Column %
	69	30	126	143	45	149	154	716	N
4	10	4	18	20	6	21	22	100	Row %
	5	3	7	9	12	12	13	8	Column %
	8	4	19	21	21	39	62	174	N
5	5	2	11	12	12	22	36	100	Row %
	1	0	1	1	6	3	5	2	Column %
	22	10	35	57	36	74	148	382	N
6	6	3	9	15	9	19	39	100	Row %
	2	1	2	4	10	6	13	4	Column %
	6	4	11	15	18	29	104	187	N
7. Both have degrees	3	2	6	8	10	16	56	100	Row %
	0	0	1	1	5	2	9	2	Column %
	1438	988	1923	1519	365	1265	1154	8652	N
Total	17	11	22	18	4	15	13	100	Row %
	100	100	100	100	100	100	100	100	Column %

Note: Cohort members educational attainment categories are as follows:

1. No qualifications

2. Below O-level, NVQ 1 [Sub-secondary]

3. 1–4 O-level passes, NVQ2 [Secondary—low performance]

4. 5+ O-level passes or 1 A-level pass, NVQ 3 [Secondary—high perf.]

5. 2+ A-level passes [Higher secondary]

6. Tertiary sub-degree qualification, NVQ 4 [Lower tertiary]

7. Degree, NVQ 5 or 6, higher degree [Higher tertiary]

Table 51. Cross-tabulation of Parental CG Status Quartiles by Cohort Members' Educational Attainment at age 38

| Parental CG Status | *Cohort Member's Educational Attainment at Age 38* | | | | | | | Total | |
	1	2	3	4	5	6	7		
1st Quartile	594	374	667	503	94	338	231	2801	N
	21	13	24	18	3	12	8	100	Row %
	44	41	37	34	27	28	21	34	Column %
2nd Quartile	351	232	435	294	58	236	182	1788	N
	20	13	24	16	3	13	10	100	Row %
	26	26	24	20	16	19	16	22	Column %
3rd Quartile	263	198	436	355	65	298	202	1817	N
	14	11	24	20	4	16	11	100	Row %
	19	22	24	24	18	24	18	22	Column %
4th Quartile	151	102	259	312	137	349	512	1822	N
	8	6	14	17	8	19	28	100	Row %
	11	11	14	21	39	29	45	22	Column %
Total	1359	906	1797	1464	354	1221	1127	8228	N
	17	11	22	18	4	15	14	100	Row %
	100	100	100	100	100	100	100	100	Column %

Note: Cohort members educational attainment categories are as follows:

1. No qualifications

2. Below O-level, NVQ 1 [Sub-secondary]

3. 1–4 O-level passes, NVQ2 [Secondary—low performance]

4. 5+ O-level passes or 1 A-level pass, NVQ 3 [Secondary—high perf.]

5. 2+ A-level passes [Higher secondary]

6. Tertiary sub-degree qualification, NVQ 4 [Lower tertiary]

7. Degree, NVQ 5 or 6, higher degree [Higher tertiary]

Table 52. Cross-tabulation of Parental CAMSIS Status Quartiles by Cohort Members' Educational Attainment at age 38

| Parental CAMSIS | *Cohort Member's Educational Attainment at Age 38* | | | | | | | Total | |
	1	2	3	4	5	6	7		
	413	282	518	417	60	275	176	2141	N
1st Quartile	19	13	24	19	3	13	8	100	Row %
	30	31	29	28	17	23	16	26	Column %
	345	252	474	360	61	269	197	1958	N
2nd Quartile	18	13	24	18	3	14	10	100	Row %
	25	28	26	25	17	22	17	24	Column %
	427	244	501	330	87	302	230	2121	N
3rd Quartile	20	12	24	16	4	14	11	100	Row %
	31	27	28	23	25	25	20	26	Column %
	174	128	304	357	146	375	524	2008	N
4th Quartile	9	6	15	18	7	19	26	100	Row %
	13	14	17	24	41	31	47	24	Column %
	1359	906	1797	1464	354	1221	1127	8228	N
Total	17	11	22	18	4	15	14	100	Row %
	100	100	100	100	100	100	100	100	Column %

Note: Cohort members educational attainment categories are as follows:

1. No qualifications

2. Below O-level, NVQ 1 [Sub-secondary]

3. 1–4 O-level passes, NVQ2 [Secondary—low performance]

4. 5+ O-level passes or 1 A-level pass, NVQ 3 [Secondary—high perf.]

5. 2+ A-level passes [Higher secondary]

6. Tertiary sub-degree qualification, NVQ 4 [Lower tertiary]

7. Degree, NVQ 5 or 6, higher degree [Higher tertiary]

11   Cognitive Ability and Educational Attainment

This section examines the association between early-life cognitive ability, measured at age 10, and educational attainment later in life. Table 53 reports mean cognitive ability scores across categories of highest qualification achieved at the ages of 20 and 38.[7] Mean cognitive ability scores rise with increased educational attainment, with the exception of those whose highest qualification at age 38 is equivalent to a tertiary sub-degree qualification; for these individuals, early-life cognitive ability is, on average, less than it is for those who achieved 2+ A-level passes. The same was found for the BCS70 cohort.

Table 53. Mean Cognitive Ability by Educational Attainment at Ages 20 and 38

	Age 20			Age 38		
	Mean	Std. Dev.	N	Mean (SD)	Std. Dev.	N
1. No qualifications	-0.74	1.31	3,552	-0.84	1.29	1,817
2. Below O-level, NVQ 1	-0.57	1.03	1,351	-0.57	1.08	1,178
3. 1–4 O-level passes, NVQ 2	0.03	1.13	2,711	-0.05	1.13	2,226
4. 5+ O-level passes or 1 A-level pass, NVQ 3	0.54	1.06	2,371	0.37	1.12	1,734
5. 2+ A-level passes	1.27	0.88	1,891	1.04	1.00	421
6. Tertiary sub-degree qualification, NVQ 4	0.46	1.12	573	0.56	1.06	1,467
7. Degree, NVQ 5 or 6, higher degree	—	—	1	1.26	0.91	1,348
Total	0.05	1.33	12,450	0.13	1.31	10,191

Following Bukodi and Goldthorpe (2013), Table 54 takes an alternative approach using the same information; mean scores are examined according to whether individuals passed various qualification thresholds. The first threshold compares those who attained some level of qualification (categories 2–7 in Table 53 above) rather than none (category 1), the second threshold compares those who attained qualifications at NVQ 1 or higher (categories 3–7) with those whose attainment is lower (categories 1–2), and so on.

---

[7] Educational attainment up until age 20 is also considered because this is the latest age for which we have attainment information for the LSYPE sample.

Table 54. Mean Cognitive Ability by Educational Attainment Thresholds at Ages 20 and 38

	Age 20			Age 38		
	Mean	Std. Dev.	N	Mean	Std. Dev	N
Threshold i.						
1	-0.74482311	1.3093233	3552	-0.83930862	1.2869016	1817
2–7	0.36565542	1.2022939	8898	0.3359706	1.2130902	8374
Threshold ii.						
1–2	-0.69637414	1.2401056	4903	-0.73465217	1.2155597	2995
3–7	0.53296842	1.153963	7547	0.48480785	1.1684187	7196
Threshold iii.						
1–3	-0.43830672	1.2516546	7614	-0.4446732	1.226024	5221
4–7	0.81581009	1.0665878	4836	0.72636476	1.1037329	4970
Threshold iv.						
1–4	-0.20689029	1.2781261	9985	-0.24101801	1.2505504	6955
5–7	1.0847017	1.0014016	2465	0.916154	1.0488815	3236
Threshold v.						
1–5	0.02899409	1.3381353	11,876	-0.16780145	1.2728417	7376
6–7	0.45933187	1.1162454	574	0.89737035	1.0548799	2815
Threshold vi.						
1–6	—	—	—	-0.04684914	1.269956	8843
7	—	—	—	1.2631165	0.9134537	1348

Table 55 reports the means of cognitive ability by the type of threshold.

Table 55. Mean Cognitive Ability Scores by Type of Educational Threshold (Age 38)

	KS5 Transition					
	Academic only			Academic and vocational		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
No	-0.16	1.25	7,896	-0.24	1.25	6,955
Yes	1.13	0.97	2,295	0.92	1.05	3,236

	HE Transition					
	Academic only			Academic and vocational		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
No	-0.04	1.27	8,904	-0.05	1.27	8,843
Yes	1.28	0.90	1,287	1.26	0.91	1,348

Appendix

Table 56. Cross-tabulation between Parental Social Class and Parental Education (Composite measure)

	1. No qualifications	2	3	4	5	6	7. Both have degrees	Total	
Higher managerial	81	71	112	110	36	102	74	586	N
	13.82	12.12	19.11	18.77	6.14	17.41	12.63	100	Row %
	2.65	4.28	7	18.03	24.16	32.38	46.84	7.76	Column %
Lower managerial	193	163	231	145	49	101	62	944	N
	20.44	17.27	24.47	15.36	5.19	10.7	6.57	100	Row %
	6.31	9.82	14.43	23.77	32.89	32.06	39.24	12.5	Column %
Intermediate	206	141	178	106	26	24	9	690	N
	29.86	20.43	25.8	15.36	3.77	3.48	1.3	100	Row %
	6.73	8.49	11.12	17.38	17.45	7.62	5.7	9.14	Column %
Small employers	383	222	226	82	14	31	5	963	N
	39.77	23.05	23.47	8.52	1.45	3.22	0.52	100	Row %
	12.52	13.37	14.12	13.44	9.4	9.84	3.16	12.75	Column %
Lower supervisory	746	407	335	77	8	24	4	1,601	N
	46.6	25.42	20.92	4.81	0.5	1.5	0.25	100	Row %
	24.38	24.52	20.92	12.62	5.37	7.62	2.53	21.2	Column %
Semi routine	570	283	215	53	9	22	3	1,155	N
	49.35	24.5	18.61	4.59	0.78	1.9	0.26	100	Row %
	18.63	17.05	13.43	8.69	6.04	6.98	1.9	15.29	Column %
Routine	881	373	304	37	7	11	1	1,614	N
	54.58	23.11	18.84	2.29	0.43	0.68	0.06	100	Row %
	28.79	22.47	18.99	6.07	4.7	3.49	0.63	21.37	Column %
Total	3,060	1,660	1,601	610	149	315	158	7,553	N
	40.51	21.98	21.2	8.08	1.97	4.17	2.09	100	Row %
	100	100	100	100	100	100	100	100	Column %