MSc Quantitative Research Methods

Closing the Gap?

A longitudinal analysis of role of the experience of poverty in explaining the socio-economic gradient of cognitive ability and educational attainment in the early years in Scotland

This research examines the so-called 'socio-economic gap in attainment' - a current policy priority in Scotland. It uses evidence from the Scottish sample of the longitudinal Millennium Cohort Study to establish the existence of, and the factors associated with, the socio-economic gap and gradient in cognitive ability and attainment in the early years of childhood and primary school in Scotland, with a particular emphasis on the experience of poverty.

The literature was reviewed to identify potential risk factors associated with both experiencing poverty and child development outcomes. Concerns about trends in inter-generational income mobility and the extent to which early childhood experiences determine early cognitive ability and future educational and economic outcomes were also considered as a driver of much of this research.

This research has confirmed that a cognitive or attainment gap exists beyond the age of 5 into the early years of primary school on several measures of socio-economic classification. There is evidence that the gradient and therefore the gap is getting larger – in particular between those who have experienced no poverty and those who have experienced persistent poverty.

Whilst the change in this gradient from age 3 to 5 associated with experience of poverty is to a large extent mediated through prior ability levels and the general home environment, the association between poverty and the gradient from age 5 to 7 is less well explained – with experience of poverty continuing to be negatively associated with educational progress in the early years of primary school over and above any of the characteristics controlled for in the regression analysis.

1 Introduction and background

1.1 The policy context and the need for clarity

'Closing the attainment gap' is an issue of considerable political interest north of the border. However, this gap is one that lacks a unifying definition. One recent influential report (Ellis & Sosu, 2014) used at least eight definitions for the gap, including:

> *a persistent gap in attainment between pupils from the richest and poorest households in Scotland (p.3)*
>
> *the achievement gap associated with poverty. (p.3 and p.6)*
>
> *poverty-linked educational inequality (p.6)*

The report also suggested closing this gap could mean:

> *(to) raise attainment or to address low achievement associated with poverty (p.3)*
>
> *raising achievement in/among economically disadvantaged groups (p.4 and p.6)*
>
> *improving the educational outcomes for pupils from economically disadvantaged homes (p.6)*
>
> *raising attainment to close the poverty gap (p.6)*
>
> *(knowing) what works for pupils from economically disadvantaged homes (p.6).*

It is clear that there is considerable ambiguity in how to define the gap, how to measure it and therefore how to frame an appropriate policy goal to addresses it (The Royal Society of Edinburgh, 2015).

The latter have appeared in various forms in recent Scottish Government policy documents. The National Improvement Framework (Scottish Government, 2016a), for example, specifies as one of its targets:

> *Achieving equity: ensuring every child has the same opportunity to succeed, with a particular focus on closing the poverty-related attainment gap. (p.3)*

It defines this gap as:

> *..the difference in attainment between the most and least disadvantaged children. (p.15)*

Indeed, an early draft of the National Improvement Framework was more specific, wanting to achieve a:

> *year on year improvement in closing the attainment gap (The Scottish Government, 2015, p. p. 10)*

This was defined as the percentage point gap in attainment at various points in a child's school education between those in the most deprived areas compared to all other children defined using the Scottish Index of Multiple Deprivation[1]. The final draft of the Framework was less prescriptive in regards to measurement, however, simply stating that:

> *We all need more robust and consistent evidence which will help us in raising attainment and closing the gap. We need to know the size of the attainment gap at different ages and stages, across Scotland…we need to know whether the attainment gap is narrowing over time...*

This would include:

> *..the breakdown of children's progress by deprivation over time (using the Scottish Index of Multiple Deprivation)" (Scottish Government, 2016a, pp. 16-17)*

[1] See section 3.4.2 for more information

More recently, the First Minister has talked about tackling the attainment gap *between schools,* (BBC, 2016), which seems to be shifting the focus to the role of schools in addressing any differences between pupils.

1.2 Towards a more research-based definition

Definitions matter, however, and there is general agreement that it is imperative to improve the understanding of the gap by improving the evidence base in Scotland (Ellis & Sosu, 2014; The Royal Society of Edinburgh, 2015).

Scottish Government research publications, rather than policy documents have been clearer about what is meant by a gap. They generally refer to it as:

> *"the unequal socio-economic patterning of outcomes and risk factors that disadvantage less affluent children" (Scottish Government, 2015, p. 1)*

This definition focuses on the descriptive aspects of an outcome gap in terms of measuring inequality and the sources of this inequality, the policy documents focusing more on addressing the inequality or closing the gap to achieve equity.

Focusing on cognitive and educational outcomes, it will therefore be necessary to:

- Establish to what extent these outcomes vary by socio-economic characteristics of pupils and their families (however defined)
- Determine how this relationship changes over time (either by observing and comparing different cohorts or examining the same cohort over time)
- Identify those (risk) factors which are most strongly associated with or play a causal role in driving these variations

1.3 Child development outcomes and socio-economic factors

The relationship between a child's cognitive and educational development and progress and family socio-economic factors has been clearly established in UK and elsewhere (Blow, Goodman, Kaplan, Walker, & Windmeijer, 2005; Brooks-Gunn & Duncan, 1997; Greg J Duncan & Brooks‐Gunn, 2000; Haveman & Wolfe, 1995; Mayer, 1997; Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2008).

There are many strands to this research attempting to identify the causal factors linking child development outcomes more widely to socio-economic status (SES). Those from sociology disciplines favour overlapping theories such as *the bioecological model* which focuses on proximal processes, person characteristics, environmental contexts and time periods (Bronfenbrenner & Morris, 2006). Proximal factors are those mechanisms through which distal factors such as socio-economic or demographic characteristics drive child development. By controlling for these proximal factors, the relationship between socio-economic background and the outcome considered should disappear.

Alternatives are the *parental stress models*: poverty is stressful and how parents cope with stressful events affect development trajectories (Mayer, 1997). (Haveman & Wolfe, 1995) discuss the importance of *role models* (parents and older siblings) - including the role of parental employment (Ruhm (2000) cited in Blow et al., 2005) and two-parent families (Seltzer, 1994) - and *neighbourhood effects* as transmission mechanisms (Leventhal & Brooks-Gunn, 2000). Finally, there are the *economic deprivation* or *welfare culture* theories – which focus on the existence of harmful effects of both the dependence and stigma associated with being on welfare or in long-term poverty on children's aspirations and attainments (Brooks-Gunn & Duncan, 1997; Greg J Duncan & Brooks‐Gunn, 2000; Macaulay, 1977).

An empirical strand of this literature has focused on the very early years and how differences in children's cognitive ability by family circumstances and environment emerge *before* they have even started primary school or pre-school (Bradley et al., 1989; Dearden, Sibieta, & Sylva, 2011; Klebanov, Brooks-Gunn, McCarton, & McCormick, 1998; Leventhal & Brooks-Gunn, 2000). Gaps related to childhood disadvantage emerge early, persist over time and have a lasting influence, predicting future educational outcomes (Blanden & Machin, 2010; Feinstein, 2003) and labour market performance in adulthood (Gregg & Machin, 1998).

Important interactions have also been found between development trajectories and socio-economic status, showing the unfair advantage of a higher status family background is compounded over time (Feinstein, 2004) and even influences the next generation's cognitive achievement (Gregg & Machin, 1998), however the dynamics of child cognitive ability and attainment have to be interpreted cautiously due to the issue of regression to the mean (Jerrim & Vignoles, 2013).

1.4 Implications for intergenerational mobility

Much of this research has stemmed from concerns about intergenerational mobility i.e. the extent to which future life chances and opportunities, specifically higher education, occupation and earnings, are related to social origins. This has attracted the attention of economists (Becker & Tomes, 1986) who characterise families as an optimising decision-making unit: parents make investments in their children's human capital (time, material resources, education, etc.) and (imperfectly) pass on cultural and genetic endowments and assets in order to improve their children's life chances.

With perfect capital markets (and information), families would invest optimally in their children regardless of socio-economic background and the only source of intergenerational inequality would be through the transfer of parental endowments and assets or tastes. However, credit market failures are likely to exist in education, particularly for investments in young children (Barr, 2004). This leaves poorer families credit constrained whilst rich families are able to invest more in education resulting in better results, jobs, incomes, health and happiness for their children (McMahon, 2010).

Refinements to this theory emphasize the importance of non-cognitive skills such as self-esteem, persistence, concentration, resilience, motivation, ambition as well as 'know-how', contacts and networks, which are distributed unequally amongst different socio-economic groups, in determining future outcomes (Carneiro & Heckman, 2003; Heckman, Stixrud, & Urzua, 2006).

Blanden, Gregg, & Machin (2005) found that levels of intergenerational income mobility in the UK were on a par with those of the US (an intergenerational partial correlation of 0.271 and 0.289, respectively), but lower than Canada or Nordic countries (0.143 for Canada, Sweden and Denmark[2]). Comparing the 1958 and 1970 British birth cohorts, they also found evidence that levels of intergenerational mobility (measured as relative income mobility) had fallen in the UK (intergenerational partial correlation rose from 0.166 to 0.286).

[2] Note that the higher the level of intergenerational partial correlation of incomes, the more a son's wage is determined by his father's and therefore the lower the level of social (income) mobility. Males have traditionally been considered in this type of analysis to remove complications associated with females' labour market participation when raising a family.

In later analysis, Blanden, Gregg, & Macmillan (2007) went on to confirm that over 80% of the decline in mobility between the 1958 and 1970 cohorts can be accounted for by an increase in educational inequality due to better access to higher education, attainment at age 16 and labour market attachment of higher socio-economic groups. They also highlighted the increasing importance of non-cognitive skills in driving changes in mobility mediated through educational attainment.

There is evidence from the UK that gaps in achievement at age 16 by family income have started to narrow over the last decade (Gregg and Macmillan 2010). And whilst more recent findings have shown that income mobility has fallen as the 1958 and 1970 cohorts have aged, there is evidence that levels of educational inequality amongst more recent cohorts on a range of measures in the UK are falling as average educational achievement rises (Blanden & Macmillan, 2014). Nevertheless, these gaps remain substantial, particularly at high levels of attainment, and with educational qualifications being such a strong determinant of later life income and opportunities, such achievement gaps create a major obstacle to future levels of social mobility in the UK (Goodman, Gregg, & Washbrook, 2011).

1.5 Evidence from Scotland

Using data from the Scottish longitudinal study, Growing Up in Scotland (GUS), Bradshaw (2011) has shown that children from higher socio-economic groups - whether this is measured by household income, parental education or social class - have higher cognitive ability[3] on average at both 3 and 5, than children whose parents have lower incomes, lower educational qualifications and/or are from a lower social class.

[3] as measured by vocabulary test scores

At age 5, 20% of children in the highest income quintile had below average ability compared to 54% in the lowest income quintile, with the largest difference in ability at age 5 being associated with parental education levels (Scottish Government, 2015). The latter was also most strongly associated with the *change* in cognitive ability between 3 and 5 - the (within-cohort[4]) attainment gap between those with high levels of parental education and those with low levels of education widening between age 3 and 5 (Bradshaw, 2011). Note, however that the difference in vocabulary ability was sensitive to the choice of socio-economic measure – with a narrowing of the gap using income groups and no change using social class. There is also evidence that this gap is falling as the between-cohort gap in vocabulary scores at age 3 has narrowed[5](Scottish Government, 2015).

Evidence from Primary 1 (P1)[6] using data from 2012/13 showed that the relationship between *mean* cognitive development and deprivation, as measured by the Scottish Index of Multiple Deprivation (SIMD), was significant, if weak (correlations of less than 0.3), and masked significant variation in the *actual* scores by income quintile (Tymms, Merrell, & Buckley, 2015). Pupils from the least deprived quintile made more progress in total cognitive development over P1 than those from the most deprived (the equivalent of over 1.3 months more progress), though this varied by the particular construct considered. There was evidence that progress in P1 varied significantly by school, but no evidence that some schools were better at improving equity (i.e. reducing the gap) than others. There was also little evidence that the relationship between deprivation and cognitive development at the start of P1 had changed over the subsequent two cohorts analysed.

---

[4] Birth cohort 1: Children born in 2004/05

[5] From a 7.8 point gap between children in the lowest and highest income quintiles in birth cohort 1 (born in 2004/05) to a 6.3 point gap in birth cohort 2 (born in 2010/11)

[6] first year of primary school in Scotland when children are generally aged 5

In a nationally representative repeated cross-sectional survey of pupils in P4 (aged around 8), P7 (11) and S2 (13), which assesses pupil's performance in numeracy and literacy in alternate years, pupils from the least deprived areas showed statistically significantly higher performance than pupils from the most deprived areas – with larger gaps at the higher stages (Scottish Government, 2016a). This was true across all stages measured and in both numeracy (2015) and literacy (2014). Some of the gaps have also changed over time: the P4 writing performance gap decreased between 2012 and 2014, the P4 numeracy gap increased between 2011 and 2015 and the S2 reading performance gap increased between 2012 and 2014 (Scottish Government, 2016b).

As in England, there is recent evidence that the attainment gap is narrowing when considering national qualifications (between 2011/12 and 2013/14) and leaver destinations. However the gaps remain significant at higher levels of qualifications: 80% of school leavers from the 20% least deprived areas gained one or more qualifications at SCQF level 6[7] or higher in 2013/14 compared with 39% of those from the 20% most deprived areas. This gap continues into further and higher education: 73.9% of 2013/14 school leavers from the least deprived quintile were in further or higher education in March 2015 compared to only 55.6% of leavers from the most deprived quintile (Scottish Government, 2016a).

[7] broadly equivalent to A-levels

International comparisons using the OECD's Programme for International Student Assessment (PISA), which tests children at 14 years of age, show that the degree to which socio-economic status predicts performance in Scotland irrespective of which school was attended, stands very close to the OECD average (and became smaller in Scotland between 2009 and 2012). The spread of achievement by socio-economic background in Scotland is also narrower than in the OECD as a whole and although socio-economically advantaged students outperform their disadvantaged peers in general terms, many disadvantaged students succeed at school and achieve high levels in the PISA assessments (OECD, 2015).

1.6 Summary and research gap

Given the above, it is unsurprising, therefore, that the attainment gap by family background has played, and continues to play, an important part in government policy in Scotland as in the rest of the UK. Successive Governments have sought evidence on how early on the gap is found, how it evolves through childhood, how important it is in determining future outcomes and the role that early intervention, the education system and student funding policies can play in overcoming – or at the very least reducing – these inequalities. Policies that have sought to address this problem directly include the Sure Start programme and the Pupil Premium in England[8] and the Early Years Framework[9] and Scottish Attainment Challenge in Scotland[10].

[8] https://www.gov.uk/guidance/pupil-premium-information-for-schools-and-alternative-provision-settings - last accessed July 2016

[9] http://www.gov.scot/Topics/People/Young-People/early-years/delivery/framework - last accessed July 2016

[10] http://www.educationscotland.gov.uk/inclusionandequalities/sac/index.asp - last accessed July 2016

White (1982), in his early meta-analysis, showed not only that the observed relationship between an individual aspect of socio-economic circumstances and cognitive development varies depending on the precise choice of SES risk factor (income, parental education levels, job classification or other) but that it increases as more aspects of SES are included in a study. Hence if one is to stand any chance of identifying the separate contribution of any particular association one requires reliable data on correlated aspects of parental socio-economic status (Greg J. Duncan, Yeung, Brooks-Gunn, & Smith, 1998).

This research will focus on income-based measures of socio-economic status – in particular a family's experience of poverty. As family income can be volatile and in order to capture the cumulative impacts of being in persistent poverty, it is important to capture permanent or inter-temporal measures of income (Blanden & Machin, 2010; Gregg & Machin, 1998).

Using work that has taken place in Scotland (Bradshaw, 2011) and the UK (Dearden et al., 2011) as a starting point, the present research will add to the literature and policy debate on the nature, dynamics and associated risks of the gap in cognitive ability and educational attainment in the early years of childhood in Scotland by family background in several ways.

First it will determine the size of the socio-economic gap and gradient[11] in cognitive ability and education attainment in young children in Scotland. Second, it will assess the particular role of experience of poverty in shaping this gap or gradient. Third it will take into account any risk factors associated with poverty and ability or attainment to control for any potential confounding factors to determine poverty's remaining independent contribution (if any). Fourth, it uses the Scottish sample of the Millennium Cohort Study, a longitudinal dataset, to take into account the duration of experiencing poverty. Fifth, it investigates the role of aspects of the home environment as potential protective factors, reducing the risk effect of experience of poverty over and above the other control factors included in the model. Sixth, it extends the analysis to include the first two years of primary school to identify any continuing impact of early home environment factors and of being at school on the socio-economic gradient of educational achievement.

The structure of the remainder of the paper is as follows. Section 2 introduces the research questions which will form the basis of the analysis. Section 3 outlines the methodology used – in particular the data source, the analytic sample and strategy and the choice of outcome variable, measure of experience of poverty and other associated risk/causal factors. Section 4 presents the results from the two main regressions and compares them. Section 5 deals with conclusions and policy implications. Finally, section 6 highlights limitations and areas for future research.

[11] In this analysis a gap is understood to be the difference in outcome at a specific point in time whereas a gradient is how this gap evolves over time.

2 Research Questions

The following research questions address some of the gaps in the research identified in the literature review, with particular reference to Scotland.

- What evidence is there of a socio-economic gap in the cognitive skills and attainment of young Scottish children from age 3 through to age 7, covering the early years of primary school? (RQ1)
- How does the socio-economic gradient in cognitive skills and attainment influence how the identified gap widens or narrows over time? (RQ2)
- Which family or child characteristics positively or negatively influence the size of the gradient and does any association with experience of poverty remain after these have been taken into account? (RQ3)
- Do the influence of these characteristics and the independent association of experience of poverty change between age 3 to 5 and age 5 to 7? (RQ4)

3 Methodology

3.1 Data source and analytic sample

The data is drawn from the Millennium Cohort Study (MCS). The MCS is an on-going survey of 18,818 babies born between September 2000 and January 2002 into 18,552 families in the UK (Hansen, 2014)[12] . Data collections took place when the children were aged 9 months, 3, 5,7 and 11 years. The age 14 survey was in the field throughout 2015/early 2016 and a new survey of the cohort when age 17 will take place in 2018.

In this research we use data from the first four sweeps of data collection from age 9 months in 2001/02 to age 7 in 2007/08. Data have been collected from parents, cohort members and teachers using personal interviews, self-completion questionnaires and a range of administered tests and assessments. The dataset includes information on socio-economic and socio-demographic family characteristics, children's cognitive, social, emotional and behavioural development, their health and wellbeing and that of their family. The survey includes data from all four countries of the UK. At 9 months, 18,552 families were interviewed representing 18,818 children (taking into account twins and triplets).

The longitudinal sample of families in Scotland who participated in each of the first three or four waves of data collection formed the basis of the analytic sample for this research. At 9 months, 2336 families were interviewed representing 2370 children. Age 3 this had fallen to 1814 families (1841 children), at age 5 it remained at 1814 families (1840 children) and by age 7 had fallen to 1628 families (1650 children).

[12] Note the Scottish sample was taken from births between 24 November 2000 and 11 January 2002

The analysis was restricted to families where the natural mother was the main respondent at sweep 1 (S1 - 99% of all families in Scotland), and to only one cohort member per family (i.e. twins and triplets were dropped, n=34, 27, 26 and 26 in sweeps 1 to 4 respectively). Taking into account complete information across all other measures included in the analysis, the complete case analytic sample size was n=1359 for the age 5 analysis (74.9% of families) and n=1229 for the age 7 analysis (75.5% of families). The common complete case analytic sample across both age analyses is n=1160 families.

3.2 Analytic strategy

Data analysis was carried out in Stata 13. In order to establish the existence of a socio-economic gap and the gradient (RQ1 and RQ2), weighted grouped means of t-test scores (with a mean of 50 and standard deviation of 10) were calculated at age 3, 5 and 7 using the common analytic sample (n=1160), over the following three different measures of socio-economic status:

- OECD equivalised income quintile at S1
- Experience of poverty between S1 up to and including S4
- Residing in an area in the lowest quintile for SIMD at 9 months (S1)

Additionally, following Dearden et al., (2011), a table showing the dynamics of the top and the bottom groups in attainment by experience of poverty over the course of the three sweeps is produced.

In order to establish those factors which have the biggest association with change in ability and attainment and to unpick and explain the relationship between of experience of poverty on these, three separate regression models are estimated at age 5 and age 7 (RQ3). The first describes the initial direct relationship between experience of poverty and cognitive ability/attainment – analagous to grouped means (model 1). We then add in household levels of education and prior ability (model 2). Finally, we add in all of the other control measures outlined below (model 3). By including prior attainment we are using a 'value-added' model (Blanden, Greaves, Gregg, Macmillan, & Sibieta, 2015). Alternative specifications using other potential measures of socio-economic conditions and ability/attainment were also estimated to check the robustness of the results to the principal specification.

By comparing the results from our regressions using data from age 5 to that from age 7, we will be able to establish how the influence of factors changes from ages 5 to 7 (RQ4).

OLS estimations and descriptive statistics were calculated with the Stata *svy* command using the appropriate Scotland single country analysis weights to take account of systematic non-response in surveys when carrying out inference to the population (Ketende & Jones, 2011).

3.3 Outcome measures: early cognitive and academic progress

A range of cognitive tests and tests of educational attainment have been included in the MCS from age 3 (Connelly, 2013). This analysis will focus on some of the British Ability Scales II test scores, in particular the naming vocabulary test scores at age 3 and 5, and the word reading score at age 7[13]. The former is closely related to more general measures of cognitive ability (Blanden & Machin, 2010), whereas the latter is a test of educational achievement. Although a child's performance in the word reading test should correlate with their cognitive ability, the former should not be seen as a test of cognitive ability (Connelly, 2013, citing Elliot et al., 1996). This should be borne in mind when looking at trends over time and in interpreting regression results with differing dependent variables and controls for prior ability. The scores adjust for the age of the cohort member and difficulty of the test. The test scores are converted to the same scale with t-scores (mean of 50 and standard deviation of 10) being used in analysing the socio-economic gaps and gradients and z-scores (a mean of zero and a standard deviation of one) being used in the regressions. Equivalent percentile rank scores have been derived[14] and used in the alternative specifications noted above.

Graphical examination and statistical testing of the dependent variables for both complete case samples was carried out. It was found that the dependent variables were sufficiently close to a normal distribution to permit reliable statistical inference to take place[15].

---

[13] In the naming test, the child is shown a series of pictures of objects and is asked to name them. In the word reading test they are asked to read a series of words presented on a card.

[14] using the known distribution of the appropriate normed sample

[15] For more information see sections 7.1 and 7.2 in the Appendix.

3.4 Measures of socio-economic classification

Three separate measures of socio-economic classification were used in this analysis: experience of poverty, Scottish equivalised income quintiles and the Scottish Index of Multiple Deprivation (SIMD).

3.4.1 Experience of poverty

Using the same approach as Parsons, Schoon, & Vignoles (2013) to capture longitudinal experience of worklessness, an experience of poverty indicator was generated by counting the number of times the household equivalised family income (Hansen, 2014) – up to and including the sweep the outcome measure is taken from – had fallen below 60% of the UK median income level, the Scottish government's (Scottish Government, 2007) and UK government's (DWP, 2016) definition of being in relative poverty. As the data is longitudinal, we are able to identify families who were:

- Never in poverty at any sweep
- Moved in and out of poverty (intermittent poverty)
- In poverty at each sweep (permanent poverty)

Using this as the key explanatory variable allows the analysis to take account of the cumulative impact of socio-economic conditions, as highlighted in the literature, but with an additional emphasis on the negative effects of poverty. It should be noted however, that as we don't continuously observe income, it is not possible to say whether a family whose observed income puts them in permanent relative poverty did not experience periods when they were not in poverty. Any findings reached on the effect of poverty are therefore subject to this caveat.

3.4.2 Scottish Index of Multiple Deprivation (SIMD)

Like other UK area based measures of relative multiple deprivation, the SIMD[16] combines scores on several domains of importance in measuring deprivation. Overall (i.e. combined domain) rank deciles were used to create a two-category variable for being in the bottom quintile of SIMD areas or not, a key group comparator used to measure progress against poverty-related targets in Scotland's Child Poverty Strategy (Scottish Government, 2014). The inclusion of this variable allows us to control for any neighbourhood effects referred to above.

3.4.3 Other socio-economic controls

In light of the need to include reliable controls in our regression model for aspects of parental socio-economic status which are likely to be correlated with experience of poverty, but are not exact substitutes, pairwise correlations between alternative measures and experience of poverty were investigated over two time periods (Table 1 and Table 2). Whilst not strictly appropriate when dealing with the relationship between categorical variables, they give an indication of where possible multicollinearity exists and where it does not, maximising the ability of our model to capture the independent association of our chosen measure of SES, experience of poverty, with any change in ability/attainment.

Table 1 Pairwise correlations with S1-S3 experience of poverty

Pairwise correlations with S1-S3 experience of poverty								
	S1-S3 experience of worklessness	S1 highest household parental qualification	S1 Main respondent NS-SEC Current job (5 classes)	S1 Partner NS-SEC Current job (5 classes)	S1 - in bottom quintile, SIMD	S1 OECD income weighted quintiles	S1 Housing tenure	S1 Overcrowded home
Correlation Coefficient	0.69	-0.52	0.43	0.32	0.36	-0.73	0.49	0.28
Obs	1352	1359	824	1004	1359	1359	1359	1359

Using age 3-5 analytic sample as a base, unweighted sample size reported

[16] See http://www.gov.scot/Topics/Statistics/SIMD/BackgroundMethodology for more information on how the SIMD is calculated – last accessed July 2016.

Table 2 Pairwise correlations with S1-S4 experience of poverty

Pairwise correlations with S1-S4 experience of poverty								
	S1-S4 experience of worklessness	S1 highest household parental qualification	S1 Main respondent NS-SEC Current job (5 classes)	S1 Partner NS-SEC Current job (5 classes)	S1 - in bottom quintile, SIMD	S1 OECD income weighted quintiles	S1 Housing tenure	S1 Overcrowded home
Correlation Coefficient	0.70	-0.50	0.43	0.31	0.36	-0.71	0.45	0.33
Obs	1223	1229	758	935	1229	1368	1229	1229

Using age 5-7 analytic sample as a base, unweighted sample size reported

It is clear that experience of poverty is highly correlated with experience of worklessness and income – therefore these will not be included in the model. Whilst household parental qualification is also reasonably highly correlated with poverty, the importance of this characteristic in explaining child cognitive ability/attainment in numerous empirical studies independently of income effects e.g (Bradshaw, 2011; Dearden et al., 2011) argues for its inclusion. Job classification will not be included as it is likely to be highly correlated with parental qualifications[17]. Also, any independent effects associated with job type are likely to be mediated through the home environment, parental values, stress and material wants – all of which are controlled for elsewhere (Parcel & Menaghan, 1990). Overcrowding (Parsons et al., 2013) and housing tenure at time of birth will be included to capture parental stress effects (see section 1.3 above).

3.5 Remaining control variables

There are a large number of potential additional controls in the MCS to include in a regression equation. In order to reduce this to a parsimonious number, the choice of variables was influenced by examining bilateral correlations and cross tabulations of those variables capturing the factors identified in the literature above – those having the strongest association being chosen over others in similar domains.

[17] pairwise correlation between NS-SEC 5 category and highest household qualification is around 0.50 for the mother and 0.52 for the partner

3.5.1 Family demographic characteristics

Included covariates, all taken at age 9 months in order to capture initial conditions (Gregg & Machin, 1998), are marital status, whether only English is spoken in the household and whether any of the carers is black or minority ethnic[18].

3.5.2 Child characteristics

The child's gender and age in months will be included. The latter because the achievement test age standardisation is carried out over a 3 month period so an additional control for age should be present (Connelly, 2013). Birth weight of the child, prior cognitive ability and the prior strength and difficulties score (SDQ) will also be included[19]. Only the responses from main carers for the latter were used in the analysis. These cover issues such as emotional problems, conduct problems, hyperactivity/inattention and peer problems. By combining these it is possible to derive an SDQ total difficulties score for the child [20]. Including prior SDQ scores allows us to control for prior non-cognitive skills levels (Dearden et al., 2011; Heckman et al., 2006; Parsons et al., 2013).

3.5.3 Maternal characteristics

The Rutter Malaise Inventory (Johnson, 2012, citing Rutter, M., Tizard, J. & Whitmore, K. (1970)) is a measure of parental psychosocial distress and is the sum of dichotomous responses to 9 questions administered at 9 months. A higher score indicates higher distress. As only the scores for natural mothers were used, it can be viewed as an indicator of maternal post-natal depression or distress.

---

[18] The small numbers of minority ethnic respondents in Scotland mean that it is not possible to obtain effects by minority ethnic group.

[19] The strength and difficulties questionnaire[19] is a series of 25 questions asked in MCS sweeps 2, 3 and 4 (Johnson, J. (2012).

[20] Other questions, which cover positive attributes such as ability to share, are combined to form a pro-social scale – however only the former will be used in the analysis.

The child/parent relationship scale, completed at age 3 of the child, is formed from 15 self-administered questions on a 5-point scale and involve the respondent's feelings and beliefs about her relationship with her child, and about the child's behaviour towards the mother. The higher the total score, the more positive the relationship (Johnson, 2012). The mother's age at birth of the cohort member will also be included given the negative associations with child development of younger mothers and the protective effects of older mothers found in earlier empirical work (see, for example (Bradshaw, 2011)).

3.5.4 Home environment

Indicators of the home learning environment covering a range of areas were collected at age 3 in the MCS. In order to preserve sample size, however, rather than creating an index of all factors (de la Rochebrochard, 2012), variables were included in the model to indicate frequency of being read to and whether anyone from the household takes the cohort member to the library. A simple additive index of parent teaching activities was derived to capture whether the cohort member was taught the alphabet, numbers or songs in the household at age 3 (Parsons et al., 2013). By only using measures at age 3, we capture initial conditions and minimise problems of endogeneity resulting from positive or negative experiences of the education system (Blanden et al., 2015; White, 1982).

Simple indicators of parenting style included are whether the child has regular term-time bedtimes at age 5 and age 7 and whether they have regular mealtimes at age 5[21].

[21] Whether the child had regular mealtimes was not collected at age 7 but the characteristic at age 5 was still included in the age 7 regression.

3.6 Non-response bias: age 5 sample

The dependent and a selection of key independent variables were investigated to determine whether any non-response bias exists between all the available data and the completed cases used for the analytic sample. Table 3 and Table 4 show that, compared to all data, the analytic sample seems to contain children who are, on average, slightly more able and who come from families which are possibly slightly more well off or educated. Hence it may be that those children who perform relatively badly and/or come from poorer or less educated families may be less likely to respond. This may mean that our results underestimate the association of poverty with the change in cognitive ability or attainment. Nevertheless, the 95% confidence intervals of all means and proportions considered overlap which suggests non-response bias is within acceptable limits.

Table 3 Non-response bias, continuous variables: age 5 sample

	All cases				Completed cases			
	mean	95% CI lower	95% CI upper	n	proportion	95% CI lower	95% CI upper	n
S3 BAS Naming Vocab Score	0.62	0.54	0.70	1778	0.72	0.64	0.80	1359
S2 BAS Naming Vocab Score	0.29	0.21	0.38	1525	0.31	0.22	0.40	1359
S2 SDQ main difficulties	9.01	8.65	9.37	1546	8.79	8.43	9.15	1359
Cohort Age at interview (months) (S3)	63.20	63.00	63.39	1812	63.24	63.02	63.47	1359
Cohort Birth Weight (kg)	3.41	3.38	3.43	1811	3.42	3.40	3.45	1359
Child/Parent Relationship Scale (S2)	64.60	64.16	65.03	1520	64.77	64.30	65.25	1359
Mother's age at birth (years)	28.73	28.17	29.29	1812	29.02	28.37	29.68	1359
Parent teaching score (S2)	2.66	2.63	2.70	1589	2.67	2.63	2.71	1359

Means calculated using appropriate single country weights, unweighted sample size reported.

Table 4 Non-response bias, categorical variables: age 5 sample

	All sample			Completed cases		
	%	95% CI lower	95% CI upper	%	95% CI lower	95% CI upper
Experience of Poverty (S1-S3)						
Never	59.6%	54.6%	64.5%	61.3%	56.3%	66.0%
Intermittent	25.7%	22.3%	29.4%	25.2%	21.7%	29.1%
Persistent	14.7%	12.0%	17.9%	13.6%	10.9%	16.8%
n	1579			1359		
Highest Household Qualification (S1)						
None	7.4%	5.7%	9.7%	5.3%	3.6%	7.6%
NVQ level1	3.4%	2.4%	4.8%	2.7%	1.8%	4.0%
NVQ level 2	22.1%	19.3%	25.1%	21.3%	17.9%	25.0%
NVQ level 3	22.9%	20.8%	25.2%	23.5%	21.0%	26.2%
NVQ level 4	35.1%	31.3%	39.0%	38.1%	33.9%	42.4%
NVQ level 5	8.1%	6.6%	10.1%	8.7%	7.0%	10.8%
Overseas	1.1%	0.6%	1.8%	0.5%	0.2%	1.0%
n	1811			1359		
Marital Status (S1)						
married	58.9%	55.1%	62.7%	62.0%	57.7%	66.2%
cohabiting	25.2%	22.7%	27.9%	23.9%	21.2%	26.8%
single parent	15.9%	13.5%	18.6%	14.1%	11.6%	17.0%
n	1810			1359		
In lowest SIMD quintile (S1)						
No	80.3%	74.3%	85.2%	82.8%	77.2%	87.3%
Yes	19.7%	14.8%	25.7%	17.2%	12.7%	22.8%
n	1810			1359		
Cohort member gender						
Male	51.0%	48.7%	53.2%	50.8%	47.8%	53.7%
Female	49.1%	46.8%	51.3%	49.3%	46.3%	52.2%
n	1812			1359		
Regular meal times (S3)						
no - never	1.6%	1.1%	2.5%	1.4%	0.8%	2.5%
yes - sometimes	4.2%	3.1%	5.7%	4.0%	2.8%	5.8%
yes - usually	94.2%	92.5%	95.5%	94.6%	92.6%	96.1%
n	1808			1359		

Proportions calculated using appropriate single country weights.
Unweighted sample size reported.

3.7 Non-response bias: age 7 sample

The picture is similar for the 5-7 sample (Table 5 and Table 6) with evidence that the completed cases sample contains children who are more able on average and families that are less poor, on average (though the completed cases sample seems to contain more families in persistent poverty than the full sample here). Whilst the caveat about possible bias in the results noted above remains, as before all variables considered have overlapping confidence intervals lending support to our ability to make reasonable inferences about the population from our sample.

Table 5 Non-response bias, continuous variables: age 7 sample

	All cases				Completed cases			
	mean	95% CI lower	95% CI upper	n	mean/ proportion	95% CI lower	95% CI upper	n
S4 BAS Naming Vocab Score	0.69	0.60	0.69	1591	0.75	0.66	0.84	1229
S3 BAS Naming Vocab Score	0.59	0.50	0.69	1529	0.69	0.60	0.77	1229
S3 SDQ main difficulties	7.09	6.81	7.37	1533	6.77	6.48	7.06	1229
Cohort Age at interview (months) (S4)	87.17	86.96	87.37	1624	87.22	86.99	87.44	1229
Cohort Birth Weight (kg)	3.42	3.39	3.45	1626	3.43	3.40	3.46	1229
Child/Parent Relationship Scale (S2)	64.37	63.93	64.81	1384	64.54	64.08	64.99	1229
Mother's age at birth	28.18	27.60	28.75	1627	28.68	28.00	29.35	1229
Parent teaching score (S2)	2.66	2.62	2.70	1442	2.66	2.61	2.70	1229

Means calculated using appropriate single country weights, unweighted sample size reported.

Table 6 Non-response bias, categorical variables: age 7 sample

	All sample			Completed cases		
	%	95% CI lower	95% CI upper	%	95% CI lower	95% CI upper
Experience of Poverty (S1-S4)						
Never	55.9%	50.9%	60.8%	59.6%	54.6%	64.5%
Intermittent	32.7%	28.7%	36.9%	26.2%	22.5%	30.3%
Persistent	11.4%	8.8%	14.7%	14.2%	11.1%	17.8%
n	1368			1229		
Highest Household Qualification (S1)						
None	8.1%	5.9%	11.0%	5.4%	3.7%	7.9%
NVQ level1	3.5%	2.4%	5.2%	3.2%	2.1%	4.8%
NVQ level 2	22.9%	20.1%	26.0%	21.7%	18.3%	25.5%
NVQ level 3	23.8%	21.4%	26.3%	23.6%	21.1%	26.3%
NVQ level 4	33.0%	29.4%	36.8%	37.1%	33.1%	41.3%
NVQ level 5	7.6%	6.0%	9.4%	8.5%	6.7%	10.6%
Overseas	1.2%	0.7%	1.9%	0.5%	0.2%	1.2%
n	1626			1229		
Marital Status (S1)						
married	55.5%	51.3%	59.6%	60.0%	55.5%	64.4%
cohabiting	27.3%	24.4%	30.5%	23.9%	21.2%	28.2%
single parent	17.2%	14.4%	20.3%	14.1%	11.6%	18.4%
n	1625			1229		
In lowest SIMD quintile (S1)						
No	78.1%	71.7%	83.5%	82.0%	76.2%	86.6%
Yes	21.9%	16.5%	28.4%	18.0%	13.4%	23.8%
n	1627			1229		
Cohort member gender						
Male	51.1%	48.6%	53.7%	51.1%	48.0%	54.2%
Female	48.9%	46.3%	51.4%	48.9%	45.9%	52.0%
n	1627			1229		
Regular term-time bedtimes (S4)						
no - never	2.3%	1.6%	3.3%	2.2%	1.4%	3.4%
yes - sometimes	5.0%	4.1%	6.2%	4.7%	3.6%	6.1%
yes - usually	92.7%	91.3%	93.8%	93.1%	91.3%	94.6%
n	1612			1229		

4 Results

4.1 Socio-economic gap and gradient (RQ1&2)

Figure 1 shows the mean test scores at age 3, 5 and 7 for cohort members in the bottom, middle and top income quintile at sweep 1 in the common complete case sample (n=1160). The gap in cognitive ability/attainment between the lowest income quintile (q1) and the top quintile (q5) is statistically significant at each sweep. The mean gap grows between age 3 (S2) and age 5 (S3), from 5.4 to 8.5 points but then remains broadly stable between age 5 and 7 (S4). The gap between those at the middle of the income distribution (q3) and the top (q5) becomes statistically significant at age 5 (5.0 points) and then increases slightly after the first few years of school (5.6 points).

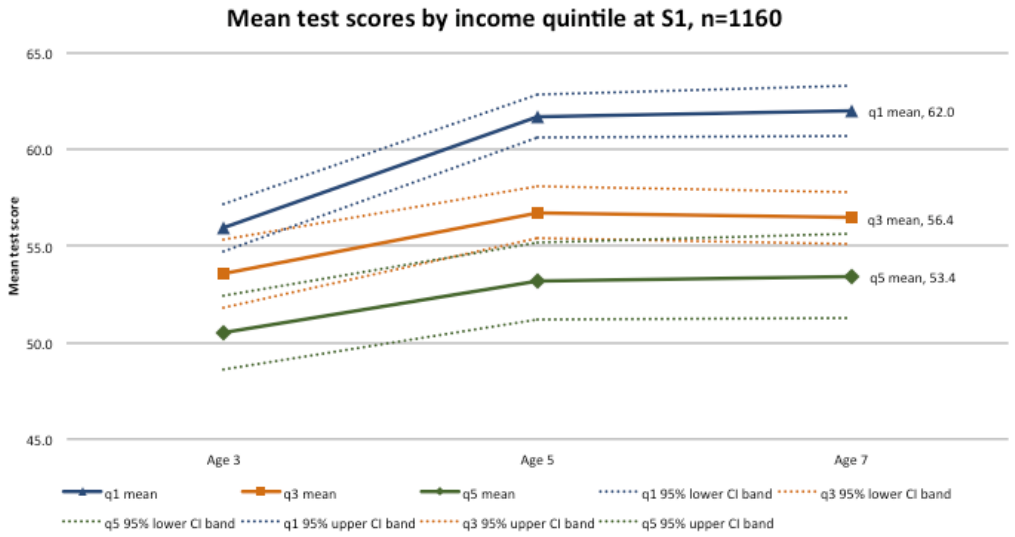

Figure 1 Mean test scores by income quintile (S1): age 3 to age 7

Table 7 shows the cumulative experience of poverty up to and including sweep 2, 3 and 4 for the common analytic sample:

Table 7 Cumulative experience of poverty by sweep, common analytic sample

	S1-S2	S1-S3	S1-S4
Never	68%	64%	58%
Intermittent	16%	24%	32%
Persistent	16%	12%	10%
n	1160	1160	1160

Note: weighted percentages, unweighted sample size

As expected, the experience of poverty increases as the number of time periods increases: 32% of families had experienced at least one period of relative poverty by S2, with this increasing to 36% by S3 and 42% by S4.

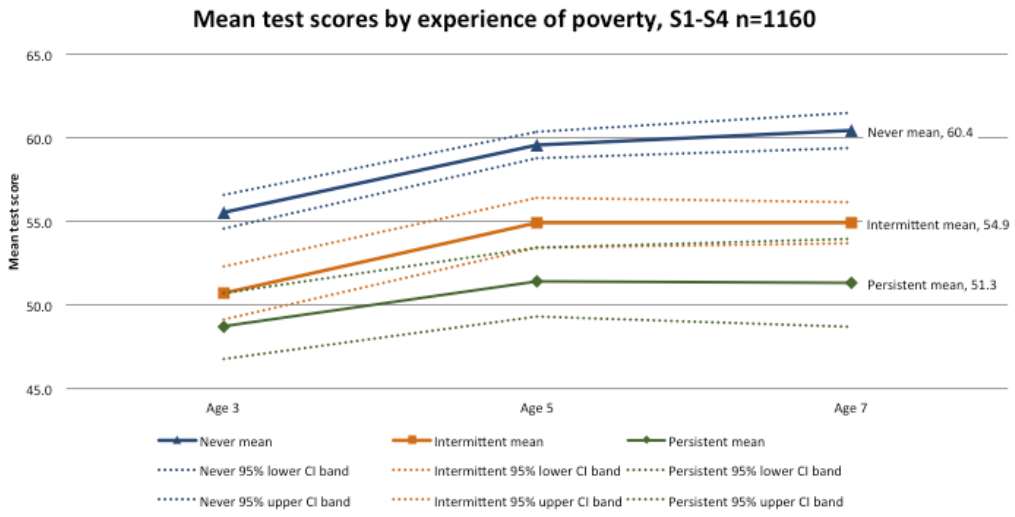

Figure 2 Mean test scores by experience of poverty (S1-S4): age 3 to age 7

Figure 2 shows the relationship between experience of poverty over S1 to S4 and cognitive performance or attainment. We can see a statistically significant and increasing gap exists between the scores of children whose families have never experienced poverty, and those have experienced either intermittent or persistent poverty. The gap between the intermittent and persistent category is not significant in any sweep at the 95% level. The mean gap in test scores between those who have never experienced poverty and those that are in persistent poverty grows from 6.8 points at age 3 (S2) to 8.2 points at age 5 (S3) to 9.1 points at age 7 (S4). Between those children in intermittent poverty and never in poverty, the mean gap drops marginally from 4.8 points at age 3 (S2) to 4.6 points at age 5 (S3) but then increases to 5.5 points by age 7 (S4).

Looking finally at mean test scores grouped by whether a family is located in one the lowest quintile areas of the SIMD at 9 months (around 17.4% of families in the common analytic sample) yields the following results (Figure 3):

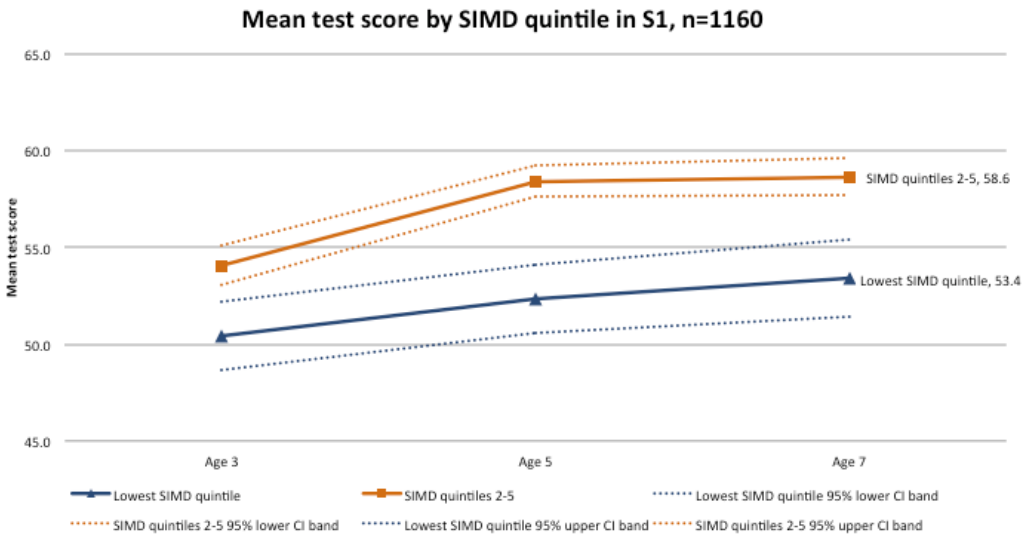

Figure 3 Mean test score by SIMD quintile (S1): age 3 to age 7

This shows that a statistically significant gap in mean scores exists between those children located in area of multiple deprivation at 9 months and those that are not, at age 3, 5 and 7. The mean gap gets bigger between the age of 3 and 5 (3.6 points to 6.1 points) but then narrows again by the age of 7 (to 5.2 points).

Note that the standardised mean scores seem to be on an upward trend – this could be due to those families with lower scores being more likely to drop out which is dealt with by using the same sample at each sweep and updated sample weights. It may also suggest that Scottish test scores are improving with respect to the standardised sample (which is UK based).

To examine the dynamics of movement over time of children in various parts of the distribution we can look at the following table (Table 8).

Table 8 Dynamics of test scores by experience of poverty

	Escape from bottom 20%	Fall from top 20%
Age 3-5		
Never	86%	22%
Intermittent	74%	41%
Persistent	77%	58%
Age 3-7		
Never	94%	31%
Intermittent	79%	54%
Persistent	79%	37%
Age 5-7		
Never	76%	34%
Intermittent	76%	53%
Peristent	51%	51%

Note: weighted percentages, unweighted n=1160

This shows the percentage of children in the bottom or the top quintile for performance at the lower age who 'escape' or 'fall' respectively by the upper age (Dearden et al., 2011). Whilst caution should be used in interpreting the percentages due to regression to the mean effects (Jerrim & Vignoles, 2013) and small sample sizes (particularly for those children in persistent poverty), the general trend is clear: those who experience poverty are less likely to escape relative low levels of cognitive ability/attainment or maintain relative high levels of cognitive ability/attainment over time, than those children who have never experienced poverty.

For example, by age 7, over 90% of those children who have never experienced poverty and were in the bottom 20% of cognitive ability at age 3 have managed to 'escape'. This is nearer 80% for those who have experienced poverty of any type. Similarly, just over 30% of those who have never experienced poverty can no longer maintain their position in the top quintile of performance from age 3 to age 7, compared to over 50% for those in intermittent poverty.

Comparing the figures between 3 to 7 and 5 to 7, those children falling from the top 20% is relatively stable across the never and intermittent categories, however the proportions escaping relatively low performance are now quite similar in the first few years of school.

4.2 Regression equation

The estimated equations at age 5 and 7 are as follows (subsets for Models 1 & 2):

$$y_{it} = \alpha_t + \beta_t s_t + \gamma_{1t} y_{it-1} + \gamma_{2t} e_1 + \delta_{jt}\, c_{ijt} + \varepsilon_{it}$$

where y_{it}= z-score BAS attainment/ability at age 5 or 7 for cohort member i

s_t = experience of poverty up to and including the relevant age (Models 1, 2 & 3)

y_{it-1}= ability at prior age (Models 2 & 3)

e_1= highest level of qualification in the household at cohort age 9 months (Models 2 & 3)

c_{ijt}= a vector of additional control variables (Model 3 only)

α_t= constant

ε_{it}= error term

4.3 Ability at age 5 (RQ3)

4.3.1 Descriptive statistics

Table A in the Appendix shows the grouped mean value of BAS Vocabulary z-score at age 5 for each of the covariates for the age 5 complete case analytic sample. Continuous variables have been grouped into quintiles or another appropriate groupings[22]. An F-test of equality of the grouped means was carried out and the p-value reported. There is little evidence in the data that, before taking account of other characteristics, having a least one non-white parent or speaking a language other than English in the home is associated with different levels of cognitive ability at age 5. This is mainly due to small sample sizes generating large confidence intervals for these groups. In terms of maternal characteristics, there is little evidence of a relationship between maternal distress around 9 months or their experience of long-standing illness. In terms of child characteristics, having a long-standing illness, birth weight and gender are not strongly associated with cognitive ability. Other relationships are as expected with any unusual results being explained by wide confidence intervals crossing zero.

4.3.2 Multiple regression results and key findings at age 5

Turning now to the multiple regression results from the OLS regression for attainment at age 5, these can be seen in Table 9.

[22] Note that test data is grouped according to the relevant normed sample.

Table 9 OLS regression results at age 5

	Model 1	Model 2	Model 3
	b/se	b/se	b/se
Experience of Poverty (S1-S3)			
Base category: Never			
Intermittent	-0.427***	-0.160**	-0.051
	(0.06)	(0.07)	(0.07)
Persistent	-0.732***	-0.329***	-0.166
	(0.08)	(0.10)	(0.11)
S2 BAS Naming Vocab z-score		0.382***	0.353***
Range: -3 - 3		(0.03)	(0.03)
Cohort Age at assessment (months) (S3)			-0.009
Range: 52.8-72.2			(0.01)
Highest Household Qualification (S1)			
Base category: None			
NVQ level 1		0.178	0.188
		(0.20)	(0.19)
NVQ level 2		0.122	0.039
		(0.11)	(0.11)
NVQ level 3		0.141	-0.014
		(0.11)	(0.12)
NVQ level 4		0.301**	0.078
		(0.13)	(0.14)
NVQ level 5		0.534***	0.301**
		(0.14)	(0.15)
Overseas		-0.173	-0.196
		(0.30)	(0.28)
In lowest SIMD quintile (S1)			-0.256***
0=no, 1=yes			(0.06)
Child Non-cognitive Skills (SDQ) (S2)			-0.007
Range: 0-30			(0.01)
At least one parent BME (S1)			0.011
0=no, 1=yes			(0.17)
Housing Tenure (S1)			
Base category: home owner			
Rent(social)			0.065
			(0.08)
Rent(private)			-0.199
			(0.16)
Other			0.014
			(0.09)
Overcrowded home (S1)			-0.263***
0=<1, 1=1+per room			(0.09)
Language spoken at home (S1)			-0.489**
0=Eng. Only 1=Also another lang.			(0.19)

	Model 1	Model 2	Model 3
	b/se	b/se	b/se
Marital Status (S1)			
Base category: Married			
cohabiting			-0.055
			(0.07)
single parent			-0.015
			(0.10)
Child/Parent Relationship Scale (S2)			0.005
Range: 36-75			(0.01)
Mother depression score (S1)			0.003
Malaise scale, Range: 0-9			(0.01)
Mother's age at birth			0.006
Range: 14 to 49			(0.00)
Mother experience of long term illness (S1-S3)			
Base category: Never			
Intermittent			-0.013
			(0.07)
Persistent			0.109
			(0.08)
Child is read to (S2)			0.088***
Range: 1=Not at all…6=Every day			(0.03)
Take to the library (S2)			0.115**
0=no, 1=yes			(0.05)
Parent teaching score (S2)			-0.006
Range: 0-3			(0.06)
Regular meal times (S3)			
Base category: No - never			
Yes - sometimes			0.496**
			(0.22)
Yes - usually/always			0.408**
			(0.18)
Regular term-time bedtime (S3)			
Base category: No - never			
Yes - sometimes			-0.126
			(0.20)
Yes - usually/always			0.099
			(0.13)
Child has a LS illness (S3)			-0.051
0=No, 1= Yes			(0.07)

	Model 1	Model 2	Model 3
	b/se	b/se	b/se
Cohort Birth Weight (kg)			-0.068
Range: 0.57 to 6.55			(0.06)
Cohort member female			-0.123**
0=boy, 1=girl			(0.06)
constant	0.927***	0.461***	0.05
	(0.04)	(0.13)	(0.66)
obs	1359	1359	1359
R-sqr	0.076	0.262	0.310
F-test	<0.001***	<0.001***	<0.001***

* p<0.1, ** p<0.05, ***p<0.01

As the dependent variable is standardised with a mean of zero and standard deviation (s.d.) of 1, the derived coefficients show the effect size on ability in terms of standard deviations. Model 1 shows that intermittent and persistent poverty have a substantial and significant negative association with changes in ability scores at age 5, with persistent poverty having the strongest association. Model 2 shows that this is substantially, but not exclusively, mediated through levels of parental education and prior ability. Following the introduction of the full range of controls in model 3, however, experience of poverty no longer has a statistically significant independent association with changes in cognitive ability between 3 and 5 years old.

Of the other socio-economic risk factors associated with experiencing poverty, only a very high level of parental education was a protective factor whereas being in an area of relative deprivation and suffering overcrowding were significantly and negatively associated with changes in cognitive ability before school. Whilst it seems that speaking more than one language at home has a strong negative association with cognitive development, this is unduly influenced by the small number of families in this category (n=16).

In terms of child characteristics, prior cognitive ability was identified as a strong protective factor, however there is no evidence that lower non-cognitive skills in terms of a higher SDQ total difficulties score are associated with cognitive development in model 3. Being a girl is associated with slower cognitive development between 3 and 5 than boys, *ceteris paribus*.

In terms of the home environment, the strongest protective factor for families is having regular meal times. Looking specifically at the home learning environment, being read to at age 3 (especially if read to frequently) is the strongest protective factor and being taken to the library at age 3 is also positively associated with faster cognitive development to age 5. More general parent-teaching activities or regular bed times do not have a significant association with cognitive development over this time period.

4.3.3 Group significance tests and alternative specifications

Wald tests of joint significance were undertaken on grouped variables in model 3. It was impossible to reject the null of no group significance for all these, including the experience of poverty categories, apart from highest household qualification (p-value = 0.03 – significant at the 5% level) and regular meal times (p-value=0.07 – significant at the 10% level)

Alternative specifications of model 3 were run as follows to check the robustness of the results to the choice of income measure and outcome measure (results in Table B in the Appendix presented along with model 3 results for comparison purposes):

- OECD equivalised income quintile at sweep 3 rather than experience of poverty
- percentile ranks of test scores as the dependent and explanatory variable rather than z-scores

Apart from some minor changes in significance and effect level (reflecting the differing dependent variable in the percentile regression and the fact that small changes in standardised score can have a large impact on percentile rank in the middle of the distribution), the results were remarkably robust to the different specifications. The one interesting result was the loss of significance on the NVQ level 5 category and the significance of the highest income quintile in the equivalised income quintile specification. This suggests that the association of higher parental qualifications with child cognitive development may be mediated through a pure income effect not captured in the 'never experienced poverty' group.

Using r^2, we can see that model 3 was able to explain around 31% of the variation in cognitive development at age 5 (compared to around 8% and 26% for models 1 and 2 respectively). We also comfortably reject the null of no explanatory power for the additional variables going from model 2 to model 3 at the 1% level (F test of joint significance – p-value <0.001). Standard checks of normality and heteroscedasticity of the residuals were carried out and the influence of outliers on the results was checked.[23]

4.4 Attainment at age 7 (RQ3)

4.4.1 Descriptive Statistics

As for the age 5 regression, Table A shows the grouped mean value of BAS Reading z- scores at age 7 for each of the covariates for the age 7 complete cases. A similar pattern of results can be observed here as for age 5. Notable differences are the stronger gradient of the means for prior non-cognitive skills at age 7, a stronger correlation with parental teaching and a more consistent increase in mean test scores by frequency of being read to.

4.4.2 Multiple regression results and key findings at age 7

We now turn to the results from the OLS multiple regression at age 7 (Table 10).

[23] See section 7.5 in Appendix for more detail

Table 10 OLS regression results at age 7

	Model 1	Model 2	Model 3
	b/se	b/se	b/se
Experience of Poverty (S1-S4)			
Base category: Never			
Intermittent	-0.578***	-0.295***	-0.248***
	(0.07)	(0.08)	(0.09)
Persistent	-0.877***	-0.434***	-0.383**
	(0.13)	(0.12)	(0.16)
S3 BAS Naming Vocab z-score		0.350***	0.316***
Range: -3 - 3		(0.04)	(0.04)
Cohort Age at assessment (months) S4			-0.025**
Range: 77.4-96.5			(0.01)
Highest Household Qualification (S1)			
Base category: None			
NVQ level 1		0.028	0.155
		(0.31)	(0.26)
NVQ level 2		0.034	0.07
		(0.18)	(0.16)
NVQ level 3		0.156	0.136
		(0.20)	(0.19)
NVQ level 4		0.25	0.209
		(0.19)	(0.19)
NVQ level 5		0.381	0.303
		(0.25)	(0.23)
Overseas		-0.131	0.103
		(0.18)	(0.17)
In lowest SIMD quintile (S1)			-0.003
0=no, 1=yes			(0.11)
Child Non-cognitive Skills (SDQ) (S3)			-0.038***
Range: 0-28			(0.01)
At least one parent BME (S1)			0.199
0=no, 1=yes			(0.16)
Housing Tenure (S1)			
Base category: home owner			
Rent(social)			-0.103
			(0.09)
Rent(private)			0.151
			(0.24)
Other			0.061
			(0.15)
Overcrowded home (S1)			0.039
0=<1, 1=1+per room			(0.13)
Language spoken at home (S1)			0.136
0=Eng. Only 1=Also another lang.			(0.20)

	Model 1	Model 2	Model 3
	b/se	b/se	b/se
Marital Status (S1)			
Base category: Married			
cohabiting			-0.026
			(0.07)
single parent			-0.004
			(0.10)
Child/Parent Relationship Scale (S2)			-0.001
Range: 36-75			(0.01)
Mother depression score (S1)			0.003
Malaise scale, Range: 0-9			(0.02)
Mother's age at birth			-0.001
Range: 15 to 49			(0.01)
Mother experience of long term illness (S1-S4)			
Base category: Never			
Intermittent			-0.043
			(0.07)
Persistent			-0.059
			(0.10)
Child is read to (S2)			-0.019
Range: 1=Not at all…6=Every day			(0.04)
Take to the library (S2)			0.133*
0=no, 1=yes			(0.08)
Parent teaching score (S2)			0.109*
Range: 0-3			(0.06)
Regular meal times (S3)			
Base category: No - never			
Yes - sometimes			0.25
			(0.28)
Yes - usually/always			0.149
			(0.22)
Regular term-time bedtime (S4)			
Base category: No - never			
Yes - sometimes			-0.279
			(0.24)
Yes - usually/always			-0.298
			(0.19)

	Model 1	Model 2	Model 3
	b/se	b/se	b/se
Child has a LS illness (S4)			-0.071
0=No, 1= Yes			(0.10)
Cohort Birth Weight (kg)			-0.048
Range: 0.57 to 6.55			(0.05)
Cohort member female			0.073
0=boy, 1=girl			(0.06)
constant	1.034***	0.486**	3.053**
	(0.05)	(0.19)	(1.17)
obs	1229	1229	1229
R-sqr	0.087	0.193	0.240
F-test	<0.001***	<0.001***	<0.001***

* p<0.1, ** p<0.05, ***p<0.01

Experience of poverty has a significant negative association with educational progress between age 5 and 7 in model 1 with a stronger effect for persistent than intermittent poverty. Model 2 shows this association is partly, but not fully mediated through levels of parental education and prior cognitive ability. When a full range of controls in included in model 3, a significant and negative association between educational development in the early years of primary school and intermittent and persistent poverty remains.

In terms of child characteristics, prior cognitive ability was identified as a strong protective factor, however, lower non-cognitive skills in terms of a higher SDQ total difficulties score and being older are negatively associated with cognitive development in model 3 – though the age coefficient may simply be capturing the lack of precision in the age standardisation process.

In terms of the home environment, being taken to the library at age 3 and parental teaching activities were protective factors associated with educational progress from 5 to 7, however this was only at a low level of significance.

4.4.3 Group significance tests and alternative specifications

Wald tests of joint significance were undertaken on grouped variables in model 3. These yielded no surprising results though here, in contrast to the age 5 regression, we were able to reject the null of joint-zero coefficients on the experience of poverty variables at the 5% level (p-value=0.018).

As in the age 5 regression, similar alternative specifications were run to examine the robustness of the results (Table D)[24]. Using percentile ranks reduces the level of significance of the experience of poverty factors and having a regular bedtime becomes significant at the 10% level. Also, girls seem to develop more quickly when using percentile ranks. However, using standardised scores is a more reliable way to compare impacts due to bunching of percentiles at the mean (Connelly, 2013). As with the age 5 regression, using current income quintiles, only the top income quintile is significant at the 5% level (effect size 0.3). Otherwise the results are remarkably similar across specifications.

Using r^2, we can see that model 3 was able to explain around 24% of the variation in educational progress between age 5 and 7 (compared to around (9% and 19% for models 1 and 2 respectively). We also comfortably reject the null of no explanatory power for the additional variables going from model 2 to model 3 at the 1% level (F test of joint significance – p-value <0.001). As for the age 5 regression, standard checks of normality and heteroscedasticity of the residuals were carried out and the influence of outliers on the results was checked.[25]

4.5 Comparison of age 5 and age 7 results (RQ4)

Figure 4 shows the evolution of regression coefficients on the experience of poverty measures across the different specifications and sweeps. Only coefficients which are significant at the 5% level or below are shown.

---

[24] Equivalised income quintiles at S4 were used in the age 7 regression

[25] See section 7.9 in Appendix for more detail

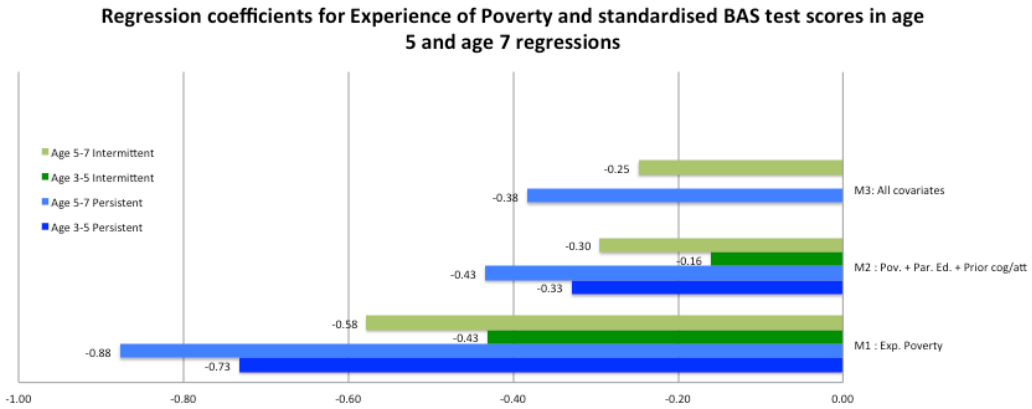

Figure 4 Comparison of regression coefficients on experience of poverty measures: age 5 and age 7 regressions

Whilst comparisons between sweeps should bear in mind that the measures used are different it is still helpful to talk about broad trends in the findings. It seems plausible to suggest the following:

- experience of poverty is more strongly associated with lower levels of cognitive/educational development from age 5 to 7, than from age 3 to 5, for each of the specifications considered
- persistent experience of poverty has a consistently larger negative association with development than intermittent poverty
- the introduction of controls which are associated with experience of poverty and cognitive/educational development reduce the size of the former's association with the latter.
- nevertheless, in the fully specified models, the negative association between educational development and experience of poverty remains substantial and significant between the ages of 5-7 – with evidence being stronger for intermittent than persistent poverty.

Considering the remainder of the results from model 3 for the age 5 regression and the age 7 regression, they suggest the characteristics associated with the development of cognitive ability before school and then educational attainment as the child enters primary school, change. Inclusion of the full range of controls at age 5 is able to reduce the explanatory power of the experience of poverty variables to the point where they are no longer statistically significant in our model. Prior ability has a strong positive association with changes in ability between age 3 and 5 and there is some evidence of continuing positive association of parental education levels in development from age 3 to age 5 for very highly educated households. Family circumstances are also associated with the dynamics of cognitive ability between ages 3 and 5, with location in an area of relative deprivation and experience of overcrowding at birth all being associated with a negative effect on cognitive development. However, having a stable home environment in general – as measured by regular meal times is a strong protective factor associated with improvements in cognitive ability.

In terms of the home learning environment, being read to often and being taken to the library are stronger protective factors than specific parent teaching activities for cognitive development between 3 and 5.

The pictures changes, somewhat during the first two years of primary school in Scotland from 5 years to 7. Inclusion of the full range of controls no longer explains away differences in educational progress[26] associated with experience of poverty – leaving a residual negative association. Prior cognitive ability continues to have a strong protective role in early educational attainment to age 7 as in cognitive development to age 5, however, there is no evidence of a continuing association of levels of parental education suggesting the influence of this may have crystallised earlier (or may be felt again later) or are mediated through the experience of poverty factor.

[26] Strictly speaking we cannot talk about gains in attainment since it is cognitive ability which is measured at age 5 and educational attainment or achievement at age 7

Family conditions considered, including neighbourhood effects, are not associated with early educational development and evidence of the association of the home learning environment or parental approach is weak, at best. This may be because the earlier tests of vocabulary are more strongly affected by the home environment than a reading test (Blanden et al., 2015). What does seem to be emerging is an association of prior non-cognitive skills with educational progress. Certainly, poorer performance on the SDQ total difficulties questionnaire at age 5 is associated with poorer progress in the reading test by age 7, *ceteris paribus*. Nevertheless, the conclusion remains that there seem to be aspects of experiencing poverty's association with educational progress at the start of primary school which are not captured by our other controls.

It is also important to note that a large percentage of the variation in cognitive ability at age 5 or attainment at age 7 remains unexplained in the fully-specified models (around 70% in the age 5 model and 75% in the age 7 model). There are clearly many more factors other than experience of poverty underlying the variation in educational/cognitive performance of children.

5 Conclusions and policy implications

Addressing the so-called 'socio-economic gap in attainment' is a current policy priority in Scotland – though what this actually means is quite unclear (Denholm, 2016). This research has improved the evidence base on this gap using evidence from the Scottish sample of the Millennium Cohort Study up to age 7. This is in line with recommendations in previous research to address this lack of an evidence base (Ellis & Sosu, 2014)

The aim of the analysis was to establish the existence of and the factors associated with the socio-economic gap and gradient in cognitive ability and attainment in the early years of childhood and primary school in Scotland, with a particular emphasis on the experience of poverty.

The literature search focused on the theoretical underpinning of this relationship and what the possible causal factors, of both the level of any gap and the evolution of this gap, could be. Concerns about trends in inter-generational income mobility within nations and the extent to which these are determined - through access to higher education, better paid jobs and lower levels of unemployment - by family socio-economic and demographic circumstances and early childhood experiences drive much of this research.

Theories of child development, education and economics have been linked together to attempt to understand how family circumstances and early childhood experiences determine developments in cognitive ability and educational attainment in children in the early years and how this can be used to predict economic and educational outcomes into early adulthood and beyond (Feinstein, 2003).

Earlier research on Scottish data identified a socio-economic gap in cognitive development between age 3 and age 5 and emphasised the impact of parental education on pre-school developments in cognitive ability (Bradshaw, 2011). Similar gaps have been found at a UK level using MCS data – though using different measures of socio-economic status (Blanden & Machin, 2010; Dearden et al., 2011; Parsons et al., 2013).

The current research has confirmed that this cognitive or attainment gap exists in Scotland beyond the age of 5 into the early years of primary school on several measures of socio-economic classification. (RQ1)

Whilst the mean performance of children at all socio-economic levels in Scotland is on a general upward trend, there is evidence that the gradient and therefore the gap is getting larger – in particular between those who have experienced no poverty and those who have experienced persistent poverty (or those at the highest income quintiles versus everyone else) – (RQ2).

Whilst the change in this gradient from age 3 to 5 associated with experience of poverty is to a large extent mediated through prior ability levels and the general home environment, the association between poverty and the gradient from age 5 to 7 is less well explained – with experience of poverty continuing to be negatively associated with educational progress in the early years of primary school over and above any of the characteristics controlled for in the regression analysis (RQ3 and RQ4).

Any policy implications drawn are subject to the caveats that the sample size used is relatively small and one must be careful when drawing conclusions about causality from associations. Nevertheless, given that the analysis uses a rich dataset and controls for many of the potential confounding factors identified in the literature, policy prescriptions may justified.

The results suggest that from age 3 to age 5, it is prior ability, very high levels of parental education and wider home circumstances and environment (rather than specific learning activities, *per* se) which are most strongly associated with gains in cognitive ability. This argues for the importance of early intervention from birth to influence early child development and cognitive ability (before 3 years old) and the potential role of parental classes in establishing good parenting styles early on for those groups most at risk. There may also be a role for government in addressing wider neighbourhood aspects of poverty captured in area deprivation given its (negative) association with changes in cognitive ability by age 5. From age 5 to age 7, policies designed to more directly alleviate families in relative poverty – particularly if this is persistent - may have a positive impact on educational attainment, even if the causal mechanism is unclear. However, the increasing association of a child's non-cognitive skills with progress in educational attainment argues for the importance of schools to work closely with parents to address any behaviour concerns early on.

6 Limitations and areas for future research

The problem of establishing causality in the light of potential unobserved mediating variables is a significant factor limiting this research and research in this area more widely. In particular it is difficult to observe parental endowments – and income may be a poor indicator of this (Brooks-Gunn & Duncan, 1997), though using inter-temporal measures of income or poverty, as we do here, should mitigate this somewhat. Taken further, some have suggested that 'poor' parents are somehow different from 'rich' parents in ways that are unobservable but that influence both income and child outcomes i.e. there is no real link between the two (Haveman & Wolfe, 1995).

A more fruitful avenue may be to suggest that the transmission mechanisms between different factors and child outcomes are different for different socio-economic groupings. For example, research in Scotland has suggested that improving infant-maternal attachment and having advanced communication skills at an early age are key protective factors for cognitive development for those parents with no or lower qualifications, whereas for the children of more highly educated parents only infant-maternal attachment mattered and poor early communication was not a barrier to later improvement in vocabulary (Bradshaw, 2011). It would be interesting to explore interactions between characteristics and experience of poverty to see if the mechanisms through which they are associated with performance differ.

It may well be that school quality, being an unobserved variable in our analysis, is having an impact on the results with children in relative poverty more likely to attend lower quality schools. Certainly evidence from Primary 1 (Tymms et al., 2015) and pre-school (Bradshaw, Lewis, & Hughes, 2014) in Scotland and the UK (Melhuish et al., 2008; Sylva et al., 2008) shows that quality matters.

It is clear that conclusions from the results are complicated by the fact that the domains and the tests administered at different ages in the MCS are not the same. Cognitive ability can be estimated by vocabulary skills but is a wider concept. Similarly, cognitive ability is related to but not synonymous with educational attainment. Hence caution should be applied to these results which start off considering the former and then necessarily move on to the latter. It may prove fruitful, therefore, to examine the relationship between the other tests of ability included in the MCS and experience of poverty and also to extend the analysis to age 11 and beyond. This could also be compared with results from GUS when cognitive results for their second birth cohort become available in the future.

7 Appendices

7.1 Tests for normality of dependent variable: complete cases

The dependent variable for complete cases in the age 5 regression and age 7 regression was examined using the standard range of plots. Statistical tests were also applied to test for normality[27]. For the age 5 data, graphical inspection yielded no cause for concern (see Figure A, Figure B and Figure C). For the age 7 data, some truncation can be noted at the top of the distribution (see Figure D, Figure E and Figure F) which influences the rejection of the statistical tests for normality at the 5% level. However, in large samples such as this, even small deviations from non-normality can result in significant tests (Ghasemi & Zahediasl, 2012). Therefore it would seem that our dependent variable is sufficiently close to a normal distribution to permit reliable statistical inference to take place.

[27] Shapiro-Wilk test and a skewness/kurtosis test in Stata using the *swilk* and *sktest* commands respectively

7.2 Normality plots of outcome variables

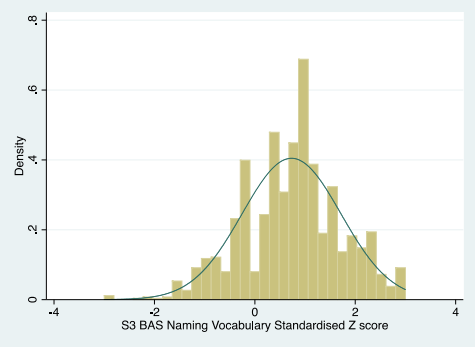

Figure A Histogram of S3 BAS Naming Vocabulary z-score

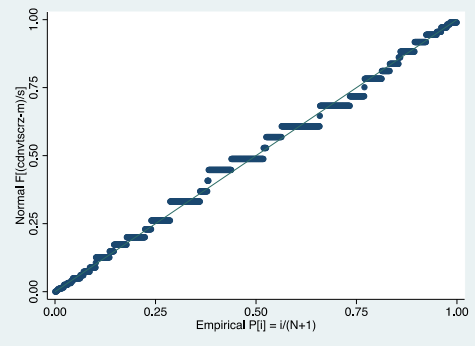

Figure B Normal probability plot for S3 BAS Naming Vocabulary z-score

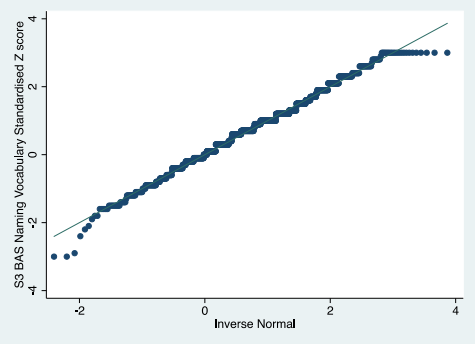

Figure C Qnorm plot of S3 BAS Vocabulary z-score

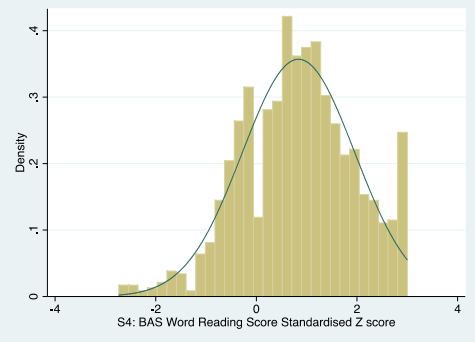

Figure D Histogram of S4 BAS Reading z-score

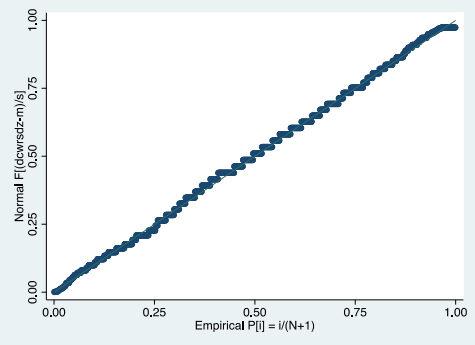

Figure E Normal probability plot of S4 BAS Reading z-score

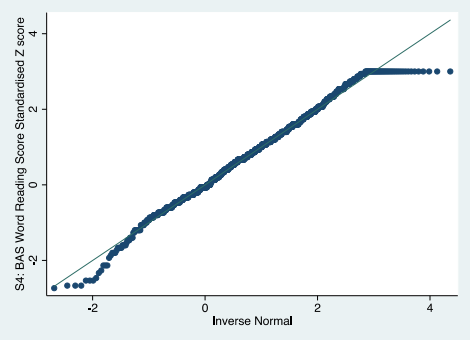

Figure F Qnorm plot of S4 BAS Reading z-score

7.3 Descriptive statistics results: age 5 regression

Table A Grouped mean BAS Vocabulary z-score: age 5

	Mean	Linearised s.e	95% CI Lower	95% CI Upper	F-Test equivelance of means
Experience of Poverty (S1-S3)					
Never	0.927	0.042	0.843	1.011	<0.001***
Intermittent	0.500	0.066	0.368	0.631	
Persistent	0.195	0.071	0.053	0.337	
S2 BAS Naming Vocab z-score					
Lowest quintile	0.078	0.089	-0.100	0.256	<0.001***
2nd quintile	0.301	0.068	0.164	0.438	
3rd quintile	0.627	0.079	0.469	0.786	
4th quintile	0.856	0.053	0.750	0.963	
5th quintile	1.310	0.048	1.213	1.406	
Cohort Age at assessment (months) S3					
Lowest quintile	0.907	0.064	0.779	1.035	0.009***
2nd quintile	0.643	0.059	0.525	0.760	
3rd quintile	0.592	0.073	0.445	0.739	
4th quintile	0.758	0.071	0.616	0.899	
5th quintile	0.760	0.072	0.617	0.903	
Highest Household Qualification (S1)					
None	0.131	0.101	-0.071	0.334	<0.001***
NVQ level 1	0.319	0.173	-0.027	0.664	
NVQ level 2	0.481	0.068	0.345	0.617	
NVQ level 3	0.571	0.049	0.473	0.668	
NVQ level 4	0.945	0.047	0.851	1.040	
NVQ level 5	1.240	0.078	1.083	1.397	
Overseas	-0.017	0.244	-0.505	0.472	
In lowest SIMD quintile (S1)					
No	0.822	0.040	0.743	0.901	<0.001***
Yes	0.229	0.063	0.103	0.355	
Child Non-cognitive Skills (SDQ) (S2)					
Lowest quintile	0.976	0.054	0.869	1.084	<0.001***
2nd quintile	0.782	0.073	0.637	0.927	
3rd quintile	0.728	0.062	0.604	0.852	
4th quintile	0.560	0.057	0.447	0.673	
5th quintile	0.369	0.080	0.209	0.529	
At least one parent BME (S1)					
No	0.725	0.039	0.647	0.802	0.237
Yes	0.490	0.199	0.092	0.888	
Housing Tenure (S1)					
Home owner	0.871	0.040	0.790	0.952	<0.001***
Rent(social)	0.373	0.056	0.261	0.486	
Rent(private)	0.402	0.245	-0.089	0.892	
Other	0.595	0.091	0.413	0.777	
Overcrowded home (S1)					
No overcrowding	0.784	0.038	0.707	0.860	<0.001***
Overcrowding	0.130	0.097	-0.064	0.324	
Language spoken at home (S1)					
English only in the home	0.726	0.038	0.649	0.803	0.010**
Another language spoken in the home	0.109	0.237	-0.365	0.583	

	Mean	Linearised s.e	95% CI Lower	95% CI Upper	F-Test equivelance of means
Marital Status (S1)					
Married	0.847	0.039	0.769	0.926	<0.001***
Cohabiting	0.567	0.072	0.424	0.711	
Single parent	0.417	0.079	0.259	0.574	
Child/Parent Relationship Scale (S2)					
Lowest quintile	0.541	0.072	0.397	0.686	<0.001***
2nd quintile	0.606	0.075	0.455	0.756	
3rd quintile	0.690	0.058	0.575	0.806	
4th quintile	0.853	0.054	0.745	0.961	
5th quintile	0.950	0.063	0.824	1.075	
Mother depression score (S1)					
Lowest quintile	0.768	0.048	0.672	0.863	0.277
2nd quintile	0.706	0.068	0.571	0.842	
3rd quintile	0.776	0.068	0.640	0.913	
4th quintile	0.651	0.108	0.436	0.866	
5th quintile	0.595	0.080	0.434	0.755	
Mother's age at birth					
Less than 20	0.348	0.136	0.076	0.620	<0.001***
20-29	0.606	0.049	0.508	0.703	
30-39	0.826	0.046	0.735	0.918	
40 and over	0.897	0.133	0.631	1.164	
Mother experience of long term illness (S1-S3					
Never	0.750	0.041	0.667	0.833	0.365
Intermittent	0.648	0.067	0.514	0.782	
Persistent	0.727	0.096	0.535	0.918	
Child is read to (S2)					
Not at all	-0.640	0.428	-1.496	0.216	<0.001***
Less often	0.214	0.160	-0.106	0.533	
Once or twice a month	0.086	0.206	-0.326	0.499	
Once or twice a week	0.414	0.069	0.276	0.553	
Several times a week	0.551	0.075	0.401	0.702	
Every day	0.863	0.039	0.785	0.942	
Take to the library (S2)					
No	0.573	0.049	0.475	0.672	<0.001***
Yes	0.879	0.047	0.785	0.972	
Parent teaching score (S2)					
0 - Very low	-	-	-	-	<0.001***
1 - Low	0.712	0.212	0.288	1.136	
2 - Medium	0.731	0.066	0.598	0.863	
3 - High	0.717	0.041	0.635	0.799	
Regular meal times (S3)					
No - never	-0.256	0.222	-0.700	0.187	<0.001***
Yes - sometimes	0.490	0.118	0.255	0.725	
Yes - usually/always	0.744	0.040	0.665	0.824	
Regular term-time bedtime (S3)					
No - never	0.376	0.150	0.075	0.677	0.003***
Yes - sometimes	0.197	0.193	-0.188	0.582	
Yes - usually/always	0.765	0.040	0.685	0.844	

	Mean	Linearised s.e	95% CI Lower	95% CI Upper	F-Test equivelance of means
Child has a LS illness (S3)					
No	0.742	0.039	0.664	0.820	0.081*
Yes	0.617	0.073	0.470	0.764	
Cohort Birth Weight (kg)					
Lowest quintile	0.731	0.074	0.583	0.880	0.878
2nd quintile	0.687	0.087	0.512	0.861	
3rd quintile	0.752	0.065	0.622	0.881	
4th quintile	0.693	0.067	0.558	0.827	
5th quintile	0.740	0.069	0.601	0.878	
Cohort member female					
Boy	0.729	0.049	0.630	0.828	0.772
Girl	0.711	0.050	0.611	0.810	

The weighted sample size for the category 'Very Low' in parent teaching score (S2)
is too small to calculate a mean value for ability at age 5
Note: Weighted means, unweighted sample size = 1359
* p<0.1, ** p<0.05, ***p<0.01

7.4 Alternative specification results: age 5 regression

Table B Alternative specification results: age 5

	Model 3 Original Specification	Model 3 Percentile Ranks	Model 3 OECD income
	b/se	b/se	b/se
Experience of Poverty (S1-S3)			
Base category: Never			
Intermittent	-0.051	-2.02	
	(0.07)	(1.93)	
Persistent	-0.166	-4.849	
	(0.11)	(3.11)	
OECD Single country income quintile (S3)			
Base category: Lowest quintile			
2nd quintile			0.031
			(0.09)
3rd quintile			0.12
			(0.12)
4th quintile			0.145
			(0.11)
5th quintile			0.256**
			(0.11)
S2 BAS Naming Vocab z-score	0.353***		0.354***
Range: -3 - 3	(0.03)		(0.03)
S2 BAS Naming Vocab percentile rank		0.320***	
Range: 0.1-99.9		(0.02)	
Cohort Age at assessment (months)(S3)	-0.009	-0.236	-0.009
Range: 52.8-72.2	(0.01)	(0.20)	(0.01)
Highest Household Qualification (S1)			
Base category: None			
NVQ level 1	0.188	3.777	0.196
	(0.19)	(5.23)	(0.19)
NVQ level 2	0.039	0.344	0.068
	(0.11)	(3.10)	(0.11)
NVQ level 3	-0.014	-0.918	0.004
	(0.12)	(3.63)	(0.12)
NVQ level 4	0.078	0.846	0.063
	(0.14)	(3.92)	(0.14)
NVQ level 5	0.301**	5.362	0.252
	(0.15)	(4.25)	(0.15)
Overseas	-0.196	-9.567	-0.204
	(0.28)	(8.94)	(0.29)
In lowest SIMD quintile (S1)	-0.256***	-7.562***	-0.253***
0=no, 1=yes	(0.06)	(1.62)	(0.06)
Child Non-cognitive Skills (SDQ) (S2)	-0.007	-0.29	-0.007
Range: 0-30	(0.01)	(0.20)	(0.01)
At least one parent BME (S1)	0.011	-0.653	0.03
0=no, 1=yes	(0.17)	(4.80)	(0.17)

	Model 3 Original Specification	Model 3 Percentile Ranks	Model 3 OECD income
Housing Tenure (S1)			
Base category: home owner			
Rent(social)	0.065	2.546	0.064
	(0.08)	(1.91)	(0.08)
Rent(private)	-0.199	-7.899*	-0.203
	(0.16)	(4.07)	(0.15)
Other	0.014	1.006	0.016
	(0.09)	(2.51)	(0.09)
Overcrowded home (S1)	-0.263***	-6.454**	-0.250**
0=<1, 1=1+per room	(0.09)	(2.81)	(0.10)
Language spoken at home (S1)	-0.489**	-12.024*	-0.495***
0=Eng. Only 1=Also another lang.	(0.19)	(6.23)	(0.19)
Marital Status (S1)			
Base category: Married			
cohabiting	-0.055	-2.122	-0.038
	(0.07)	(1.90)	(0.07)
single parent	-0.015	-0.795	-0.015
	(0.10)	(2.73)	(0.10)
Child/Parent Relationship Scale (S2)	0.005	0.124	0.006
Range: 36-75	(0.01)	(0.13)	(0.01)
Mother depression score (S1)	0.003	0.204	0.002
Malaise scale, Range: 0-9	(0.01)	(0.39)	(0.01)
Mother's age at birth	0.006	0.127	0.005
Range: 14 to 49	(0.00)	(0.13)	(0.00)
Mother experience of long term illness (S1-S3)			
Base category: Never			
Intermittent	-0.013	-0.401	-0.012
	(0.07)	(1.67)	(0.07)
Persistent	0.109	2.971	0.128
	(0.08)	(2.25)	(0.08)
Child is read to (S2)	0.088***	2.619***	0.084***
Range: 1=Not at all…6=Every day	(0.03)	(0.84)	(0.03)
Take to the library (S2)	0.115**	2.295*	0.118**
0=no, 1=yes	(0.05)	(1.22)	(0.05)
Parent teaching score (S2)	-0.006	0.529	-0.007
Range: 0-3	(0.06)	(1.35)	(0.06)
Regular meal times (S3)			
Base category: No - never			
Yes - sometimes	0.496**	14.970**	0.517**
	(0.22)	(6.32)	(0.21)
Yes - usually/always	0.408**	12.254**	0.435**
	(0.18)	(5.58)	(0.18)

	Model 3 Original Specification	Model 3 Percentile Ranks	Model 3 OECD income
Regular term-time bedtime (S3)			
Base category: No - never			
Yes - sometimes	-0.126	-3.225	-0.158
	(0.20)	(5.41)	(0.20)
Yes - usually/always	0.099	3.651	0.078
	(0.13)	(3.55)	(0.13)
Child has a LS illness (S3)	-0.051	-0.785	-0.059
0=No, 1= Yes	(0.07)	(1.63)	(0.07)
Cohort Birth Weight (kg)	-0.068	-1.81	-0.067
Range: 0.57 to 6.55	(0.06)	(1.43)	(0.06)
Cohort member female	-0.123**	-3.082**	-0.119**
0=boy, 1=girl	(0.06)	(1.49)	(0.06)
constant	0.05	35.216**	-0.123
	(0.66)	(17.23)	(0.67)
obs	1359	1359	1359
R-sqr	0.310	0.306	0.313
F-test	<0.001***	<0.001***	<0.001***

7.5 Diagnostic tests: age 5 regression

The residuals of model 3 were checked for normality using the normal range of plots (Figure G, Figure H and Figure I) and statistical tests. These showed little evidence of non-normality. A scatter plot of residuals against fitted values showed no evidence of heteroscedasticity and few outliers (Figure J). The estimation was rerun without these 2 observations, however as there was no *a priori* reason to exclude these outliers, these results have not been reported.

7.6 Residual plots: age 5 regression

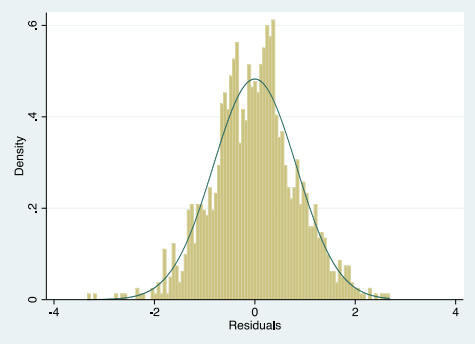

Figure G Histogram of residuals: model 3 age 5 regression

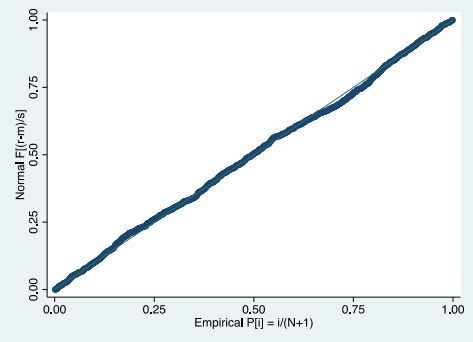

Figure H Normal probability plot of residuals: model 3 age 5 regression

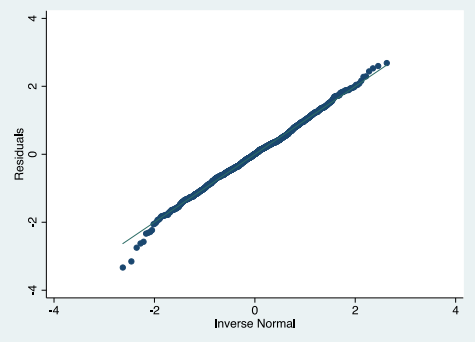

Figure I Qnorm plot of residuals: model 3 age 5 regression

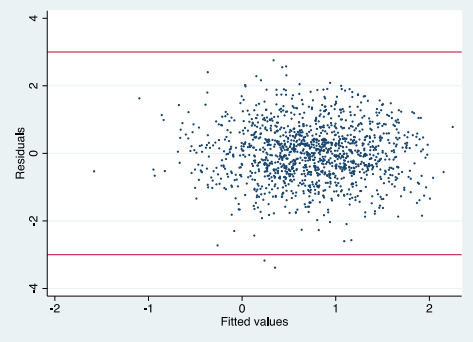

Figure J Scatterplot of residuals against fitted values: model 3 age 5 regression

7.7 Descriptive statistics results: age 7 regression

Table C Grouped mean BAS Vocabulary z-score: age 7

	Mean	Linearised s.e	95% CI Lower	95% CI Upper	F-Test equivelance of means
Experience of Poverty (S1-S4)					
Never	1.034	0.051	0.932	1.136	<0.001***
Intermittent	0.457	0.056	0.344	0.569	
Persistent	0.158	0.125	-0.092	0.407	
S3 BAS Naming Vocab z-score					
Lowest quintile	-0.156	0.149	-0.453	0.141	<0.001***
2nd quintile	0.354	0.099	0.157	0.551	
3rd quintile	0.387	0.087	0.213	0.561	
4th quintile	0.706	0.080	0.546	0.867	
5th quintile	1.104	0.056	0.993	1.216	
Cohort Age at assessment (months) S4					
Lowest quintile	0.802	0.085	0.632	0.973	0.012**
2nd quintile	0.889	0.085	0.719	1.059	
3rd quintile	0.786	0.066	0.654	0.918	
4th quintile	0.780	0.093	0.595	0.966	
5th quintile	0.505	0.086	0.332	0.677	
Highest Household Qualification (S1)					
None	0.144	0.192	-0.240	0.528	<0.001***
NVQ level 1	0.335	0.236	-0.136	0.807	
NVQ level 2	0.452	0.075	0.302	0.603	
NVQ level 3	0.663	0.065	0.534	0.793	
NVQ level 4	1.008	0.061	0.885	1.131	
NVQ level 5	1.263	0.141	0.982	1.545	
Overseas	0.064	0.205	-0.346	0.473	
In lowest SIMD quintile (S1)					
No	0.838	0.047	0.744	0.933	<0.001***
Yes	0.372	0.098	0.175	0.568	
Child Non-cognitive Skills (SDQ) (S3)					
Lowest quintile	1.051	0.073	0.904	1.197	<0.001***
2nd quintile	0.949	0.076	0.796	1.101	
3rd quintile	0.813	0.085	0.643	0.983	
4th quintile	0.651	0.085	0.481	0.821	
5th quintile	0.153	0.076	0.000	0.305	
At least one parent BME (S1)					
No	0.750	0.045	0.659	0.840	0.196
Yes	0.959	0.162	0.635	1.284	

	Mean	Linearised s.e	95% CI Lower	95% CI Upper	F-Test equivelance of means
Housing Tenure (S1)					
Home owner	0.947	0.046	0.856	1.039	<0.001***
Rent(social)	0.279	0.066	0.147	0.412	
Rent(private)	0.671	0.344	-0.017	1.359	
Other	0.687	0.153	0.382	0.992	
Overcrowded home (S1)					
No overcrowding	0.794	0.048	0.699	0.889	0.005***
Overcrowding	0.403	0.124	0.154	0.652	
Language spoken at home (S1)					
English only in the home	0.753	0.045	0.662	0.844	0.386
Another language spoken in the home	0.934	0.201	0.531	1.337	
Marital Status (S1)					
Married	0.932	0.047	0.838	1.026	<0.001***
Cohabiting	0.558	0.069	0.420	0.697	
Single parent	0.371	0.088	0.195	0.546	
Child/Parent Relationship Scale (S2)					
Lowest quintile	0.541	0.075	0.390	0.691	<0.001***
2nd quintile	0.529	0.080	0.369	0.689	
3rd quintile	0.851	0.094	0.663	1.039	
4th quintile	0.970	0.077	0.816	1.125	
5th quintile	0.960	0.105	0.751	1.169	
Mother depression score (S1)					
Lowest quintile	0.811	0.063	0.686	0.936	0.004***
2nd quintile	0.737	0.069	0.598	0.876	
3rd quintile	0.938	0.077	0.783	1.093	
4th quintile	0.589	0.116	0.357	0.821	
5th quintile	0.524	0.099	0.326	0.722	
Mother's age at birth					
Less than 20	0.171	0.159	-0.147	0.489	<0.001***
20-29	0.629	0.059	0.511	0.747	
30-39	0.935	0.051	0.832	1.038	
40 and over	0.626	0.179	0.268	0.985	
Mother experience of long term illness (S1-S4)					
Never	0.833	0.055	0.722	0.944	0.039**
Intermittent	0.671	0.069	0.532	0.809	
Persistent	0.555	0.115	0.325	0.785	

	Mean	Linearised s.e	95% CI Lower	95% CI Upper	F-Test equivelance of means
Child is read to (S2)					
Not at all	-0.435	0.678	-1.792	0.922	0.029**
Less often	0.488	0.312	-0.137	1.112	
Once or twice a month	0.490	0.229	0.032	0.948	
Once or twice a week	0.575	0.091	0.393	0.757	
Several times a week	0.587	0.084	0.418	0.756	
Every day	0.859	0.059	0.741	0.978	
Take to the library (S2)					
No	0.589	0.061	0.467	0.710	<0.001***
Yes	0.938	0.054	0.830	1.046	
Parent teaching score (S2)					
0 - Very low	-0.474	0.113	-0.699	-0.249	<0.001***
1 - Low	0.499	0.207	0.085	0.913	
2 - Medium	0.762	0.068	0.626	0.899	
3 - High	0.768	0.050	0.668	0.869	
Regular meal times (S3)					
No - never	-0.010	0.238	-0.486	0.465	0.007***
Yes - sometimes	0.599	0.211	0.176	1.022	
Yes - usually/always	0.777	0.045	0.688	0.866	
Regular term-time bedtime (S4)					
No - never	0.804	0.198	0.409	1.199	0.401
Yes - sometimes	0.553	0.173	0.206	0.899	
Yes - usually/always	0.763	0.046	0.671	0.856	
Child has a LS illness (S3)					
No	0.802	0.044	0.713	0.890	0.026**
Yes	0.552	0.108	0.336	0.769	
Cohort Birth Weight (kg)					
Lowest quintile	0.802	0.079	0.645	0.959	0.729
2nd quintile	0.674	0.102	0.471	0.877	
3rd quintile	0.808	0.065	0.679	0.937	
4th quintile	0.774	0.068	0.638	0.910	
5th quintile	0.715	0.093	0.528	0.902	
Cohort member female					
Boy	0.714	0.054	0.607	0.821	0.205
Girl	0.797	0.058	0.681	0.913	

Weighted means, unweighted sample size = 1229
* p<0.1, ** p<0.05, ***p<0.01

7.8 Alternative specification results: age 7 regression

Table D Alternative specification results: age 7

	Model 3 Original Specification	Model 3 Percentile Ranks	Model 3 OECD income
	b/se	b/se	b/se
Experience of Poverty (S1-S4)			
Base category: Never			
Intermittent	-0.248***	-5.628**	
	(0.09)	(2.37)	
Persistent	-0.383**	-9.005*	
	(0.16)	(4.51)	
OECD Single country income quintile (S4)			
Base category: Lowest quintile			
2nd quintile			0.028
			(0.11)
3rd quintile			0.007
			(0.12)
4th quintile			0.122
			(0.14)
5th quintile			0.309**
			(0.13)
S3 BAS Naming Vocab z-score	0.316***		0.316***
Range: -3 - 3	(0.04)		(0.04)
S3 BAS Naming Vocab percentile rank		0.274***	
Range: 0.1-99.9		(0.04)	
Cohort Age at assessment (months) S4	0.103	2.931	0.131
Range: 77.4-96.5	(0.17)	(5.09)	(0.20)
Highest Household Qualification (S1)			
Base category: None			
NVQ level 1			
NVQ level 2	0.155	6.227	0.164
	(0.26)	(7.06)	(0.27)
NVQ level 3	0.07	4.228	0.145
	(0.16)	(4.61)	(0.17)
NVQ level 4	0.136	4.876	0.208
	(0.19)	(5.28)	(0.20)
NVQ level 5	0.209	7.133	0.266
	(0.19)	(5.30)	(0.20)
Overseas	0.303	9.172	0.301
	(0.23)	(6.23)	(0.24)
In lowest SIMD quintile (S1)	-0.003	0.974	0.001
0=no, 1=yes	(0.11)	(3.18)	(0.11)
Child Non-cognitive Skills (SDQ) (S3)	-0.038***	-0.933***	-0.038***
Range: 0-28	(0.01)	(0.20)	(0.01)

	Model 3 Original Specification	Model 3 Percentile Ranks	Model 3 OECD income
	b/se	b/se	b/se
At least one parent BME (S1)	0.199	5.694	0.168
0=no, 1=yes	(0.16)	(4.27)	(0.16)
Housing Tenure (S1)			
Base category: home owner			
Rent(social)	-0.103	-3.375	-0.162*
	(0.09)	(2.53)	(0.08)
Rent(private)	0.151	0.873	0.084
	(0.24)	(5.60)	(0.24)
Other	0.061	0.445	0.055
	(0.15)	(3.88)	(0.16)
Overcrowded home (S1)	0.039	1.642	0.009
0=<1, 1=1+per room	(0.13)	(3.52)	(0.13)
Language spoken at home (S1)	0.136	1.492	0.089
0=Eng. Only 1=Also another lang.	(0.20)	(4.95)	(0.20)
Marital Status (S1)			
Base category: Married			
cohabiting	-0.026	-1.948	-0.04
	(0.07)	(1.70)	(0.07)
single parent	-0.004	-0.441	-0.074
	(0.10)	(2.85)	(0.10)
Child/Parent Relationship Scale (S2)	-0.001	-0.048	-0.001
Range: 36-75	(0.01)	(0.15)	(0.01)
Mother depression score (S1)	0.003	0.192	0
Malaise scale, Range: 0-9	(0.02)	(0.43)	(0.02)
Mother's age at birth	-0.001	-0.004	-0.001
Range: 15 to 49	(0.01)	(0.15)	(0.01)
Mother experience of long term illness (S1-S4)			
Base category: Never			
Intermittent	-0.043	-1.818	-0.04
	(0.07)	(1.73)	(0.07)
Persistent	-0.059	-1.234	-0.049
	(0.10)	(2.77)	(0.10)
Child is read to (S2)	-0.019	-0.659	-0.023
Range: 1=Not at all…6=Every day	(0.04)	(1.01)	(0.04)
Take to the library (S2)	0.133*	3.711*	0.117
0=no, 1=yes	(0.08)	(1.95)	(0.08)
Parent teaching score (S2)	0.109*	2.278*	0.107*
Range: 0-3	(0.06)	(1.34)	(0.06)

	Model 3 Original Specification	Model 3 Percentile Ranks	Model 3 OECD income
	b/se	b/se	b/se
Regular meal times (S3)			
Base category: No - never			
Yes - sometimes	0.25	7.131	0.251
	(0.28)	(7.55)	(0.29)
Yes - usually/always	0.149	7.406	0.178
	(0.22)	(6.09)	(0.24)
Regular term-time bedtime (S4)			
Base category: No - never			
Yes - sometimes	-0.279	-6.609	-0.303
	(0.24)	(6.07)	(0.24)
Yes - usually/always	-0.298	-8.400*	-0.316*
	(0.19)	(4.81)	(0.18)
Child has a LS illness (S4)	-0.071	-3.045	-0.083
0=No, 1= Yes	(0.10)	(2.37)	(0.10)
Cohort Birth Weight (kg)	-0.048	-0.549	-0.049
Range: 0.57 to 6.55	(0.05)	(1.34)	(0.05)
Cohort member female	0.073	3.476**	0.077
0=boy, 1=girl	(0.06)	(1.58)	(0.06)
constant	3.053**	97.882***	2.856**
	(1.17)	(33.24)	(1.15)
obs	1229	1229	1229
R-sqr	0.240	0.236	0.240
F-test	<0.001***	<0.001***	<0.001***

* p<0.1, ** p<0.05, ***p<0.01

7.9 Diagnostic tests: age 7 regression

As before, the residuals of model 3 were checked for normality using the normal range of plots (Figure K, Figure L and Figure M) and statistical tests. These showed little evidence of non-normality. A scatter plot of residuals against fitted values showed no evidence of heteroscedasticity and few outliers (Figure N). The estimation was rerun without these 3 observations which had a minimal impact on the results, however, again, as there was no *a priori* reason to exclude these outliers, these results have not been reported.

7.10 Residual plots: age 7 regression

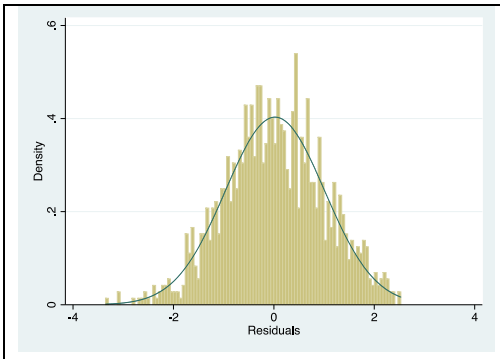

Figure K Histogram of residuals: model 3 age 7 regression

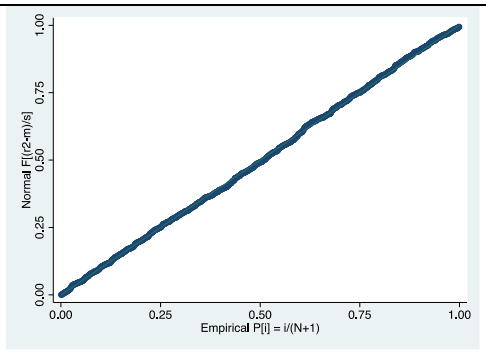

Figure L Normal probability plot of residuals: model 3 age 7 regression

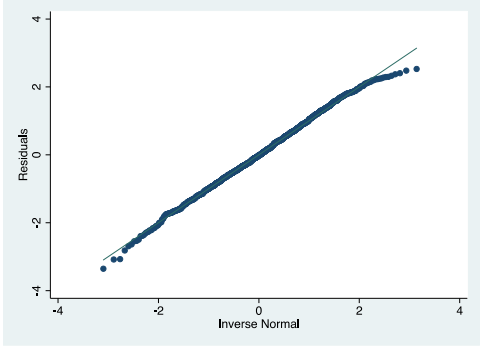

Figure M Qnorm plot of residuals: model 3 age 7 regression

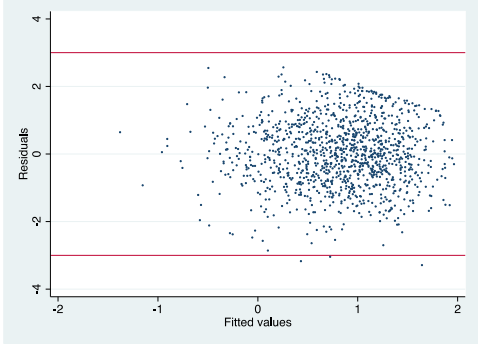

Figure N Scatterplot of residuals against fitted values: model 3 age 7 regression

8 References

Barr, N. A. (2004). *The economics of the welfare state* (4th ed. ed.). Oxford: Oxford : Oxford University Press.

BBC. (2016). Sturgeon announces school reform summit as part of 'bold' education plan. Retrieved 15th August, 2016, from http://www.bbc.co.uk/news/uk-scotland-scotland-politics-36372337

Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics, 4*(3, Part 2), S1-S39.

Blanden, J., Greaves, E., Gregg, P., Macmillan, L., & Sibieta, L. (2015). *Understanding the improved performance of disadvantaged pupils in London*: Centre for Analysis of Social Exclusion, LSE.

Blanden, J., Gregg, P., & Machin, S. (2005). Intergenerational Mobility in Europe and North America. *Report supported by the Sutton Trust, Centre for Economic Performance, London School of Economics*.

Blanden, J., Gregg, P., & Macmillan, L. (2007). Accounting for intergenerational income persistence: noncognitive skills, ability and education. *The Economic Journal, 117*(519), C43-C60.

Blanden, J., & Machin, S. (2010). Intergenerational inequality in early years assessments. In K. Hansen, H. Joshi & S. Dex (Eds.), *Children of the 21st Century–The First Five Years* (pp. 153-168). Great Britain: The Policy Press.

Blanden, J., & Macmillan, L. (2014). Education and intergenerational mobility: Help or hindrance?

Blow, L., Goodman, A., Kaplan, G., Walker, I., & Windmeijer, F. (2005). How Important is Income in Determining Children's Outcomes? A Methodology Review of Econometric Approaches. *Institute for Fiscal Studies*.

Bradley, R. H., Caldwell, B. M., Rock, S. L., Ramey, C. T., Barnard, K. E., Gray, C., et al. (1989). Home environment and cognitive development in the first 3 years of life: A collaborative study involving six sites and three ethnic groups in North America. *Developmental psychology, 25*(2), 217.

Bradshaw, P. (2011). *Growing Up in Scotland: Changes in child cognitive ability in the pre-school years*.

Bradshaw, P., Lewis, G., & Hughes, T. (2014). *Growing Up in Scotland: Characteristics of pre-school provision and their association with child outcomes*. Retrieved from http://www.gov.scot/Resource/0045/00453130.pdf.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. *Handbook of child psychology*.

Brooks-Gunn, J., & Duncan, G. J. (1997). The effects of poverty on children. *The future of children*, 55-71.

Carneiro, P., & Heckman, J. (2003). Human Capital Policy. *IDEAS Working Paper Series from RePEc*.

Connelly, R. (2013). *Millennium Cohort Study Data Note : Interpreting Test Scores*. London: Institute of Education.

de la Rochebrochard, E. (2012). *Millennium Cohort Study: Data Note 1: The home learning environment at Age 3*. London: Institute of Education.

Dearden, L., Sibieta, L., & Sylva, K. (2011). The socio-economic gradient in early child outcomes: evidence from the Millennium Cohort Study. *Longitudinal and Life Course Studies, 2*(1), 19-40.

Denholm, A. (2016). Scottish Government plans to close school attainment gap "seriously unclear" says Royal Society. *Herald Scotland* Retrieved 12th August, 2016, from http://www.heraldscotland.com/news/13412062.Scottish_Government_plans_to_close_school_attainment_gap__seriously_unclear__says_Royal_Society/

Duncan, G. J., & Brooks‐Gunn, J. (2000). Family poverty, welfare reform, and child development. *Child development, 71*(1), 188-196.

Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How Much Does Childhood Poverty Affect the Life Chances of Children? *American Sociological Review, 63*(3), 406-423.

DWP. (2016). *Households below average income (HBAI): An analysis of the income distribution 1994/95 to 2014/15*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/532416/households-below-average-income-1994-1995-2014-2015.pdf.

Ellis, S., & Sosu, E. (2014). Closing the attainment gap in Scottish education. *Joseph Rowntree Foundation, York*.

Feinstein, L. (2003). Inequality in the early cognitive development of British children in the 1970 cohort. *Economica, 70*(277), 73-97.

Feinstein, L. (2004). Mobility in pupils' cognitive attainment during school life. *Oxford Review of Economic Policy, 20*(2), 213-229.

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism, 10*(2), 486-489.

Goodman, A., Gregg, P., & Washbrook, E. (2011). Children's educational attainment and the aspirations, attitudes and behaviours of parents and children through childhood. *Longitudinal and Life Course Studies, 2*(1), 1-18.

Gregg, P., & Machin, S. (1998). *Child disadvantage and success or failure in the youth labour market*. London.

Hansen, K. (2014). *Millennium Cohort Study: A Guide to the Datasets (Eighth Edition) - First, Second, Third, Fourt and Fifth Surveys*. London: Institute of Education.

Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of economic literature, 33*(4), 1829-1878.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). *The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior*: National Bureau of Economic Research.

Jerrim, J., & Vignoles, A. (2013). Social mobility, regression to the mean and the cognitive development of high ability children from disadvantaged homes. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176*(4), 887-906.

Johnson, J. (2012). *Millennium Cohort Study: Psychological, Developmental and Health Inventories. User Guide to the Data*. London: Institute of Education.

Ketende, S. C., & Jones, E. M. (2011). *The Millennium Cohort Study: A User Guide to Analysing MCS using STATA*. London: Institute of Education.

Klebanov, P. K., Brooks-Gunn, J., McCarton, C., & McCormick, M. C. (1998). The contribution of neighborhood and family income to developmental test scores over the first three years of life. *Child development*, 1420-1436.

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. *Psychological bulletin, 126*(2), 309.

Macaulay, J. (1977). Stereotyping child welfare. [journal article]. *Society, 14*(2), 47-51.

Mayer, S. (1997). *What money can't buy: Family income and children's life chances*. Cambridge, MA: Harvard University Press.

McMahon, W. W. (2010). The external benefits of education. *The Economics of Education*, 68-79

Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj‐Blatchford, I., & Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues, 64*(1), 95-114.

OECD. (2015). *Improving schools in Scotland - an OECD perspective*: OECD.

Parcel, T. L., & Menaghan, E. G. (1990). Maternal working conditions and children's verbal facility: Studying the intergenerational transmission of inequality from mothers to young children. *Social Psychology Quarterly*, 132-147.

Parsons, S., Schoon, I., & Vignoles, A. (2013). Parental worklessness and children's early school achievement and progress. *Longitudinal and Life Course Studies, 5*(1), 19-41.

Scottish Government. (2007). Poverty methodology. Retrieved 26 August, 2016, from http://www.gov.scot/Topics/Statistics/Browse/Social-Welfare/IncomePoverty/PovertyMethodology

Scottish Government. (2014). *Child Poverty Measurement Framework - The Wider Evidence Base*. Retrieved from http://www.gov.scot/Resource/0045/00457209.pdf.

Scottish Government. (2015). *Tackling inequalities in the early years: Key Messages from 10 years of the Growing Up in Scotland study*.

Scottish Government. (2016a). *National Improvement Framework for Scottish Education: 2015 Interim Framework Report*. Retrieved from http://www.gov.scot/Resource/0049/00491743.pdf.

Scottish Government. (2016b). *Scottish Survey of Literacy and Numeracy (2015) Numeracy*

Seltzer, J. A. (1994). Consequences of marital dissolution for children. *Annual Review of Sociology*, 235-266.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). *Final report from the primary phase: Pre-school, school and family influences on children's development during Key Stage 2 (7-11)*.

The Royal Society of Edinburgh. (2015). *The Draft National Improvement Framework for Education: a report of a discussion which brought together education researchers and Scottish Government Officials*. Edinburgh: The Royal Society of Edinburgh.

The Scottish Government. (2015). *Creating a smarter Scotland: A draft National Improvement Framework for Scottish Education*.

Tymms, P., Merrell, C., & Buckley, H. (2015). *Children's development at the start of school in Scotland and the progress made during their first school year: An analysis of PIPS baseline and follow-up assessment data*. Retrieved from http://www.gov.scot/Resource/0049/00490869.pdf.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3), 461-481.

9 Useful websites

Growing up in Scotland, http://growingupinscotland.org.uk/ , last accessed August 2016

Millenium Cohort Study, http://www.cls.ioe.ac.uk, last accessed August 2016

The Scottish Government, http://www.gov.scot/, last accessed August 2016

UK Data Service, https://www.ukdataservice.ac.uk/, last accessed July 2016