



The link between childhood reading skills and adult outcomes: analysis of a cohort of British children

IFS Briefing Note BN169

Claire Crawford
Jonathan Cribb

The Link between Childhood Reading Skills and Adult Outcomes: Analysis of a Cohort of British Children¹

Claire Crawford² and Jonathan Cribb³

© The Institute for Fiscal Studies, April 2015
ISBN: 978-1-909463-85-1

Executive summary

This briefing note builds on previous work by Crawford and Cribb (2013) to investigate the link between children's reading skills at age 10 and their outcomes as adults using data from the British Cohort Study (a survey of individuals born in one week of April 1970). We find that reading skills are associated with significant increases in gross hourly wages and gross weekly earnings, particularly at older ages (ages 38 and 42), but less consistent evidence for strong links between reading skills in childhood and other outcomes in adulthood, including the likelihood of being in work, self-reported health status and the intergenerational transmission of reading skills. We also find some suggestive evidence that the link between reading skills in childhood and wages and earnings in adulthood is stronger amongst those from poor backgrounds. Overall, this note provides suggestive evidence that improving reading skills in childhood may be one route through which earnings potential in adulthood could be increased, although it should be noted that these estimates are associations rather than evidence of causality.

1. Introduction

This briefing note investigates the link between children's reading skills and adult outcomes using data from the British Cohort Study. It builds on previous work by Crawford and Cribb (2013)⁴ – who explored the links between reading and

¹ The authors are grateful to Save the Children for funding this research.

² University of Warwick and Institute for Fiscal Studies.

³ Institute for Fiscal Studies.

⁴ C. Crawford and J. Cribb, *Reading and Maths Skills at Age 10 and Earnings in Later Life: A Brief Analysis using the British Cohort Study*, Research Report REPO3, Centre for Analysis of Youth Transitions (CAYT), London, 2013,

mathematics skills at age 10 and labour market outcomes at ages 30, 34 and 38 using the same data – in five key ways: first, we investigate labour market outcomes at older ages (up to age 42); second, we investigate a wider range of adult outcomes, including self-reported health status and the likelihood of reading to one’s own children; third, we make comparisons at different parts of the reading distribution – in particular between those scoring in the bottom 25% versus the middle 50% or top 25%; fourth, we investigate whether the importance of reading skills varies between children from poor and non-poor backgrounds; fifth, we explore the robustness of our results to the use of more flexible regression techniques (specifically the use of a technique known as propensity score matching).

2. Data and methodology

To conduct our analysis, we use data from the British Cohort Study (BCS), which tracks individuals born in a particular week of April 1970 through their lives, up to and including the latest survey in 2012, when the individuals were aged 42. Of particular interest for our study is the fact that BCS cohort members were tested on their reading skills at age 10 using a shortened version of the Edinburgh Reading Test. We calculate the percentage of questions that they answered correctly and split children into three groups on the basis of this total score: the top 25%, middle 50% and bottom 25%. Our analysis shows the advantage of scoring in the middle 50% or the top 25% (relative to the bottom 25%) on a range of outcomes at ages 30, 34, 38 and 42: the likelihood of being in work, gross hourly wages, gross weekly earnings, self-reported health status and the frequency of reading to one’s children. Appendix Table 1 shows the average outcomes amongst individuals in our sample at each age.

We estimate the relationship between reading scores at age 10 and later outcomes using a model of the following basic form:

$$y_i = \alpha + \gamma_1 (\text{reading middle } 50\%)_i + \gamma_2 (\text{reading top } 25\%)_i + \mathbf{X}'_{1i}\boldsymbol{\beta}_1 + \mathbf{X}'_{2i}\boldsymbol{\beta}_2 + \varepsilon_i.$$

In this model, the main coefficients of interest are γ_1 and γ_2 , which can be interpreted as the advantage associated with scoring in the middle 50% versus the bottom 25% in terms of reading skills at age 10 on the outcome in question and as the advantage associated with scoring in the top 25% versus the bottom 25% in terms of reading skills at age 10 on the outcome in question, respectively.

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/190625/Reading_and_maths_skills_at_age_10_and_earnings_in_later_life.pdf.

X'_{1i} and X'_{2i} are different sets of control variables (specified below). For each outcome at each age, we show three sets of results:

- **Raw differences (specification 1).** This specification shows the link between reading skills and the outcome of interest without taking into account any of the other ways in which children with different reading skills may differ from each other. It is equivalent to a model without X'_{1i} and X'_{2i} .
- **Differences accounting for family background (specification 2).** This specification shows the link between reading skills and the outcome of interest amongst children from similar backgrounds (e.g. of the same gender and ethnicity, whose parents have similar education and work histories, and so on, which form part of X'_{1i}). This second specification includes X'_{1i} but not X'_{2i} .
- **Differences accounting for other skills (specification 3).** Reading skills are not the only types of skills that are likely to matter for the outcomes we consider; how well children do in maths and other cognitive tests, as well as their non-cognitive skills (e.g. confidence and self-esteem) may also be important. These skills are included in X'_{2i} . The final specification therefore considers the effect of reading skills over and above these other types of skills, and includes both X'_{1i} and X'_{2i} .

The exact variables that we control for in specifications 2 and 3 are listed in Appendix Table 2.

We summarise the differences in characteristics between individuals with low, medium and high reading skills at age 10 in Table 1. There are large and statistically significant differences between the groups. Compared with children scoring in the bottom 25% in terms of reading skills at age 10, those scoring in the top 75% are significantly less likely to be male, to have been born prematurely and to have had special educational needs, and significantly more likely to have been breastfed and to come from higher socio-economic backgrounds (as measured by the education levels of parents, probability that their parents own their home and various other measures). Individuals scoring in the top 75% in terms of reading skills at age 10 also have better other types of skills as well: for example, they are more likely to have higher maths skills and higher measures of non-cognitive skills (such as higher self-esteem, greater self-perceived ability and better behaviour). Given these differences, it is vital to account for the other ways in which children with different reading skills differ from each other when conducting our analysis.

Table 1. Average characteristics of individuals with different reading scores

Characteristic	Average of each characteristic for those with:			Difference in average between individuals with:		Sample size
	Low reading score	Middle reading score	High reading score	Middle and low scores	High and low scores	
White	0.946	0.977	0.988	0.031 ^{***}	0.042 ^{***}	10,764
Special educational needs	0.030	0.003	0.000	-0.027 ^{***}	-0.030 ^{***}	10,548
Male	0.562	0.505	0.482	-0.057 ^{***}	-0.079 ^{***}	11,646
Parents owned house	0.448	0.622	0.785	0.174 ^{***}	0.337 ^{***}	10,687
Number in bedroom at age 5	1.932	1.699	1.555	-0.233 ^{***}	-0.377 ^{***}	9,500
Age father left education	15.104	15.663	16.688	0.559 ^{***}	1.584 ^{***}	8,095
Age mother left education	15.219	15.639	16.544	0.421 ^{***}	1.326 ^{***}	8,055
Mother's age at birth:						
Under 20	0.121	0.095	0.057	-0.026 ^{***}	-0.064 ^{***}	11,566
20–24	0.376	0.370	0.315	-0.007	-0.062 ^{***}	11,566
25–29	0.275	0.309	0.374	0.033 ^{***}	0.099 ^{***}	11,566
30–34	0.143	0.149	0.167	0.006	0.024 ^{**}	11,566
35+	0.084	0.077	0.086	-0.006	0.003	11,566
Father employed at age 10	0.891	0.940	0.969	0.050 ^{***}	0.078 ^{***}	9,968
Mother employed at age 10	0.506	0.544	0.539	0.038 ^{***}	0.033 ^{**}	10,195
Mother married at birth	0.924	0.944	0.958	0.020 ^{***}	0.035 ^{***}	11,634
Mother very interested in child's education aged 10	0.334	0.548	0.778	0.214 ^{***}	0.444 ^{***}	9,742
Father very interested in child's education aged 10	0.335	0.537	0.765	0.202 ^{***}	0.430 ^{***}	6,726
Height in cm at age 10	137.171	138.801	139.709	1.630 ^{***}	2.539 ^{***}	10,046
Birthweight						
Under 2.5kg	0.087	0.056	0.038	-0.030 ^{***}	-0.048 ^{***}	11,641
2.5–4.5kg	0.901	0.931	0.945	0.030 ^{***}	0.044 ^{***}	11,641
Over 4.5kg	0.012	0.013	0.017	0.001	0.004	11,641
Premature birth	0.073	0.050	0.033	-0.024 ^{***}	-0.040 ^{***}	11,144
Ever breastfed	0.294	0.353	0.462	0.060 ^{***}	0.168 ^{***}	11,625
Mother smoked pre-pregnancy	0.656	0.586	0.506	-0.070 ^{***}	-0.151 ^{***}	11,579
Maths score at age 10	-0.983	0.092	0.943	1.075 ^{***}	1.926 ^{***}	11,609
Locus of control score at 10	-0.547	0.022	0.576	0.569 ^{***}	1.124 ^{***}	11,424
Self-esteem score at 10	-0.257	-0.018	0.349	0.240 ^{***}	0.607 ^{***}	11,467
Self-perceived ability at 10	-0.314	0.028	0.307	0.342 ^{***}	0.621 ^{***}	11,448
Behaviour scale at 10	0.151	-0.020	-0.164	-0.171 ^{***}	-0.316 ^{***}	10,634
Positive activities score at 10	-0.237	0.060	0.201	0.296 ^{***}	0.438 ^{***}	10,717
Cognitive test at 5	-0.399	0.020	0.336	0.419 ^{***}	0.736 ^{***}	9,953
Pattern recognition score at 10	-0.752	0.044	0.769	0.796 ^{***}	1.521 ^{***}	11,483

*** indicates that the differences are significantly different from zero at the 1% level, ** at the 5% level and * at the 10% level.

We also look at whether (conditional on their characteristics) the link between reading skills in childhood and later life outcomes differs for individuals who grew up in poor versus non-poor families. To do this, we compare the results for those children who grew up in the poorest 25% of families (as measured at age 10) with the results for those who grew up in the richest 75% of families. Choosing the poorest 25% of children corresponds to the relative child poverty rate (25.6%) based on a poverty line of 70% of median income in 1980 (the year that the individuals in our data turned 10).⁵

For our results to be interpreted as the causal effect of reading skills on adult outcomes, we have to assume that cohort members with higher or lower test scores do not differ in unobserved ways that also affect their outcomes. This is a strong assumption to make; thus we interpret our results, even from specification 3, as evidence of strong associations rather than causal effects.

We also conduct some robustness checks (referred to where relevant throughout the note) to assess how sensitive these results are to making different assumptions about sample selection or method.⁶

3. Results

Table 2 shows the results for the three labour market outcomes: the likelihood of being in work, gross hourly wages and gross weekly earnings at ages 30, 34, 38 and 42. The top panel shows the link between reading skills at age 10 and the likelihood of being in work (relative to being unemployed or out of the labour market) at these ages. The middle panel shows the link between reading skills and gross hourly wages, and the bottom panel the link between reading skills and gross weekly earnings.⁷

⁵ The proportion of children living in poverty on the basis of the standard 60% of median income line was 16.6% in our sample. Unfortunately, this is too small a group for us to be likely to be able to identify significantly different results for poor and non-poor children. For more details, see the IFS poverty and inequality spreadsheet, http://www.ifs.org.uk/tools_and_resources/incomes_in_uk.

⁶ We also run our analysis using a technique known as propensity score matching. This reduces in two ways the risk that we mistakenly compare individuals who look very different from each other: first, by allowing us to account for characteristics more flexibly; second, by dropping individuals scoring in the bottom 25% in the reading test who look so different from those scoring in the middle 50% and top 25% that we cannot find an appropriate match. The appendix describes the intuition and assumptions underlying propensity score matching in more detail.

⁷ The wages and earnings outcomes are logged, which means that the coefficient estimates can be interpreted as the percentage change in wages or earnings associated with moving from the bottom 25% in terms of reading skills at age 10 to the middle 50% or the top 25% in terms of reading skills.

Table 2. Link between childhood reading skills and labour market outcomes

<i>Dependent variable: Specification:</i>	In work at age 30			In work at age 34			In work at age 38			In work at age 42		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Reading score in middle 50%	0.093*** (0.011)	0.074*** (0.011)	0.034*** (0.012)	0.066*** (0.012)	0.056*** (0.012)	0.031** (0.013)	0.054*** (0.012)	0.046*** (0.012)	0.014 (0.013)	0.084*** (0.011)	0.065*** (0.011)	0.030** (0.013)
Reading score in top 25%	0.137*** (0.012)	0.100*** (0.013)	0.029* (0.016)	0.090*** (0.013)	0.071*** (0.013)	0.025 (0.017)	0.064*** (0.013)	0.049*** (0.014)	-0.010 (0.017)	0.091*** (0.012)	0.055*** (0.013)	-0.009 (0.017)
Observations	8,353	8,353	8,353	7,212	7,212	7,212	6,611	6,611	6,611	7,239	7,239	7,239

<i>Dependent variable: Specification:</i>	Log hourly wage at age 30			Log hourly wage at age 34			Log hourly wage at age 38			Log hourly wage at age 42		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Reading score in middle 50%	0.148*** (0.015)	0.119*** (0.014)	0.003 (0.016)	0.167*** (0.018)	0.125*** (0.017)	0.008 (0.019)	0.179*** (0.019)	0.147*** (0.019)	0.036* (0.021)	0.207*** (0.019)	0.163*** (0.018)	0.033* (0.020)
Reading score in top 25%	0.339*** (0.016)	0.255*** (0.017)	0.040* (0.022)	0.375*** (0.020)	0.260*** (0.021)	0.042 (0.026)	0.426*** (0.022)	0.330*** (0.022)	0.118*** (0.028)	0.446*** (0.022)	0.334*** (0.022)	0.091*** (0.028)
Observations	6,046	6,046	6,046	5,171	5,171	5,171	4,455	4,455	4,455	4,953	4,953	4,953

<i>Dependent variable: Specification:</i>	Log weekly earnings at 30			Log weekly earnings at 34			Log weekly earnings at 38			Log weekly earnings at 42		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Reading score in middle 50%	0.139*** (0.022)	0.130*** (0.020)	-0.016 (0.022)	0.177*** (0.026)	0.159*** (0.023)	0.016 (0.025)	0.180*** (0.029)	0.169*** (0.025)	0.043 (0.028)	0.245*** (0.028)	0.202*** (0.024)	0.047* (0.028)
Reading score in top 25%	0.391*** (0.024)	0.315*** (0.023)	0.046 (0.029)	0.422*** (0.029)	0.329*** (0.027)	0.064* (0.035)	0.456*** (0.033)	0.377*** (0.030)	0.139*** (0.037)	0.498*** (0.033)	0.378*** (0.031)	0.092** (0.040)
Observations	6,046	6,046	6,046	5,171	5,171	5,171	4,455	4,455	4,455	4,988	4,988	4,988

*** indicates that the estimates are significantly different from zero at the 1% level, ** at the 5% level and * at the 10% level. (1) shows the raw difference, (2) controls for family background and (3) controls for family background and other skills.

The first column in the top panel shows that, relative to those scoring in the bottom 25% in terms of reading skills at age 10, those scoring in the middle 50% are, on average, 9.3 percentage points more likely to be in work at age 30 and those scoring in the top 25% are, on average, 13.7 percentage points more likely to be in work at this age. Accounting for the fact that children with different reading skills may come from different family backgrounds (in column 2) reduces the coefficient estimates somewhat, but they remain large and significant. Once we account for differences in other types of skills (in column 3), however, the estimated importance of reading skills falls sharply – although good reading skills remain significantly positively associated with the likelihood of being in work at age 30 in this model. These results suggest that those scoring in the middle 50% of reading skills at age 10 are, on average, 3.4 percentage points more likely to be in work at age 30 than those scoring in the bottom 25%, and those scoring in the top 25% are, on average, 2.9 percentage points more likely to be in work at age 30 than those scoring in the bottom 25%.

This pattern – of large raw differences that are substantially reduced after accounting for differences in family background and other skills – is reasonably consistent across ages. The remaining differences are, however, relatively small, given that the average employment rate across individuals in this cohort is over 80% at the ages we consider (see Appendix Table 1). It is also worth noting that reading skills are not significantly associated with the likelihood of being in work in our preferred specification (specification 3) at all ages, and indeed the pattern changes depending on whether we impose a common sample (see Appendix Table 3) and the type of model we use.⁸ There is thus little consistent evidence that higher reading skills at age 10 confer a significant advantage in terms of the likelihood of being in work in adulthood (at least not at the ages we consider).

The middle and bottom panels of Table 2 focus on the link between reading skills at age 10 and gross hourly wages and gross weekly earnings respectively. Hourly wages provide an indication of an individual's productivity: their value to an employer per hour worked. Weekly earnings combine this measure of an individual's productivity with their choice of how many hours to work per week. A high-wage individual working part-time and a low-wage individual working full-time may therefore end up with similar weekly earnings, but of course they will have very different hourly wages.

⁸ These results are available from the authors on request.

Despite their potentially different interpretations, however, the differences in hourly wages and weekly earnings by reading skills at age 10 are broadly similar. The raw differences in wages and earnings are large: we estimate that those scoring in the middle 50% in terms of reading skills at age 10 earn, on average, 15% more per hour and 14% more per week at age 30 than those in the bottom 25% in terms of reading skills at age 10. For those with the best reading skills, the differences are even starker: they earn, on average, 34% more per hour and 39% more per week than those with the weakest reading skills at age 30. Moreover, these differences tend to increase with age.

As was the case when examining differences in employment status, however, the inclusion of controls for family background reduces these estimates somewhat, with further substantial reductions occurring once we account for the other types of skills that individuals with different reading skills possess. Indeed, once we compare children from the same types of families with similar levels of other skills, the differences in wages and earnings at ages 30 and 34 are relatively small and not always significantly different from zero. There is some indication that this may change with age, however: at age 38, for example, it appears that individuals scoring in the top 25% in terms of reading skills at age 10 earn, on average, 12% more per hour and 14% more per week than those scoring in the bottom 25% in terms of reading at age 10. These differences are slightly smaller, but remain sizeable and significant, at age 42.

The fact that the raw differences increase with age, together with the fact that the gaps after accounting for other factors remain significant at later ages only, may perhaps suggest that individuals with strong reading skills experience steeper age-earnings profiles (i.e. that their wages increase more rapidly with experience). The same patterns are also found when we estimate the differences in wages and earnings for those with different childhood reading skills using a more flexible modelling technique (propensity score matching),⁹ thus lending greater credibility to these estimates.

Table 3 explores the link between early reading skills and some non-labour-market outcomes in adulthood. In particular, it shows whether an individual reports that they are in good, very good or excellent health (relative to fair or poor health, and labelled 'good health' in the table) at ages 30, 34, 38 and 42. It additionally shows whether parents who were good readers themselves as children are more likely to read with their own children more frequently as adults. This information is asked

⁹ These results are available from the authors on request.

Table 3. Link between childhood reading skills and non-labour-market outcomes

<i>Dependent variable: Specification:</i>	In good health at age 30			In good health at age 34			In good health at age 38			In good health at age 42		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Reading score in middle 50%	0.068 ^{***} (0.011)	0.054 ^{***} (0.011)	0.027 ^{**} (0.012)	0.068 ^{***} (0.013)	0.049 ^{***} (0.013)	0.032 ^{**} (0.015)	0.047 ^{***} (0.011)	0.026 ^{**} (0.011)	0.002 (0.012)	0.094 ^{***} (0.012)	0.068 ^{***} (0.012)	0.039 ^{***} (0.013)
Reading score in top 25%	0.102 ^{***} (0.011)	0.075 ^{***} (0.012)	0.025 [*] (0.015)	0.100 ^{***} (0.014)	0.063 ^{***} (0.015)	0.030 (0.019)	0.074 ^{***} (0.012)	0.035 ^{***} (0.012)	-0.014 (0.016)	0.127 ^{***} (0.012)	0.077 ^{***} (0.013)	0.020 (0.017)
Observations	8,341	8,341	8,341	7,209	7,209	7,209	6,619	6,619	6,619	7,250	7,250	7,250

<i>Dependent variable: Specification:</i>	Reads to child daily at 34			Reads to child weekly at 34		
	(1)	(2)	(3)	(1)	(2)	(3)
Reading score in middle 50%	0.067 ^{**} (0.030)	0.028 (0.030)	0.044 (0.034)	0.041 [*] (0.023)	0.016 (0.023)	0.011 (0.027)
Reading score in top 25%	0.131 ^{***} (0.034)	0.058 (0.035)	0.081 [*] (0.045)	0.089 ^{***} (0.024)	0.039 (0.025)	0.030 (0.035)
Observations	1,764	1,764	1,764	1,764	1,764	1,764

*** indicates that the estimates are significantly different from zero at the 1% level, ** at the 5% level and * at the 10% level. 'Good health' indicates that the individual reports they are in good, very good or excellent health. (1) shows the raw difference, (2) controls for family background and (3) controls for family background and other skills.

for once when the cohort member is aged 34 and only applies to individuals with children between ages 0 and 10 at this point in time.

The top panel of Table 3 focuses on self-reported health outcomes. The first column shows that those scoring in the top 25% in terms of reading skills at age 10 are, on average, 10.2 percentage points more likely to report that they have (at least) good health at age 30 than those scoring in the bottom 25% in terms of reading skills at age 10. The difference between the middle 50% and the bottom 25% is 6.8 percentage points, on average. As was the case for labour market outcomes, however, these differences are reduced slightly once we account for family background characteristics and substantially once we account for other skills that children with different reading abilities possess.

While there remain some significant differences in self-reported health status at some ages, the patterns are not very consistent across model specifications. There is thus little consistent evidence that reading skills in childhood are strongly predictive of self-reported health status in adulthood. This conclusion is not altered when we use a (common) sample of individuals whose self-reported health is observed at all ages (see Appendix Table 4).

The bottom panel of Table 3 shows the relationship between individuals' reading skills as children and how likely they are to read to their own children as adults. The first three columns consider the link between reading skills and the likelihood of reading to one's child daily, the second three columns the likelihood of reading to one's child at least once a week. Since the probability of reading to a child varies substantially depending on the age of the child, we control for the age of the child in all specifications.

The raw results suggest that reading skills are significantly associated with both outcomes, with those scoring in the middle 50% in terms of reading skills at age 10 being, on average, 6.7 percentage points more likely to read to their children daily compared with those scoring in the bottom 25% in terms of reading skills, while those scoring in the top 25% are, on average, 13.1 percentage points more likely to do so. These differences are reduced, and no longer statistically distinguishable from zero at conventional levels (the 5% level), once we account for other differences between individuals in terms of their family background and the other skills they possess. On the basis of these results, we therefore find little concrete evidence for a strong intergenerational transmission of reading behaviour.

Our preferred results (those in column 3 of each set of results for each outcome at each age) account for a wide range of other ways in which children with varying

reading skills differ from each other. Nonetheless, it should be noted that where these results suggest that reading skills remain significantly associated with the outcome of interest even after accounting for these factors, this *does not prove* that changing children's reading skills *will lead to* them working more, earning more, perceiving themselves to be in better health or reading to their own children more frequently in adulthood. This is because we cannot be sure that we have fully captured all of the important ways in which these individuals differ from each other. These results should therefore be taken only as suggestive evidence of important links between reading skills and later outcomes.

Are good reading skills more important for children from poor backgrounds?

We also undertake analysis showing how the association between reading skills in childhood and outcomes in adulthood differs between poor children (those from families in the bottom 25% of the distribution of household income at age 10) and those from better-off families. We only present results controlling for both family background and the different skills that individuals have (i.e. using the third specification). We test not only whether the coefficients are themselves significantly different from zero, but also whether the effects of scoring in the middle 50% or top 25% (relative to the bottom 25%) differ for people who came from poor versus non-poor backgrounds. The results for labour market outcomes are shown in Table 4 and the results for non-labour-market outcomes are shown in Table 5.

We find little evidence of any significant differences between those from poor and non-poor backgrounds in terms of the relationship between reading skills and the likelihood of being in work or the likelihood of reading to one's own children in adulthood. This is perhaps unsurprising given the lack of consistent evidence that we found for these outcomes overall (as described above).

We do, however, find some evidence of significant differences in terms of wages, earnings and self-reported health status, with those from poor backgrounds generally benefiting more from improvements in childhood reading skills than those from better-off families. For example, the middle panel of Table 4 shows that, amongst children growing up in poor families, those scoring in the top 25% in terms of reading skills at age 10 have wages at age 30 that are, on average, 12.6% higher than those scoring in the bottom 25% in terms of reading skills at age 10, while the difference for children from non-poor backgrounds with high versus low reading scores is only 2.3%, on average. These effects are significantly different from each other at the 10% level. This pattern – of greater rewards from higher reading skills amongst children growing up in poorer families – is repeated at all

Table 4. Link between childhood reading skills and labour market outcomes, split by whether the individual grew up in a poor household

<i>Dependent variable:</i> <i>Group:</i>	In work at age 30		In work at age 34		In work at age 38		In work at age 42	
	Poor	Not poor	Poor	Not poor	Poor	Not poor	Poor	Not poor
Reading score in middle 50%	0.022 (0.028)	0.023 (0.015)	0.034 (0.028)	0.017 (0.016)	0.025 (0.030)	0.001 (0.015)	0.058** (0.029)	0.011 (0.015)
Reading score in top 25%	0.012 (0.040)	0.020 (0.019)	0.031 (0.038)	0.006 (0.020)	-0.022 (0.041)	-0.017 (0.020)	0.028 (0.041)	-0.023 (0.019)
Observations	1,730	5,807	1,415	5,117	1,226	4,768	1,475	5,071

<i>Dependent variable:</i> <i>Group:</i>	Log hourly wage at age 30		Log hourly wage at age 34		Log hourly wage at age 38		Log hourly wage at age 42	
	Poor	Not poor	Poor	Not poor	Poor	Not poor	Poor	Not poor
Reading score in middle 50%	-0.013 (0.035)	0.007 (0.019)	-0.001 (0.042)	0.011 (0.023)	0.045 (0.046)	0.036 (0.026)	0.041 (0.041)	0.028 (0.025)
Reading score in top 25%	0.126** (0.055)	0.023 † (0.026)	0.128** (0.063)	0.036 (0.031)	0.206*** (0.068)	0.101*** (0.034)	0.179*** (0.069)	0.071** (0.034)
Observations	1,124	4,367	980	3,732	782	3,289	954	3,560

<i>Dependent variable:</i> <i>Group:</i>	Log weekly earnings at 30		Log weekly earnings at 34		Log weekly earnings at 38		Log weekly earnings at 42	
	Poor	Not poor	Poor	Not poor	Poor	Not poor	Poor	Not poor
Reading score in middle 50%	-0.050 (0.048)	-0.002 (0.027)	-0.003 (0.051)	0.025 (0.031)	-0.028 (0.059)	0.058* (0.034)	0.065 (0.062)	0.034 (0.033)
Reading score in top 25%	0.123* (0.074)	0.037 (0.035)	0.221*** (0.078)	0.057 † (0.042)	0.155* (0.089)	0.141*** (0.045)	0.155 (0.099)	0.072 (0.046)
Observations	1,124	4,367	980	3,732	782	3,289	962	3,584

*** indicates that the estimates are significantly different from zero at the 1% level, ** at the 5% level and * at the 10% level. ††† indicates that the estimated coefficients for poor and not poor individuals are significantly different from each other at the 1% level, †† at the 5% level and † at the 10% level. All columns use specification 3.

Table 5. Link between childhood reading skills and non-labour-market outcomes, split by whether the individual grew up in a poor household

<i>Dependent variable:</i> <i>Group:</i>	In good health at age 30			In good health at age 34		In good health at age 38		In good health at age 42		
	Poor	Not poor		Poor	Not poor	Poor	Not poor	Poor	Not poor	
Reading score in middle 50%	0.090 ^{***} (0.026)	0.011 (0.014)	†††	0.060 [*] (0.032)	0.027 (0.018)	0.012 (0.031)	-0.008 (0.014)	0.096 ^{***} (0.030)	0.012 (0.015)	††
Reading score in top 25%	0.110 ^{***} (0.037)	0.011 (0.018)	††	0.025 (0.048)	0.030 (0.023)	0.001 (0.044)	-0.026 (0.018)	0.096 ^{**} (0.042)	-0.004 (0.019)	††
Observations	1,725	5,801		1,414	5,116	1,228	4,773	1,480	5,078	

<i>Dependent variable:</i> <i>Group:</i>	Reads to child daily at 34		Reads to child weekly at 34	
	Poor	Not poor	Poor	Not poor
Reading score in middle 50%	0.151 [*] (0.083)	0.015 (0.045)	-0.041 (0.069)	0.020 (0.034)
Reading score in top 25%	0.028 (0.123)	0.084 (0.059)	-0.012 (0.098)	0.046 (0.042)
Observations	333	1,270	333	1,270

*** indicates that the estimates are significantly different from zero at the 1% level, ** at the 5% level and * at the 10% level. ††† indicates that the estimated coefficients for poor and not poor individuals are significantly different from each other at the 1% level, †† at the 5% level and † at the 10% level.

other ages, but the estimates are not significantly different from each other. A similar pattern can be seen in terms of weekly earnings.

The top panel of Table 5 also provides some evidence of greater benefits from higher reading skills amongst children growing up in poor families in terms of self-reported health status in adulthood, at least at some ages. For example, it shows that, amongst those from poor backgrounds, scoring in the top 75% in terms of reading scores at age 10 is associated with a significantly higher probability of reporting being in at least good health at ages 30 and 42 compared with people scoring in the bottom 25% of reading scores at age 10. This is in contrast to people who come from less poor families, for whom there are no significant differences. There are, however, no significant links between reading scores and self-reported health for adults from poor or non-poor backgrounds at ages 34 or 38.

4. Conclusion

This briefing note has examined the association between reading skills at age 10 and a range of outcomes in adulthood. Overall, it has shown that reading skills are associated with significant increases in gross hourly wages and gross weekly earnings, particularly at older ages (ages 38 and 42), but that there is less consistent evidence for strong links between reading skills in childhood and other outcomes in adulthood, including the likelihood of being in work, self-reported health status and the intergenerational transmission of reading skills. It has also found some suggestive evidence that the links between reading skills in childhood and wages and earnings in adulthood are stronger amongst those from poor backgrounds, although it should be noted that the coefficient estimates between those from poor and non-poor backgrounds are not significantly different from each other at all ages.

Overall, this note provides suggestive evidence that improving reading skills in childhood may be one route through which earnings potential in adulthood could be improved. However, while we have made good use of the rich data at our disposal to account for as many of the differences between individuals with different reading skills as possible, we cannot be sure that there are no remaining differences that may also matter for outcomes. Where we find evidence of significant differences, our results should thus be regarded as providing suggestive evidence of strong associations between reading skills and outcomes, rather than guaranteeing that a causal relationship between the two exists.

Appendix

Propensity score matching

Propensity score matching relies on constructing a suitable comparison group on the basis of a wide range of characteristics that are observable to the researcher (i.e. available in the data at their disposal). The key assumptions underlying this approach are as follows. First, it must be assumed that, conditional on all observable characteristics included in the model, the outcomes for the treatment and comparison groups would be identical in the absence of the 'treatment' (in this case, different reading skills); this is known as the conditional independence assumption (CIA). Second, there must be some degree of common support between the characteristics of pupils in the treatment and control groups (i.e. there must be some individuals who score in the middle 50% or top 25% in terms of reading skills at age 10 who 'look' like those who score in the bottom 25% in all other ways); otherwise it will be impossible to find a suitable match for these individuals.

For the CIA to hold, the researcher must be able to observe all of the characteristics that are relevant both for determining whether the individual is in the treatment or comparison group and for determining the outcomes of interest. This means that the availability and selection of characteristics on which to match are crucial to the likelihood of the CIA holding. The larger the number of characteristics that must be included in the model, the harder it becomes to find a perfect match for each individual. One way to get around this problem is to estimate a propensity score, which is a simple way of summarising an individual's characteristics. This means that, rather than finding an exact match for each individual in the treatment group in terms of all of their observable characteristics, similar individuals can be found in terms of this summary propensity score.

The propensity score is simply the predicted probability from a model where the dependent variable is a binary variable equal to 1 if the individual is in the treatment group and to 0 if they are in the comparison group. All characteristics that are thought to predict either the likelihood of treatment or the outcomes of interest should be included in the model.

Once the propensity score has been estimated, individuals in the comparison group are weighted according to how closely matched they are to each individual in the treatment group. There are a number of different approaches to undertaking this weighting process – for example, giving weight only to those individuals in the comparison group that are closest in absolute terms to a particular individual in the treatment group (nearest-neighbour matching), allocating a fixed weight to all

individuals within a certain absolute distance (radius matching) or allocating each individual in the comparison group a weight depending on how close they are to each individual in the treatment group (weighted smoothed matching).

The outcomes of individuals scoring in the bottom 25% of reading skills are then compared with the weighted outcomes of individuals scoring in the middle 50% or top 25% in order to estimate the association between higher reading skills in childhood and later outcomes.

Appendix Table 1. Average outcomes amongst individuals in the British Cohort Study, by age

Age	Proportion of individuals observed:			
	In work	In good health	Reading to their child daily	Reading to their child weekly
30	82.1%	85.1%	-	-
34	83.7%	79.1%	51.5%	84.9%
38	85.4%	89.1%	-	-
42	86.0%	85.4%	-	-

Appendix Tables 2–4 appear on the following pages

Appendix Table 2. Variables controlled for in our analysis

<i>Variables controlled for in specification 2</i>		<i>Variables controlled for in specification 3</i>	
Variable	Description	Variable	Description
Sex	Male/female indicator	All variables in specification 2	See left-hand columns
Ethnicity	White/non-white indicator	Height at age 10	Height in cm
Special educational needs	Indicator of special educational needs	Premature birth	Indicator of gestation of less than 37 weeks
Tenure	Parents did/did not own house during individual's childhood	Birthweight	Indicators for under 2.5kg (baseline), 2.5–4.5kg, over 4.5kg
Number in bedroom	Number of people sleeping in child's bedroom at age 5	Child was breastfed	Indicator of whether mother ever breastfed child
Age father left education	Measured in years	Mother smoked	Indicator of whether mother smoked prior to pregnancy
Age mother left education	Measured in years	Maths score	Standardised Friendly Maths Test score
Mother's age at child's birth	Indicators for: under 20 (baseline), 20–24, 25–29, 30–34, 35 and over	Locus of control	Standardised variable that captures the degree to which the individual feels they are in control of their own life at age 10
Father in employment	Indicator Yes/No when child aged 10	Self-perceived ability	Standardised score measured at age 10
Mother in employment	Indicator Yes/No when child aged 10	Positive activities score	Standardised score measured at age 10
Mother married at birth	Indicator Yes/No	Self-esteem score	Standardised score based on LAWSEQ questions at age 10
Mother's interest in education	Teacher says mother very interested in education when child is 10	Disruptive behaviour score	Standardised score based on 'Conners' scale at age 10
Father's interest in education	Teacher says father very interested in education when child is 10	Cognitive test at age 5	Standardised score based on cognitive tests such as drawing a human figure
		Pattern recognition score at age 10	Standardised score based on part of the British Ability Scale tests

Appendix Table 3. Link between childhood reading skills and labour market outcomes using a common sample across ages

<i>Dependent variable: Specification:</i>	In work at age 30			In work at age 34			In work at age 38			In work at age 42		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Reading score in middle 50%	0.066*** (0.015)	0.057*** (0.014)	0.028* (0.016)	0.056*** (0.015)	0.054*** (0.014)	0.039** (0.015)	0.053*** (0.014)	0.048*** (0.013)	0.024 (0.015)	0.046*** (0.013)	0.041*** (0.013)	0.015 (0.014)
Reading score in top 25%	0.079*** (0.016)	0.056*** (0.016)	0.003 (0.020)	0.071*** (0.016)	0.063*** (0.016)	0.035* (0.020)	0.048*** (0.015)	0.036** (0.015)	-0.007 (0.019)	0.044*** (0.014)	0.032** (0.015)	-0.014 (0.018)
Observations	5,044	5,044	5,044	5,044	5,044	5,044	5,044	5,044	5,044	5,044	5,044	5,044

<i>Dependent variable: Specification:</i>	Log hourly wage at age 30			Log hourly wage at age 34			Log hourly wage at age 38			Log hourly wage at age 42		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Reading score in middle 50%	0.130*** (0.023)	0.115*** (0.023)	-0.003 (0.025)	0.139*** (0.027)	0.113*** (0.026)	-0.001 (0.028)	0.179*** (0.026)	0.161*** (0.025)	0.052* (0.027)	0.179*** (0.027)	0.158*** (0.026)	0.025 (0.028)
Reading score in top 25%	0.304*** (0.025)	0.242*** (0.027)	0.010 (0.032)	0.335*** (0.029)	0.242*** (0.030)	0.013 (0.037)	0.395*** (0.028)	0.313*** (0.028)	0.098*** (0.036)	0.427*** (0.030)	0.338*** (0.031)	0.080** (0.038)
Observations	2,481	2,481	2,481	2,481	2,481	2,481	2,481	2,481	2,481	2,481	2,481	2,481

<i>Dependent variable: Specification:</i>	Log weekly earnings at 30			Log weekly earnings at 34			Log weekly earnings at 38			Log weekly earnings at 42		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Reading score in middle 50%	0.102*** (0.033)	0.103*** (0.031)	-0.030 (0.034)	0.113*** (0.036)	0.118*** (0.032)	-0.011 (0.035)	0.166*** (0.037)	0.181*** (0.033)	0.059 (0.036)	0.173*** (0.036)	0.173*** (0.033)	0.026 (0.037)
Reading score in top 25%	0.347*** (0.035)	0.287*** (0.034)	0.026 (0.043)	0.365*** (0.038)	0.301*** (0.037)	0.047 (0.047)	0.425*** (0.040)	0.371*** (0.038)	0.136*** (0.048)	0.445*** (0.040)	0.372*** (0.039)	0.087* (0.050)
Observations	2,491	2,491	2,491	2,491	2,491	2,491	2,491	2,491	2,491	2,491	2,491	2,491

*** indicates that the estimates are significantly different from zero at the 1% level, ** at the 5% level and * at the 10% level. (1) shows the raw difference, (2) controls for family background and (3) controls for family background and other skills.

Appendix Table 4. Link between childhood reading skills and non-labour market outcomes using a common sample across ages

<i>Dependent variable: Specification:</i>	In good health at age 30			In good health at age 34			In good health at age 38			In good health at age 42		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Reading score in middle 50%	0.053 ^{***} (0.014)	0.041 ^{***} (0.014)	0.018 (0.016)	0.066 ^{***} (0.016)	0.053 ^{***} (0.016)	0.034 [*] (0.018)	0.040 ^{***} (0.012)	0.022 [*] (0.012)	-0.003 (0.014)	0.066 ^{***} (0.014)	0.046 ^{***} (0.014)	0.020 (0.016)
Reading score in top 25%	0.095 ^{***} (0.015)	0.071 ^{***} (0.016)	0.027 (0.019)	0.084 ^{***} (0.017)	0.058 ^{***} (0.018)	0.021 (0.023)	0.057 ^{***} (0.013)	0.023 (0.014)	-0.024 (0.017)	0.094 ^{***} (0.015)	0.056 ^{***} (0.015)	0.007 (0.019)
Observations	5,049	5,049	5,049	5,049	5,049	5,049	5,049	5,049	5,049	5,049	5,049	5,049

*** indicates that the estimates are significantly different from zero at the 1% level, ** at the 5% level and * at the 10% level. 'Good health' indicates that the individual reports they are in good, very good or excellent health. (1) shows the raw difference, (2) controls for family background and (3) controls for family background and other skills.