NatCen Social Research

Leading education and social research
Institute of Education
University of London

Institute for Fiscal Studies

CAYT Centre for Analysis of Youth Transitions

Reading and maths skills at age 10 and earnings in later life: a brief analysis using the British Cohort Study

CAYT Report No.3

Claire Crawford
Jonathan Cribb

The Centre for Analysis of Youth Transitions (CAYT) is an independent research centre with funding from the Department for Education. It is a partnership between leading researchers from the Institute of Education, the Institute for Fiscal Studies, and the National Centre for Social Research.

CAYT Centre for Analysis of Youth Transitions

Reading and maths skills at age 10 and earnings in later life: a brief analysis using the British Cohort Study

Claire Crawford and Jonathan Cribb

Institute for Fiscal Studies and CAYT

March 2013

Introduction

Machin and McNally (2008) used data from the British Cohort Study to investigate the relationship between reading skills at age 10 and weekly earnings at age 30.[1] This short note builds on their research in three important ways. First, we extend the analysis to look at the association of maths as well as reading skills at age 10 with earnings in later life. Second, we investigate whether the relationship between skills and later earnings varies over time, by considering the relationship between age 10 skills and earnings at ages 34 and 38 as well as age 30. Third, we study whether the results differ depending on whether we use hourly wages rather than weekly earnings, given that the association may differ if people with higher reading and maths skills choose to work different numbers of hours per week.

Methodology

To carry out this analysis we used data from the British Cohort Study (BCS), which tracks individuals born in a particular week in April 1970 through their lives. Of particular interest for our study is the fact that BCS cohort members were tested on their reading and maths skills at age 10 using the Edinburgh Reading Test and Friendly Maths Test respectively. We calculate the percentage of questions that they answered correctly in each test and then standardise each

[1] Machin, S. and S. McNally (2008) The literacy hour, *Journal of Public Economics*, Vol. 92, pp. 1441-1462.

score so that the relationships we observe can be more easily compared and are interpreted in terms of standard deviation effects.[2]

We are interested in the labour market outcomes of individuals with different reading and maths scores. The BCS contains information on gross weekly earnings, as well as the number of hours usually worked per week, from which it is possible to calculate a measure of gross hourly wages. These results are all expressed in nominal prices (although we account for the fact that individuals are interviewed at different points within a particular year).

We estimate the relationship between maths and reading scores at age 10 and earnings at later ages using a Mincer (1974) model[3], of the following form:

$$\log(earnings)_i = \alpha + \gamma_1 (maths\ score)_i + \gamma_2 (reading\ score)_i + \boldsymbol{X}_i'\boldsymbol{\beta} + \varepsilon_i \quad (1)$$

In this model, the coefficients of interest are γ_1 and γ_2, which can be interpreted as the percentage increase in gross weekly earnings (or gross hourly wages) associated with a one standard deviation increase in maths or reading test scores respectively. We estimate these effects both by including maths and reading test scores separately and together. We also estimate the model separately at ages 30, 34 and 38, and restrict the analysis to those in work at each age. As such, it does not account for the potential selection of individuals into the labour market.

$\boldsymbol{X}_i'$ is a set of control variables which varies by specification. Individuals with higher reading and maths test scores are also likely to differ in other ways from those with lower scores. We therefore consider four sets of controls in different specifications. Specification 1 follows Machin and McNally (2008) and controls for the cohort member's sex and region (at age 10). Specification 2 controls for sex, region, parent's income and education levels. Specification 3 departs from Machin and McNally's analysis and includes a larger number of other characteristics observed during the individual's childhood.[4] These additional

[2] As a robustness check, we used an alternative standardised test score measure, in which we weighted each correct answer by the proportion of children that got it wrong. For example, if three quarters of the sample answered the question correctly, then a correct answer would be weighted by 0.25. It thus gives greater weight to answering harder questions correctly. Our results were not materially affected by using this alternative test score. Results are available from the authors on request.

[3] Mincer, J. (1974), *Schooling Experience and Earnings*, National Bureau of Economic Research, Cambridge MA.

[4] We include ethnicity, housing tenure during childhood, father's occupational class, father's and mother's employment status at age 10, number of older and younger siblings at age 10, whether they are a twin, whether either parent had a major illness during the individual's childhood, mother's age at the individual's birth, mother's marital status at birth, the individual's birth weight, whether their mother drank during pregnancy, whether their mother smoked during pregnancy, whether they were breastfed and their birth weight.

controls are those that we regard as "exogenous", i.e. that we believe are not themselves affected by the child's performance on reading or maths tests.

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

Finally, Specification 4 adds the individual's highest educational qualification in order to investigate the extent to which reading and maths skills are directly associated with earnings or wages even after accounting for the cohort member's educational attainment, or whether their effects on labour market outcomes are channelled entirely via this route (which Machin and McNally (2008) argue can be seen "as a lower bound (or even an underestimate)" of the true effects).

Results

The results of our analysis are set out in Tables 1 and 2, which focus on the relationship between reading and maths skills on gross weekly earnings and gross hourly wages respectively at ages 30, 34 and 38.

The main results can be summarised as follows:

Reading

- The top left hand panel of Table 1 shows that, controlling for gender and region only, a one standard deviation increase in reading scores at age 10 is associated with earning 15.1% more per week at age 30. After controlling for a richer set of background characteristics (Specification 3), this association falls to 10.1%. This suggests that some, but not all, of the relationship between reading skills and weekly earnings can be explained by the other ways in which cohort members with higher or lower reading test scores differ from one another.
- After additionally controlling for the cohort member's highest educational qualification (Specification 4), the association between reading scores and earnings falls still further. This suggests that a substantial proportion of the association between reading test scores at age 10 and earnings at age 30 can be explained by the fact that reading scores are positively correlated with the cohort member's highest educational qualification. After accounting for this correlation, the remaining direct effect of a one standard deviation increase in age 10 reading scores on earnings at age 30 falls to 5.7%.

- Comparing the top, middle and bottom panels of Table 1, we find no significant and systematic differences in terms of the associations between reading test scores at age 10 and earnings at ages 30, 34 and 38.
- We do, however, find some evidence that models which include reading but not maths skills significantly overestimate the association between reading skills and earnings, because they do not take account of the fact that maths scores are highly predictive of reading scores (and vice versa). Once we include both reading and maths skills in our models (in the last four columns of Table 1), the association between reading scores and earnings falls substantially. For example, after accounting for a rich set of background characteristics, a one standard deviation increase in reading scores is now associated with earning 4.4% higher per week at age 30 (compared to 10.1% higher if we do not include maths skills).
- These results are broadly similar if we focus on gross hourly wages rather than gross weekly earnings (shown in Table 2).

Maths

- The top middle panel of Table 1 highlights that the associations between maths test scores and earnings are higher than they were for reading test scores. For example, a one standard deviation increase in maths test scores at age 10 is associated with earning 13.0% more per week at age 30 (controlling for a rich set of background characteristics in Specification 3), compared with 10.1% for a one standard deviation increase in reading test scores.
- This also holds true after controlling for the cohort member's highest educational qualification (in Specification 4), with maths test scores having a greater direct association with earnings at age 30 than reading test scores. This suggests that maths test scores may be more highly valued by employers than reading test scores (although this is of course a speculative suggestion).
- Interestingly, in contrast to the results for reading, the association between maths test scores and earnings does not fall very much when we additionally control for age 10 reading test scores in our models. For example, after controlling for a rich set of background characteristics (Specification 3), the association between a one standard deviation increase in reading test scores and earnings at age 30 falls from 10.1% if we only include reading test scores to 4.4% if we include maths scores as well, while the same figures for maths test scores are 13.0% and 10.5% respectively.
- As was the case for reading scores, these results do not differ markedly by age, nor whether we focus on gross weekly earnings or gross hourly wages.

Conclusions

The results of our analysis of the relationship between reading and maths scores at age 10 and wages and earnings in later life can be summarised as follows:

- After controlling for a rich set of demographic and family background characteristics (Specification 3) a one standard deviation increase in age 10 reading scores is associated with earning between 4.4% and 4.8% more during your 30s, while a one standard deviation increase in age 10 maths scores is associated with earning between 10.2% and 10.8% higher. These results are all significant at the 1% level.
- If we also control for the cohort member's highest educational qualification, which is likely to be an important route through which reading and maths scores affect later earnings, the association of higher reading scores with earnings during your 30s falls to between 1.9% and 2.5% (significant only at the 10% level), while the effect of higher maths scores is still between 7.1% and 7.5% (significant at the 1% level). This provides some suggestive evidence that employers may value maths skills more than reading skills.
- There is a strong correlation between maths and reading scores, and looking at the effect of one without controlling for the other is likely to overestimate their associations with earnings, particularly in the case of reading scores.
- There are no significant differences between the effect sizes at ages 30, 34 and 38, so returns estimated at a single point in time are likely to be reasonably representative of those at closely surrounding ages.
- Our results are reasonably robust to using hourly wages instead of weekly earnings, suggesting that increasing hours worked is not one of the key routes through which reading and maths skills generate higher returns in the labour market.

Table 1: Percentage increase in gross weekly earnings associated with a 1 standard deviation increase in test scores

Specification	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	Number of observations
Controlling for	Reading scores only				Maths scores only				Reading and maths scores				
Dependent variable: log earnings at age 30													6,029
Std reading test score at age 10	0.151***	0.116***	0.101***	0.057***					0.061***	0.048***	0.044***	0.019*	
	(0.008)	(0.009)	(0.009)	(0.009)					(0.010)	(0.010)	(0.010)	(0.010)	
Std maths test score at age 10					0.182***	0.148***	0.130***	0.084***	0.144***	0.119***	0.105***	0.073***	
					(0.008)	(0.008)	(0.009)	(0.009)	(0.010)	(0.011)	(0.011)	(0.010)	
Dependent Variable: log earnings at age 34													5,158
Std reading test score at age 10	0.164***	0.120***	0.108***	0.064***					0.071***	0.052***	0.048***	0.025**	
	(0.009)	(0.010)	(0.010)	(0.010)					(0.012)	(0.012)	(0.012)	(0.011)	
Std maths test score at age 10					0.194***	0.149***	0.136***	0.089***	0.149***	0.118***	0.108***	0.075***	
					(0.010)	(0.010)	(0.010)	(0.010)	(0.012)	(0.012)	(0.012)	(0.012)	
Dependent Variable: log earnings at age 38													4,441
Std reading test score at age 10	0.155***	0.112***	0.107***	0.059***					0.069***	0.048***	0.048***	0.021*	
	(0.010)	(0.010)	(0.011)	(0.011)					(0.013)	(0.013)	(0.013)	(0.013)	
Std maths test score at age 10					0.181***	0.137***	0.131***	0.083***	0.136***	0.108***	0.102***	0.071***	
					(0.010)	(0.011)	(0.011)	(0.011)	(0.014)	(0.013)	(0.014)	(0.013)	

Notes: *** denotes that the effect is significantly different from zero at the 1% level, ** at the 5% level, * at the 10% level. Standard errors are robust to heteroskedasticity. Specification 1 includes controls for sex and region. Specification 2 adds parents' income and education. Specification 3 adds ethnicity, housing tenure during childhood, father's occupational class, father's and mother's employment status at age 10, number of older and younger siblings at age 10, whether they are a twin, whether either parent had a major illness during the individual's childhood, mother's age at the individual's birth, mother's marital status at birth, the individual's birth weight, whether their mother drank during pregnancy, whether their mother smoked during pregnancy, whether they were breastfed and their birth weight. Specification 4 adds controls for the highest qualification attained.

Table 2: Percentage increase in gross hourly wages associated with a 1 standard deviation increase in test scores

Specification	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	Number of observations
Controlling for	Reading scores only				Maths scores only				Reading and maths scores				
Dependent variable: log earnings at age 30													6,029
Std reading test score at age 10	0.116***	0.089***	0.077***	0.047***					0.042***	0.031***	0.028***	0.011	
	(0.006)	(0.006)	(0.006)	(0.006)					(0.008)	(0.008)	(0.008)	(0.007)	
Std maths test score at age 10					0.145***	0.119***	0.105***	0.074***	0.119***	0.100***	0.089***	0.068***	
					(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)	(0.008)	
Dependent Variable: log earnings at age 34													5,158
Std reading test score at age 10	0.132***	0.096***	0.085***	0.052***					0.050***	0.035***	0.032***	0.015*	
	(0.007)	(0.007)	(0.007)	(0.007)					(0.009)	(0.008)	(0.008)	(0.008)	
Std maths test score at age 10					0.163***	0.126***	0.114***	0.080***	0.131***	0.105***	0.096***	0.071***	
					(0.007)	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	
Dependent Variable: log earnings at age 38													4,441
Std reading test score at age 10	0.141***	0.105***	0.097***	0.058***					0.058***	0.041***	0.040***	0.018**	
	(0.008)	(0.008)	(0.008)	(0.008)					(0.010)	(0.010)	(0.010)	(0.009)	
Std maths test score at age 10					0.168***	0.133***	0.124***	0.084***	0.130***	0.108***	0.101***	0.074***	
					(0.007)	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)	(0.010)	(0.010)	

Notes: *** denotes that the effect is significantly different from zero at the 1% level, ** at the 5% level, * at the 10% level. Standard errors are robust to heteroskedasticity. Specification 1 includes controls for sex and region. Specification 2 adds parents' income and education. Specification 3 adds ethnicity, housing tenure during childhood, father's occupational class, father's and mother's employment status at age 10, number of older and younger siblings at age 10, whether they are a twin, whether either parent had a major illness during the individual's childhood, mother's age at the individual's birth, mother's marital status at birth, the individual's birth weight, whether their mother drank during pregnancy, whether their mother smoked during pregnancy, whether they were breastfed and their birth weight. Specification 4 adds controls for the highest qualification attained.