Following in Your Parents' Footsteps?
Empirical Analysis of Matched Parent-Offspring Test Scores

Sarah Brown, Steven McIntosh and Karl Taylor

Department of Economics
University of Sheffield
9 Mappin Street
Sheffield S1 4DT
Great Britain

Corresponding Author: Steven McIntosh

Abstract: In this paper, we explore whether an intergenerational relationship exists between the reading and mathematics test scores, taken at age 7, of a cohort of individuals born in 1958 and the equivalent test scores of their offspring measured in 1991. Our results suggest that how the parent performs in reading and mathematics during their childhood is positively related to the corresponding test scores of their offspring as measured at a similar age. The results further suggest that the effect of upbringing is mainly responsible for the inter-generational relationship in literacy, while genetic effects seem more relevant with respect to numeracy.

Key Words: Human Capital; Intergenerational Transfers; Literacy; Numeracy.
JEL Classification: J13; J24

Word Count: 6940 + tables

January 2009

I. INTRODUCTION

There is a vast literature exploring the determinants and implications of educational attainment. Recent interest has focused on intergenerational aspects with the emphasis on the link between the educational attainment of parents and their offspring. Such a link between parents' and children's educational attainment is not surprising given the important role that parents play in the decisions regarding the human capital investments of their children.[1]

Establishing the existence of links between parents' and children's education is important for the evaluation of education policy. Typically, the public returns to education are estimated in terms of higher productivity, plus other non-economic benefits to society, across the lifetime of the individual who acquires the education. However, if there is an intergenerational transmission of education, then the additional benefits of higher education levels amongst children potentially have to be considered in any evaluation of education policy. Whether such considerations should be made depends upon the source of the intergenerational correlation.

There are at least three potential explanations for the existence of a positive intergenerational relationship in educational attainment. First, it could be the result of a genetic transmission of ability, such that more able parents have more able children, with the educational attainment of both rising accordingly. If this is the sole cause of the intergenerational attainment correlation, then any higher attainment amongst future generations can be ignored in any evaluation of the effects of raising current education levels, since the genes to be passed on have not been affected in this case.

Other potential causes of the intergenerational correlation, however, do mean that an increase in attainment now will be passed on to the next generation. Two potential routes are via the direct transfer of knowledge, and through income and lifestyle. The first route argues that

[1] For example, using the *German Socio-Economic Panel*, Dustmann (2004), finds that parental background (education and profession) has a strong influence on the secondary school track choice of children and ultimately their educational attainment.

better-educated parents will be in an improved position to assist their children, and having experienced the benefits of education themselves, will be more motivated to support and encourage their children through their own education. The more indirect route occurs through the impact of parents' education on their income, and then the benefits of such higher income for the education of their children. There is a huge literature on the economic returns to education in terms of higher wages,[2] whilst the advantages of a well-off family background for a child's educational outcomes have also been established by much research.[3] A higher income can buy many things that might improve a child's educational attainment, such as private schooling, tutors, books, other supplementary materials, ICT, even a house in a better-off neighbourhood and so access to higher performing schools and a peer group of middle-class friends.

There is a growing literature that is providing evidence on intergenerational links in education. The results of Belzil and Hansen (2003), based on a structural dynamic programming model, illustrate the importance of family background for explaining cross-sectional variation in schooling attainment: family background accounts for 68% of explained cross-sectional variations in schooling attainment, with more than half of this explained by the schooling of the mother and father. Ermisch and Francesconi (2001) analyse the *British Household Panel Survey* and find that parents' educational attainment is strongly associated with that of their children, where educational attainment is measured by formal qualifications such as GCSEs, A levels, vocational qualifications and degrees[4]. Ermisch and Francesconi (2001) find that, relative to a parent with no qualifications, mother's education has a stronger association with the child's educational attainment than the educational attainment of the father. Similar evidence exists for a range of countries. For example, using Finnish examination score data, Hakkinen *et al*. (2003) find that parents' education is one of

---

[2] See Ashenfelter *et al.* (1999) or Card (1999), amongst others, for a review of the methodology and results.
[3] See for example Blanden (2004) and Galindo-Rueda and Vignoles (2005) for recent examples for the UK.
[4] General Certificates of Secondary Education (GCSEs), which are taken at age 16 and are the main school leaving qualification in the UK, replaced Certificates of Secondary Education (CSEs) and O levels in the 1980s. CSEs were the equivalent of GCSEs grades below C and O levels were the equivalent of GCSEs grades A to C. A levels are public examinations taken at age 18, usually studying a set syllabus in one to four subjects over a two-year period. This qualification is the major determinant of eligibility for entry to higher education in the UK.

the strongest explanatory variables for student achievement. Indeed, Chevalier ***et al.*** (2007) show that a positive and statistically significant relationship between parents' and child's education exists in every one of the 17 countries they consider, using data from the *International Adult Literacy Survey*.

A number of papers have tried to distinguish between the possible reasons put forward above for a positive education intergenerational correlation. In particular, research has focussed on whether the effect is due to inherited genes ('nature') or upbringing ('nurture'). Separating the total effect of parents' education into these two components is not a straightforward process. The three approaches that have been most commonly used in the literature are the twins studies, the adoptees studies and the IV studies.

The idea behind the twins approach is to consider the differences in educational attainment between twins, and then relate those differences to differential attainment by their children. Thus the method used is usually fixed effects within twin-pairs. Since the genes passed on should be identical within twin-pairs, at least for identical twins, then the genetic effect will drop out of a fixed effects equation, so that any observed relationship between parents' and child's education can be attributed to nurturing effects. An example of such research is Behrman and Rosenzweig (2002), who collect data using their own questionnaire on a sample of twins derived from the Minnesota Twin Registry. To be included in their analysis, the responding twins must both be married and have a child aged at least 18, resulting in 212 female pairs and 122 male pairs. One issue with such studies is that there is still a non-twin parent, and if individuals assortatively mate, then a well-educated woman would marry a well-educated man, who would then pass his genes on to their children, so that a part of the influence of the woman's education is still acting genetically. Once Behrman and Rosenzweig control for within-twin fixed effects, assortative mating and endogenous earnings, they find no impact of a mother's education on her child's, though a father's education still has a positive and statistically significant effect on a child's education. Using the same data set, however, Antonovics and Goldberger (2005) show that these results are highly sensitive to sample selection and education measurement issues. Essentially, there are numerous problems with studies based on twins, such as small sample sizes, an exacerbation of measurement error issues, and a worry that if two twins have different education levels, then this perhaps did not occur by random, so there may be differences between the twins' unobserved characteristics after all.

The twins approach holds genetic effects constant and focuses on family background effects. The study of adoptees, by contrast, considers natural born and adopted children in the same family, therefore holding family background constant, and so explaining differences in education outcomes between children in the same family in terms of differences in their genetic inheritance. This approach is followed by Bjorklund ***et al.*** (2006), who use an administrative data set containing information on all adoptees born in Sweden between 1962 and 1966, therefore giving them a large sample of adoptees. The data set also contains information on the adoptees' new siblings in their adopted families, and on the adoptees' biological parents. The dual routes through which parental education can influence children can therefore be separately identified using such data, with the biological parents identifying the pure genetic component, and the adopted parents identifying the pure upbringing component. Bjorklund *et al.*'s results show that the impact of paternal education works equally through genes and upbringing, whilst for mother's education, the genetic effect dominates. Other studies of adoptees have been undertaken by Plug (2004), using American data from the Wisconsin Longitudinal Study, and Sacerdote (2002) using the UK *National Child Development Study*, the data also used below in this paper. Plug's results show that parental education effects are reduced but remain significant for adoptees, while Sacerdote finds the effect of parental education is as high for adoptees as for natural children. These results therefore suggest that the 'nurture' effect dominates. The limitation of adoptee studies is that sample sizes are often small. In addition, the methodology assumes that adopted children are placed in new families randomly, and that the adoption occurs at birth (so that no time is spent in upbringing with the natural parent), neither of which are necessarily the case.

The final approach to distinguishing between genetic and family upbringing effects, and the one adopted in this study, is to instrument parental education, to isolate the exogenous part of the variance in parents' education. Such exogenous variation in education will be orthogonal to genes, and so any impact that such education has on children's attainment can be attributed to upbringing alone. The instrument that has been used most often in the literature is changes in school leaving ages, which cause exogenous variation in the amount of education received by parents. Chevalier (2004) and Chevalier *et al.* (2005) for the UK, Oreopoulus *et al.* (2006) for the US and Black *et al.* (2005) for Norway, all follow this approach. Their results, however, are mixed, with Oreopoulus *et al.* finding that instrumenting increases the size of the parental education effects, Chevalier and Chevalier *et al.* find similar results but only for maternal education, while Black *et al.* find that instrumenting reduces the size of the effects that they obtain. There are problems related to using school leaving age as an instrument. First, if the law applies nationally, then changes in the law may be conflated with trend changes in parental attainment, which could be connected to genetic changes. This charge cannot be levied against the Oreopoulus *et al.* and Black *et al.* studies however, since both exploit variation in school leaving ages between states or regions in the US and Norway respectively. In addition, however, the use of such an instrument strategy identifies a local average treatment effect, relevant only to those who are affected by the change in the school leaving laws. The analysis in our paper therefore searches for alternative instruments for parental education.

The other innovation of this paper is that we explore the link between parental education and their children's education by examining the relationship between the academic test scores attained by each generation, rather than years of schooling or qualification attainment. Our empirical study therefore makes an interesting contribution to this area as we explore the link between parents' literacy and numeracy test scores during their childhood (i.e. academic test scores attained by the parents when they were aged 7) and the test scores obtained by their offspring. The UK Government, in common with the governments of many other countries, are pursuing active policies to raise the levels of literacy and numeracy, both amongst children currently in school,[5] and amongst adults.[6] Much research has shown the benefits to individuals of improved literacy and numeracy skills, both in terms of economic outcomes such as higher wages and employment likelihoods,[7] and in terms of non-economic outcomes such as health and family life.[8] However, if it can also be shown that improved literacy and numeracy skills will be passed on to the next generation, this adds another benefit to the list of positive outcomes of such policies.

II. MATCHED PARENT-OFFSPRING TEST SCORES

We analyse the British *National Child Development Study* (*NCDS*), which is a cohort study with a target sample of all children born in Great Britain during a given week – March 3rd to March 9th – in 1958. The *NCDS* follows the cohort of children at ages 7, 11, 16, 23, 33, 42 and 46. The unique feature of this data set in the UK context is that we are able to match the academic test scores of the *NCDS* respondents taken when they were aged 7 with the equivalent academic test scores of their offspring taken in 1991 when the children were aged 5 or above and their parents were aged 33. The children of a random 1 in 3 sub-sample of the full sample of *NCDS* respondents were tested. *All* children of the selected respondents were tested, so in numerous cases, different children in the sample have the same parent. Firstly, we describe academic tests taken by the *NCDS* respondents (i.e. the parents) and, secondly, we describe the academic tests taken by their offspring.

Academic tests taken by the parents (i.e. NCDS respondents)

At the age of 7 (i.e. in 1965), the *NCDS* respondents undertook tests in reading and arithmetic. To be specific, a test of word recognition and comprehension was taken, the Southgate Reading Test

---

[5] For example, the introduction of a 'Literacy Hour' in primary school, evaluated by Machin and McNally (2008).
[6] *Skills for Life*. For information, see DfEE (2001).
[7] See for example, McIntosh and Vignoles (2001) and Murnane *et al.* (1995).
[8] See Bynner and Parsons (1997).

(Southgate, 1962), which had a design structure specifically to identify those readers who were below the average reading standard for their age. A problem based arithmetic test (see Pringle, *et al.*, 1966 for further details) was also undertaken at age 7.

Academic tests taken by the offspring of the NCDS respondents

When the *NCDS* respondents were aged 33 (i.e. in 1991), the children of a random sample of 1 in 3 respondents participated in the Peabody Individual Achievement Tests (PIATs) in maths, reading recognition and comprehension. The PIATs, which measure the academic achievement of children aged 5 and over, are widely used and extensively validated brief assessments of academic achievement with high test-retest reliability (*National Longitudinal Survey of Youth* 1997 User Guide). The PIAT reading and maths tests have also been used to measure the educational development of children in the US (James-Burdumy, 2005) and the UK (Brown and Taylor, 2009).

The tests measure ability in mathematics and oral reading ability and constitute the main focus of our empirical analysis. Children start the test at a point that is appropriate for their age and establish a 'basal' ('ceiling') by achieving a certain number of consecutive correct (incorrect) answers. The maths test comprises multiple choice questions which increase in difficulty: early questions focus on, e.g., recognising numerals progressing to topics such as trigonometry and geometry. The reading recognition test comprises multiple choice questions starting with matching and naming letters and progressing to words. The mean age of the children taking the tests in 1991 (i.e. when the parent is aged 33) is 9, see Table 1. Higher scores in the tests represent higher levels of academic achievement. For the sample of 1,848 children who took the PIAT tests in 1991, we are able to match information on the child with detailed information relating to their parent, the original *NCDS* respondent from 1958.

<Table 1 about here>

III. METHODOLOGY

We explore the relationship between the test scores of the parents in reading and mathematics, y^{parent}, i.e. the test scores of the *NCDS* respondents when they were aged 7, and the test scores of their offspring in reading and mathematics, y^{child}, using a specification as given in equation (1) below:

$$y_i^{child} = b_1 ' X_i^{child} + g_1 ' X_i^{family} + f y_j^{parent} + e_{i1} \quad (1)$$

where *i=1…n* (*j=1…m*) identifies the child (parent). We estimate separate equations for the reading test scores and the maths test scores: for example, we regress the child's test score in maths on the parent's test score in maths.[9] In all estimations, standard errors are corrected to allow for the fact that the observations on the children are clustered within families of *NCDS* respondents, since the disturbance terms are likely to be correlated within families.

In order to ease interpretation, and to provide comparability with previous intergenerational studies, particularly those looking at income, all test scores included in the analysis are standardised. For the scores of the *NCDS* respondents, these are standardised to give a mean of zero and standard deviation of one. In the case of the children in the study, however, who unlike their parents are all of different ages, the test scores are standardised to have a mean of zero and a standard deviation of one ***within their yearly age group***. Thus, the test score of each child measures their score relative to the other children in the sample of the same age.

Our focus lies in ascertaining whether a positive relationship exists between the test scores of the parent and child, i.e. whether $f > 0$, and also the components of any such positive relationship. Three components will be considered. The children of parents with high literacy and numeracy skills could benefit through inherited genes, through higher income and educational attainment of their parents due to the latter's higher skill level, and through other positive upbringing factors related to their parents' high skill level.

For both the reading and the maths tests, three versions of equation 1 are estimated. In the first, the only control variable employed in X^{child} is the gender of the child, with the child's age also implicitly controlled for in the standardisation process described above. This specification therefore essentially estimates the raw (child age and gender specific) intergenerational relationship between parents' and children's test scores. The second specification adds numerous controls in the X^{child} vector, namely whether the child is currently in good health; whether the child has siblings; the frequency with which the child bullies other children; whether the child is a loner; and the number of books the child possesses. Family control variables employed in X^{family} include whether the child currently lives in a single parent household; whether the family owns the house outright or on a mortgage, compared to renting; total household labour income and the highest educational attainment of the parent (i.e. *NCDS* respondent).[10] Comparing the intergenerational coefficient before and after the inclusion of these variables will indicate whether this intergenerational effect occurs through the latter variables (with parents' education and income of particular interest), or whether its effect remains even after controlling for these individual and family characteristics. Summary statistics for the variables employed in equation (1) are shown in Table 1.

The final specification endogenises parents' test scores, y^{parent}, by employing the following specification:

$$y_j^{parent} = b_2 ' Z_j^{parent} + e_{j2} \qquad (2)$$

and then replacing y^{parent} in equation (1) with the parent's predicted test score, $\hat{y}_j^{parent}$ based on estimating equation (2)[11]. The vector Z^{parent} has to contain variables that affect the maths and

[9] We have also explored whether there is an association between the parent's maths (reading) test scores and the child's reading (maths) PIAT test score performance. The estimated relationships are often statistically insignificant and, in instances of statistical significance, the estimated effects are relatively small.

[10] For the formal educational qualifications of the parent, we distinguish between: GCSE; A levels; diploma; nursing/teaching qualification; and degree level education.

[11] The standard errors in this third specification are estimated by bootstrapping, since the predicted values of parents' test scores, estimated from the first stage regression, are used in this specification.

reading scores of the parents but are uncorrelated with unobserved characteristics of the parents and have no influence on the children's test scores (other than through any effect on the parent' test scores). The variables chosen were the age at which the *NCDS* respondent started full-time schooling, and the age at which they started the formal, structured learning of phonics (for reading scores) and 'sums' (for maths scores). It is argued that these variables are determined exogenously by accident of birth location and Local Education Authority (LEA) policy, rather than by endogenous deliberate decisions taken by the respondents' parents. The aim of this third specification is therefore to identify the variation in parents' test scores that is exogenous as far as their children's test scores are concerned. Thus any genetic effects are arguably removed from this specification, and any remaining intergenerational relationship can then be said to be working solely through upbringing, unrelated to formal education and income.

<Table 2 about here>

Table 2 reports the frequency distributions of the three variables to be used as instruments. As can be seen, the age at which the *NCDS* respondents started full-time school and began structured phonics/sums learning does vary across individuals. For school starting age, most of the variation is around an interval between 4½ and 5½ years of age. This reflects variation in local policies regarding school starting age, for example, the term in which the child turns 5, the school year in which the child turns 5, the term after the child turns 5 etc. Evidence to support the assertion that this variable reflects local LEA policy rather than individual family choice comes via the fact that most *NCDS* respondents in the same LEA at age 7 had the same school starting age. Thus, in the population as a whole, individuals have approximately the same chance of starting full-time school before or after their 5^{th} birthday, as shown in Table 2. However, within LEAs there is much more uniformity, with the modal size of the dominant proportion across LEAs being 100%. The population-weighted mean of the size of the dominant proportion across LEAs is 77%. In addition, a χ^2 statistic of 790 suggests a very strong relationship between LEA and school starting age, and that variation in school starting age is not randomly distributed across LEAs.

The lower two panels in Table 2 reveal that there is even more variation across respondents as to the age at which they were first systematically taught phonics and sums.[12] The χ^2 statistic between LEA and the age the respondent was first systematically taught phonics is 924, and that between LEA and the age the respondent was first systematically taught sums is 516, again suggesting that the ages at which *NCSD* respondents began systematically learning these skills is strongly related to the LEA in which they happened to be born.

We are therefore confident that the age at which the individuals in our sample began full-time school and began systematically learning literacy and numeracy skills is not a function of genetics or family background, but simply of the policies of the LEA in which they found themselves as children. By using these variables as instruments, we can therefore isolate the random variation in the parents' childhood test scores, to determine whether this random component is also passed on to their own children.

IV. RESULTS

<Table 3 about here>

Table 3 reports the results from the first stage regression, identifying the extent to which the *NCDS* respondents' test scores are determined by the education starting ages. Two columns of results are reported, for their age 7 reading scores and maths scores respectively. Considering first the reading test scores, the age at which the parents first started to learn phonics has a big influence on their reading test score, with better results recorded by the earlier starters. Not surprisingly, those who did not begin to learn phonics systematically until after their seventh birthday do substantially worse in the age 7 reading test. Only one of the school starting age coefficients is statistically significant, but this is not surprising given that starting ages were grouped in the window from age

[12] It is worth recalling that the *NCDS* respondents were born in 1958, and therefore received their schooling in the days before the UK's National Curriculum standardised learning methods across regions, to a large extent.

4½ to age 5½ in the sample. The statistically significant coefficient shows that those who began full-time schooling between the ages of 4½ and 5 score significantly better than those who began between the ages of 5 and 5½ , by about one-fifth of a standard deviation, in the age 7 reading test.

Similar results are obtained for the age 7 maths test, in the final column. Thus, those respondents who began systematic learning of ‘sums’ earlier, perform significantly better in the age 7 maths test, particularly those who started such learning before their fifth birthday. There is also again a separate independent effect of school starting age, with those who began their full-time schooling any time before their fifth birthday performing significantly better in the age 7 maths test than those who did not start until later.

The first stage regressions also controlled for the gender of the respondents. As an aside, it is perhaps of interest that girls performed better in the age 7 reading tests, and boys better in the age 7 maths test, both effects being around 0.1 of a standard deviation on the test scores and both being statistically significant.

Turning now to the second stage regressions of the inter-generational relationship, recall that three specifications are to be estimated: the first controlling just for the gender of the child, the second adding other child and family controls, and the third instrumenting the parents’ test scores, using the predicted values from the equation in Table 2. Before focussing on the intergenerational coefficients, Table 4 presents the results from the second specification, revealing the association between all of the control variables and the child’s test scores.

<Table 4 about here>

It is clear that relatively few of the included variables share a statistically significant relationship with the children’s test scores. Exceptions include higher test scores, both maths and reading, amongst children categorised as being loners, and amongst those who are given more books. On the edge of statistical significance is the relationship that test scores are lower amongst those children who bully others, in accordance with Brown and Taylor (2008). Other variables included, in particular those measuring the parents’ (i.e. *NCDS* respondents’) income and educational attainment, appear unrelated to their children's age 7 test scores. Parents' test scores are, however, important, and so we now focus on that intergenerational relationship in all 3 specifications described above. The coefficients on the parents' test scores in each equation are reported in Table 5.

<Table 5 about here>

The raw intergenerational test score correlation, controlling only for the gender of child, is 0.253 for reading test scores, as shown in the first row of results (Specification 1). This means that for a one standard deviation increase in parents' age 7 reading scores, there is an associated one-quarter of a standard deviation increase in their children's reading scores, relative to other children of the same age. This relationship is economically and statistically significant. Specification 2 adds a range of control variables for characteristics of the child, as well as parental income and the highest qualification of the *NCDS* respondents. The results show, however, that this has essentially no impact on the inter-generational correlation in reading test scores. Thus, any effect of parents' ability in reading on the reading skills of their children is not being transmitted via these control variables. In particular, the children of high-scoring parents are not scoring well themselves because their parents' ability has led to higher educational achievements or family income. The final specification uses instrumental variables to assess the inter-generational relationship, as described above. Again, the estimated coefficient on the parents' age 7 reading test score variable (in this case, instrumented) is left unchanged[13]. This specification isolates any exogenous variation in parents' reading ability due to the local education policies of where they live, and as such could not be passed on genetically to offspring. These results show, however, that having removed any genetic effect, the inter-generational relationship in reading test scores is as strong as ever, suggesting that the source of the relationship is not a genetic effect.

We note, in passing, that the size of this inter-generational effect, based on standardised test scores, is very similar to the intergenerational coefficient obtained in studies of income

[13] Though due to the higher standard errors in this specification, it is now statistically significant only at the 10% level.

intergenerational mobility (measured in log points) that have been obtained in the literature using the same *NCDS* data set. For example, Dearden *et al.* (1997) find an intergenerational income mobility coefficient of 0.24, while Blanden *et al.* (2007) obtain a figure of 0.205 using *NCDS* data. Therefore income and literacy skills seem to be passed on from one generation to the next within this cohort to a similar extent. This of course does not mean that the intergenerational persistence in income is *caused* by the intergenerational persistence in reading skills. Blanden *et al.* do consider the causes of the income persistence they observe, and conclude that cognitive skills do have a role to play, primarily through the impact that they have on qualification attainment and thence on income.

The final column in Table 5 shows the inter-generational coefficients with respect to maths test scores. The first thing to note, in Specification 1, is that the relationship for maths test scores is much smaller than for reading scores. A one standard deviation increase in parents' age 7 maths scores is associated with a one-tenth of a standard deviation increase in their children's maths score, relative to other children of the same age. The effect is still statistically significant at the 1% significance level, however. When the other control variables are added, in Specification 2, the strength of the inter-generational maths relationship is unaffected, as was the case for the reading scores. However, a significant difference between maths and reading is observed in Specification 3. When the parents' maths test scores are instrumented, the coefficient on this variable is reduced considerably, and is highly statistically insignificant. This result suggests that any exogenous increase in parental maths skills are not passed on to their children, implying that genetic effects are the dominant source of the inter-generational correlation in maths scores.

One weakness in the results presented so far, and in particular with the interpretation of Specification 3 as having stripped out any genetic effects, is that no attempt has been made to control for maths and reading ability of the *NCDS* respondent's partner. If there is assortative mating, then high ability *NCDS* respondents may have children with other high ability individuals. Thus, an exogenous increase in the ability of the first parent could lead to them attracting a more able mate, and the genes of the second parent could be passed on to their children. If the ability of the partner is not controlled for, we therefore may not have successfully removed all genetic effects from Specification 3. Unfortunately, none of the partners of *NCDS* respondents were tested for maths and reading skills as part of the survey, so we cannot control for their test scores. However, the survey does contain information on the years of schooling of the original respondents' partners, which could act at least as a proxy for their ability. Additional questions ask the respondents whether their current partner is the parent of each of their children. Any children for whom this was not the case were dropped from this part of the analysis, so that the sample was reduced to children for whom information on both of their natural parents was available. Years of schooling for the respondents' partner (when also a parent of the child in question) was then added to the list of control variables in Specifications 2 and 3. The results are shown in Table 6, where as before only the coefficients on the first parents' (*NCDS* respondents') test scores are reported.

<Table 6 about here>

The results show that controlling for the education of the 'other' parent does not alter the conclusions reached above. For reading test scores, the estimated inter-generational relationship is still very similar in all three specifications[14], whilst for maths test scores, there is a substantial fall to zero in this relationship when parents' test scores are instrumented to isolate exogenous variation in them. One slight difference between Tables 5 and 6 is in Specification 2. In Table 6, there is a larger, though still not too pronounced, fall in the estimated inter-generational relationship when the control variables are added to the equation, for both reading and maths. Thus controlling for the schooling of the 'other' parent does reduce somewhat the explanatory power of the first parents' test scores, suggesting that assortative mating does play a small part in explaining the inter-generational relationship between one parent's age 7 test scores and those of their children.

V. FINAL COMMENTS

[14] With the IV coefficient in Specification 3 again statistically significant at around the 10% level.

We contribute to the empirical literature on the inter-generational aspects of education by exploring the relationship between the academic test scores of parents and their offspring, rather than focusing on inter-generational links between formal academic qualifications. Furthermore, the academic test scores of the parents are measured when the parent was aged 7, i.e. during the parent's childhood. Our results suggest that how the parent performs in reading and maths tests as a child has a positive influence upon the corresponding reading and maths test scores of their offspring, i.e. the parent's ability in maths and reading as a child is positively associated with their offspring's ability in maths and reading. This relationship is stronger for reading than for maths.

The study goes on to investigate the cause of these inter-generational relationships. With respect to reading, nothing we try reduces the size of the inter-generational relationship, which remains essentially unchanged after adding controls for parental education and income (amongst others), and after instrumenting parents' test scores to isolate exogenous variation in such scores due to local education policy in the area in which they happen to live. Thus we can rule out the cause of the positive inter-generational relationship in reading test scores being due to parents with higher test scores achieving higher education levels and higher income, and the latter having the effect on the children's test scores. Similarly, the IV results suggest that the cause of the relationship is not a genetic effect, since even exogenous change in parents' reading scores seem to be passed on to their children, to the same extent as before. There is therefore a residual cause of the inter-generational relationship in reading scores that remains to be determined. We hypothesise that this is related to upbringing and parenting style, with parents who have a higher reading ability spending more time reading to their children and listening to them read, in essence being better reading 'teachers' to their children. The finding in Table 4 that those children who have been given more books obtain higher test scores is consistent with this hypothesis.

A different cause of the inter-generational relationship seems to be at work with respect to maths scores, however, where the estimated coefficient in the IV specification is essentially zero.

This suggests that exogenously increasing the maths skills of the parents (in this case via starting full-time school or the systematic learning of 'sums') will have no effect on the maths skills of their children. This in turn suggests that genetic effects are very important for the inter-generational transfer of maths skills, which cannot be passed on from one generation to the next via teaching in the home, in the same way that reading skills can be.

In terms of policy, the results presented above suggest that improving the reading scores and general literacy skills of one generation will also have a direct positive effect on the next generation, as the earlier generation acquire the skills to help their own offspring learn to read. With respect to maths scores, however, the results suggest that natural, genetic, ability is the most important determinant, and so raising the skills of one generation will not have an additional positive effect on the next generation as well. Each successive generation must cope with their inherited mathematical ability, or be taught the required mathematical skills anew.

ACKNOWLEDGEMENTS

We are grateful to the Data Archive, University of Essex, for supplying the *National Child Development Study*. We are also grateful to Arnaud Chevalier, Steve Machin and Ian Walker, plus conference participants at the Work, Pensions and Labour Economics Study Group (WPEG) Annual Conference in Sheffield (2008) and at the European Association of Labour Economists Annual Conference in Amsterdam (2008), for valuable comments. The normal disclaimer applies.

REFERENCES

Antonovics, K. and Goldberger, A. (2005). 'Does increasing women's schooling raise the schooling of the next generation? Comment', *American Economic Review*, Vol. 95, pp. 1738–44.

Ashenfelter, O., Oosterbeek, H. and Harmon, C. (1999). 'A review of estimates of the schooling/earnings relationship with tests for publication bias', *Labour Economics*, Vol. 6, pp. 453-470.

Behrman, J. and Rosenzweig, M. (2002). 'Does increasing women's schooling raise the schooling of the next generation?', *American Economic Review*, Vol. 92, pp. 323-334.

Belzil, C. and Hansen, J. (2003). 'Structural estimates of the intergenerational education correlation', *Journal of Applied Econometrics*, Vol. 18, pp. 679-96.

Bjorklund, A., Lindahl, M. and Plug, E. (2006). 'The origins of intergenerational associations: lessons from Swedish adoption data', *Quarterly Journal of Economics*, Vol. 121, pp. 999-1028.

Black, S., Devereux, P. and Salvanes, K. (2005). 'Why the apple doesn't fall far: understanding intergenerational transmission of human capital', *American Economic Review*, Vol. 95, pp. 437-49.

Blanden, J. (2004). 'Family income and educational attainment: a review of approaches and evidence for Britain', *Oxford Review of Economic Policy*, Vol. 20, pp. 245-263.

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

Brown, S. and Taylor, K. (2009). 'Social interaction and children's academic test scores: evidence from the National Child Development Study', *Journal of Economic Behavior and Organization*, forthcoming.

Brown, S. and Taylor, K. (2008). 'Bullying, education and earnings: evidence from the National Child Development Study', *Economics of Education Review*, Vol.27, pp. 387-401.

Bynner, J. and Parsons, S. (1997). *It Doesn't Get Any Better: The Impact of Poor Basic Skills on The Lives Of 37 Year Olds*, Basic Skills Agency, London.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter & D. Card (eds.), *Handbook of Labor Economics*, Vol. 3, Chapter 30, 1801-1863, Elsevier, Amsterdam.

Chevalier, A. (2004). 'Parental education and child's education: a natural experiment', IZA Discussion Paper Number 1153.

Chevalier, A., Harmon, C., O'Sullivan, V. and Walker, I. (2005). 'The impact of parental income and education on the schooling of their children', IZA Discussion Paper Number 1496.

Chevalier, A., Denny, K. and McMahon, D. (2007). 'A multi-country study of inter-generational educational mobility', UCD Geary Institute Working Paper 2007/25.

Dearden, L., Machin, S. and Reed, H. (1997). 'Intergenerational mobility in Britain', *Economic Journal*, Vol. 107, pp. 47-64.

Department for Education and Employment (DfEE) (2001). *Skills for Life: The National Strategy for Improving Adult Literacy and Numeracy Skills*, DfEE, London.

Dustmann, C. (2004). 'Parental background, secondary school track choice, and wages', *Oxford Economic Papers*, Vol. 56, pp. 209-30.

Ermisch, J., and Francesconi, M. (2001). 'Family matters: impacts of family background on educational attainment', *Economica*, Vol. 68, pp. 137-56.

Galindo-Rueda, F. and Vignoles, A. (2005). 'The heterogeneous effect of selection in secondary schools: understanding the changing role of ability', CEE Discussion Paper 52, Centre for the Economics of Education, LSE.

Greene, W. (2003). *Econometric Analysis* (Fifth Edition), Prentice-Hall International, New Jersey.

Hakkinen, I., Kirjavainen, T. and Uusitalo, R. (2003). 'School resources and student achievement revisited: new evidence from panel data', *Economics of Education Review,* Vol. 22, pp. 329-35.

James-Burdumy, S. (2005). 'The effect of maternal labor force participation on child development', *Journal of Labor Economics*, Vol. 23, pp. 177-211.

Machin, S. and McNally, S. (2008). 'The Literacy Hour', *Journal of Public Economics*, Vol. 92, pp. 1441-1462.

McIntosh, S. and Vignoles, A. (2001). 'Measuring and assessing the impact of basic skills on labour market outcomes', *Oxford Economic Papers*, Vol. 53, pp. 453-481.

Murnane, R., Willett, J. and Levy, F. (1995). 'The growing importance of cognitive skills in wage determination', *Review of Economics and Statistics*, Vol. 77, pp. 251-266.

National Longitudinal Survey Of Youth (1997) *User Guide*. Centre for Human Resources Research. The Ohio State University.

Oreopoulos, P., Page, M. and Huff Stevens, A. (2006). 'The intergenerational effects of compulsory schooling', *Journal of Labor Economics*, Vol. 24, pp. 729-60.

Plug, E. (2004). 'Estimating the effect of mother's schooling on children's schooling using a sample of adoptees', *American Economic Review*, Vol. 94, pp. 358-368.

Pringle, M., Bulter, N., and Davies, R. (1966). *11,000 Seven Year Olds*, Longman, London.

Southgate, V. (1962). *Southgate Group Reading Tests: Manual of Instructions*, University of London Press, London.

Sacerdote, B. (2002). 'The nature and nurture of economic outcomes', *American Economic Review,* Vol. 92, pp. 344-48.

Table 1: Summary Statistics

DEPENDENT VARIABLES (STAGES 1 & 2)	MEAN	STD. DEV.	MIN (%)#	MAX
PARENT (*NCDS* RESPONDENT)				
Maths Test Score aged 7	4.5266	2.7270	0 (*12.25%*)	10
Reading Test Score aged 7	21.3740	9.4535	0 *(10.20%)*	30
CHILD OF *NCDS* RESPONDENT				
PIAT Maths Test Score	38.1942	16.7871	0 *(5.79%)*	84
PIAT Reading Test Score	41.1737	20.2681	0 *(6.28%)*	84
CONTROL VARIABLES IN EQUATION (1)	**MEAN**	**STD. DEV.**	**MIN**	**MAX**
Age of child	8.6156	2.8274	5	18
Male child	0.4922	0.5001	0	1
Whether the child is in good health	0.8729	0.3332	0	1
Whether the child has any siblings	0.9026	0.2966	0	1
Frequency at which child bullies other children*	0.0885	0.3037	0	2
Whether the child is a loner	0.3401	0.2779	0	1
Number of books owned by child	2.7339	2.0449	0	5
Single parent family	0.2099	0.4073	0	1
Log household labour income	5.6258	2.3312	0	11.33
Home owned outright/mortgage	0.0239	0.1530	0	1
Parent's highest qualification: GCSE	0.3063	0.4611	0	1
Parent's highest qualification: A level	0.0141	0.1178	0	1
Parent's highest qualification: Diploma	0.0729	0.2601	0	1
Parent's highest qualification: Teaching/nursing	0.0318	0.1754	0	1
Parent's highest qualification: Degree	0.0604	0.2383	0	1

Notes: (i) # for the dependent variables used in the child and parent test score models of equations (1) and (2) the percentage of zero marks obtained in the test scores are shown in parenthesis; (ii) * for the bullying index, 0 denotes never bullied, 1 denotes sometimes bullied and 2 denotes often bullied.

Table 2: Frequency Distributions of the Instrumental Variables

AGE	FREQUENCY	PERCENT
AGE STARTED FULL-TIME SCHOOL:		
Under 3 ½ years	33	1.46
3 ½ - 4 years	43	1.90
4 – 4 ½ years	89	3.93
4 ½ - 5 years	1,006	44.40
5 – 5 ½ years	1,048	46.25
5 ½ - 6 years	37	1.63
over 6 years	10	0.44
AGE FIRST SYSTEMATICALLY TAUGHT PHONICS:		
Under 5 years	81	4.04
5 – 5 ½ years	615	30.64
5 ½ - 6 years	688	34.28
6 – 6 ½ years	382	19.03
6 ½ - 7 years	209	10.41
7 – 7 ½ years	32	1.59
AGE FIRST SYSTEMATICALLY TAUGHT SUMS:		
Under 5 years	29	1.47
5 – 5 ½ years	334	16.88
5 ½ - 6 years	896	45.28
6 – 6 ½ years	504	25.47
6 ½ - 7 years	176	8.89
7 – 7 ½ years	40	2.02

Table 3: First Stage Regressions. Dependent Variable is Parents' Age 7 Test Score

	READING TEST SCORE		MATHS TEST SCORE	
	COEFFICENT	T STATISTIC	COEFFICENT	T STATISTIC
START LEARNING PHONICS/SUMS:				
under 5 years	0.183	*(2.33)*	0.462	*(2.73)*
5-5 ½ years	0.190	*(5.20)*	0.081	*(1.46)*
5 ½ -6 years	(reference)		(reference)	
6-6 ½ years	-0.141	*(-3.34)*	-0.115	*(-2.35)*
6 ½ -7 years	-0.171	*(-3.28)*	-0.148	*(-2.04)*
7 – 7 ½ years	-0.754	*(-6.23)*	0.010	*(0.07)*
START FULL-TIME SCHOOL:				
under 3 ½ years	0.041	*(0.33)*	-0.034	*(-0.21)*
3 ½ - 4 years	0.121	*(-1.13)*	0.242	*(1.68)*
4 – 4 ½ years	0.116	*(1.46)*	0.282	*(2.61)*
4 ½ - 5 years	0.189	*(6.06)*	0.179	*(4.31)*
5 – 5 ½ years	(reference)		(reference)	
5 ½ - 6 years	0.055	*(0.47)*	-0.040	*(-0.26)*
over 6 years	-0.022	*(-0.08)*	-0.104	*(-0.31)*
male	-0.099	*(-3.21)*	0.139	*(3.35)*
constant	0.215	*(6.76)*	0.102	*(2.60)*
R^2	0.085		0.033	
OBSERVATIONS	1926		1894	

Table 4: The Determinants of the Test Scores of the Child: Specification 2

	READING TEST SCORE		MATHS TEST SCORE	
	COEFFICENT	T STATISTIC	COEFFICIENT	T STATISTIC
Male child	-0.031	*(-0.58)*	0.031	*(0.60)*
Whether child is in good health	-0.044	*(-0.58)*	-0.077	*(-1.04)*
Whether the child has siblings	0.042	*(0.58)*	0.106	*(1.55)*
Whether the child bullies other children	-0.149	*(-1.81)*	-0.139	*(-1.90)*
Whether the child is a loner	0.233	*(5.63)*	0.193	*(4.77)*
Number of books owned by child	0.030	*(2.49)*	0.032	*(2.66)*
Single parent family	0.053	*(0.86)*	0.075	*(1.28)*
Log household labour income	-0.005	*(-0.45)*	0.010	*(0.96)*
Home owned outright/mortgage	0.058	*(0.40)*	0.135	*(1.03)*
Parent’s highest qualification: GCSE	0.049	*(0.84)*	0.066	*(1.20)*
Parent’s highest qualification: A level	0.031	*(0.16)*	0.154	*(0.71)*
Parent’s highest qualification: Diploma	-0.014	*(-0.16)*	0.144	*(1.61)*
Parent’s highest qualification: Teach/nurse	0.210	*(1.74)*	0.195	*(1.69)*
Parent’s highest qualification: Degree	-0.063	*(-0.69)*	-0.098	*(-1.01)*
Parent’s test score at age 7	0.248	*(6.12)*	0.084	*(3.01)*
Intercept	0.223	*(-1.96)*	-0.369	*(-3.32)*
R^2	0.057		0.032	
OBSERVATIONS	1369		1353	

Table 5: Inter-Generational Coefficients: Coefficients on the Parental Age 7 Test Score. Dependent Variable is Children's Test Score

	CHILD READING TEST SCORE		CHILD MATHS TEST SCORE	
	COEFFICENT	T STATISTIC	COEFFICENT	T STATISTIC
1: Control for gender only	0.253	*(7.13)*	0.091	*(3.43)*
2: All controls	0.248	*(6.12)*	0.084	*(3.01)*
3: Parents' test score instrumented	0.255	*(1.87)*	-0.039	*(-0.30)*
OBSERVATIONS	1369		1353	

Notes: (i) Specifications 2 and 3 control for child gender and health, presence of siblings, whether child bullies, whether child is a loner, number of books owned, single parent family, household income, housing type and *NCDS* respondents' (parents') education. (ii) T statistics derived using standard errors corrected for clustering of children within families. The standard errors in specification 3 are also bootstrapped, to allow for the fact that parents' test scores were estimated in the first stage regression.

Table 6: Inter-Generational Coefficients, Controlling for Partners' Education: Coefficients on the Parental Age 7 Test Score. Dependent Variable is Children's Test Score

	CHILD READING TEST SCORE		CHILD MATHS TEST SCORE	
	COEFFICENT	T STATISTIC	COEFFICENT	T STATISTIC
1: Control for gender only	0.253	*(7.13)*	0.091	*(3.43)*
2: All controls	0.211	*(4.73)*	0.061	*(2.03)*
3: Parents' test score instrumented	0.228	*(1.64)*	0.004	*(0.03)*
OBSERVATIONS	1095		1081	

Notes: (i) Specifications 2 and 3 control for child gender and health, presence of siblings, whether child bullies, whether child is a loner, number of books owned, single parent family, household income, housing type, *NCDS* respondents' (parents') education, and partners' education. (ii) T statistics derived using standard errors corrected for clustering of children within families. The standard errors in specification 3 are also bootstrapped, to allow for the fact that parents' test scores were estimated in the first stage regression.