CENTRE FOR ECONOMIC PERFORMANCE

DISCUSSION PAPER NO. 311

DO PEER GROUPS MATTER? PEER GROUP VERSUS SCHOOLING EFFECTS ON ACADEMIC ATTAINMENT

November 1996

D. ROBERTSON and J. SYMONS

ABSTRACT

In this paper we estimate an educational production function. Educational attainment is a function of three types of inputs: peer group, parental and schooling. We find that conventional measures of school quality are not good predictors for academic attainment, once we control for peer group effects. Parental qualities also have strong effects on academic attainment. This academic attainment is then a key determinant of subsequent labour market success, as measured by earnings.

This paper is produced as part of the Centre's
Programme on Human Resources

DO PEER GROUPS MATTER? PEER GROUP VERSUS SCHOOLING EFFECTS ON ACADEMIC ATTAINMENT

D. ROBERTSON and J. SYMONS

NOVEMBER 1996

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London
WC2A 2AE

ISBN 0 7530 0945 5

DO PEER GROUPS MATTER?
PEER GROUP VERSUS SCHOOLING EFFECTS ON ACADEMIC ATTAINMENT

D. ROBERTSON and J. SYMONS

The Centre for Economic Performance is financed by the Economic and Social Research Council.

ACKNOWLEDGEMENTS

We are grateful to the ESRC Data Archive for providing access to the NCDS data, to Leon Feinstein, Hashem Pesaran and Nicholas Rau for helpful suggestions, and to seminar participants at UCL, LSE and Oxford University.

Donald Robertson is a Lecturer in Economics at Cambridge University and Jim Symons is a Reader in Economics at University College, London. Both are Research Associates at the Centre for Economic Performance, LSE.

DO PEER GROUPS MATTER?
PEER GROUP VERSUS SCHOOLING EFFECTS ON ACADEMIC ATTAINMENT

D. ROBERTSON and J. SYMONS

"I don't understand all this fuss about education. None of the Paget family can read or write, but they do very well."

William Lamb, 2nd Viscount Melbourne

1. Introduction

This paper conducts an empirical investigation of the causes of academic achievement in British children at age 11. We investigate as well the relationship between this achievement and subsequent labour market success.

There is a view in the literature of the economics of education that, to a first approximation at least, the academic attainment of children in schools is completely explained by the education, income and social class of their parents; and by the levels of these factors in the child's peer group. School quality as measured by conventional inputs such as class size, teacher experience and general school expenditure are typically found to have only minor effects. This view originates with the work of the Coleman Report (1966), and is extensively reviewed and discussed in Hanushek (1986). The Coleman report emphasised the pre-eminence of peer group effects on educational attainment. More recent studies are Summers and Wolfe (1977), Henderson, Mieszkowski and Sauvageau (1978) and Borjas (1992). One expects peers to affect achievement directly (eg helping each other with course work) and also via values, though it seems to us likely that the latter effect predominates.

Recently Card and Krueger (1992) have challenged this accepted wisdom[1]. They find a significant correlation between the returns to schooling in the states of the United States and state-wide educational quality as measured by conventional inputs. Whereas this study gives results for labour market outcomes, and the Coleman-Hanushek evidence relates to achievement, the results are clearly at variance with the earlier literature. The question is of immense significance for public policy: if Card and Krueger are right, the problem of the poor could be alleviated within one or two generations by increased school expenditures, whilst if they are wrong, such expenditures would be a complete waste.

One curiosity with the Card and Krueger work is that they find that family background is not significantly related to the returns from schooling. This is so contrary to a great number of studies that it suggests something may be wrong with their findings. One way this might occur is if the intensity of parental desire for academic success were to vary across states. In this case one would expect to see the conventional inputs well correlated with attainment, even if the direct effect of the inputs on attainment were slight. Indeed, in principle, the lavishness of expenditure on the school system could be an almost perfect proxy for this intensity: certainly better than such proxies as state-wide parental education and income.

We shall study these issues using a cohort of British children born in a particular week in 1958 (the National Child Development Study, NCDS). Section 2 describes the model we shall use to address these issues. Section 3 describes the estimation results and econometric issues in greater detail. Section 4 discusses the

problems raised by attrition in the NCDS. Section 5 provides further analysis of the peer group effects, and Section 6 concludes.

Overall our evidence points towards the *parents and peers* theory of educational attainment in children, giving only a minor role to school inputs. This attainment is then rewarded in the labour market.

2. The Educational Production Function

The central theoretical concept in this study is the educational production function. Attainment is viewed as generated by peer group inputs, parental inputs and conventional school inputs. In principle, attainment at any date will depend on the entire history of these inputs prior to that date. Attainment will also depend on inherent ability, some part of which may be genetic (and thus could be thought of as an original parental input), and some part of which may be due to chance.

Following Hanushek (1992), we write attainment at time t, A_t, to be some function

$$A_t = \Phi(\mathbf{P}_t,\mathbf{F}_t,\mathbf{S}_t) \tag{1}$$

where P_t, F_t and S_t are the streams of past peer group, family and schooling inputs respectively obtained by time t. If prior achievement is also observed, then, under certain assumptions, this function can be rewritten as

$$A_t = \phi(A_{t-1},P_t,F_t,S_t) \tag{2}$$

where P_t, F_t, and S_t now represent the inputs from t-1 to t. Parents have an intertemporal utility function defined over the ultimate welfare (indexed here by attainment) of their children, together with their own consumption and leisure. This utility function is maximised subject to both financial and time constraints and the production function ϕ. This makes clear, as is almost always the case in production theory, that the observed factor inputs are endogenous. The problem that will usually arise in estimating (2) is that parents who are exceptionally ambitious for their children (in a manner unobserved by the econometrician) will choose high levels of the productive inputs as well as aiding the child directly. This gives biases if simple regression estimates are interpreted as structural parameters of the production function. These biases apply in principle to most of the variables we investigate below, from those measuring school quality, extending even to family structure variables such as number of children or presence of a father. The usual solution is to apply an instrumental variable estimator. The outcome from the optimisation above would give reduced forms of the endogenous variables as functions of exogenous variables, most notably prices of labour and prices of schooling inputs. These we do not observe directly, but if we assume these prices are constant across local markets, regional indicator variables would then be possible instruments, provided that preferences and technology do not vary systematically over regions, and that an individual's region is not itself endogenous.[2] The shortage of plausible instruments imposes an upper limit on the amount of endogeneity that can be handled. We shall use these instruments to correct endogeneity bias where it seems most likely (the variables most easily chosen by parents), and in variables of

particular interest to this study, those reflecting school and peer group. In general we shall treat other variables, such as mother's labour force status as exogenous. Where bias does occur, its likely direction is often known, thus aiding inference.

3. The Data and Empirical Estimation

The national child development study

The National Child Development Study is an ongoing survey of a particular cohort of the UK population: every child born in the UK during one week in March 1958. Information has so far been collected at five points in their lives on a wide variety of personal, parental, educational and occupational subjects. The first sweep of the survey (NCDS1) collected information in 1965 when the children were 7 years old. This information comprised mainly parental data, a large amount of health information on the child and a certain amount of early schooling data. Subsequent sweeps took place in 1969 (NCDS2), 1974 (NCDS3), 1981 (NCDS4) and 1991 (NCDS5). At each sweep further parental, health and educational data were recorded, including scores on ability tests. At the final two sweeps a large amount of labour market data were also collected, in addition to educational attainment data. At the later sweeps, data from individuals of the same cohort not born in the UK, but now resident in the UK, were also collected. The original NCDS sample consisted of 18,359 individuals. It may be viewed as a representative subsample of the population of the UK born in the late 50s.

Unfortunately not only has there been attrition over the life of the survey, but also all variables are not available for all individuals remaining in the study. Consequently each estimation is performed over the largest possible subsets for whom the required data are available. If attrition and non-response are uncorrelated with the explanatory variables in the regressions then our estimated coefficients remain unbiased. Below we perform some simple analyses regarding the non-responding individuals that will bear on this issue.

Estimation and results

The first set of reported regressions link measured abilities of the individual at age 7 to parental and family variables. Ability is measured by test scores of maths and reading attainment. Figures 1 and 2 graph the distribution of the raw scores for the sample renormalised to 0-100. An immediate problem is the bunching of the reading scores at the upper limit. It seems clear that the reading test here is too coarse and is failing to discriminate adequately for the individuals in the upper tail of the distribution of the reading tests. This bunching of scores would clearly be detrimental to our subsequent analysis. The dataset contains a set of variables describing teachers' assessments of the abilities of the children at age seven in a range of categories/subjects and Table 1 reports these assessments as explanatory variables in a tobit regression for reading score assuming truncation at score 100. The regression results indicate that the explanatory variables have strong predictive power for the reading score. The predicted values from this regression are used in place of the raw reading score throughout the subsequent analysis. Figure 3 graphs these predicted values, normalised to 0-100. Note that the tobit-corrected predicted values produce a bell-shaped distribution of reading scores.

Our interpretation of the reading and maths scores at age 7 is that they give some measure of initial ability, and potentially some early schooling effects. This initial ability will be a combination of genetic inheritance and pre-school parental inputs which we shall not seek to disentangle. Table 2 relates these maths and (both uncorrected and tobit-corrected) reading scores at age 7 to a set of parental and schooling variables. The parental variables may loosely be grouped into variables reflecting parental quality and those reflecting parental time. We proxy parental time by whether the mother[3] works (full time or part time), the presence of a father, the number of children in the household. The quality of parental time is measured by indicators of parental education and social class. We imagine that these quality variables affect attainment both directly (eg an educated parent is better able to help with homework) and indirectly via the transmission of values (eg punctuality, self-control, ambition). We find that attainment is, in general, negatively related to the number of children in the household (at the time of the tests). Having a working mother reduces attainment at 7 (particularly if she works full time). It is problematic to interpret these results as the *ceteris paribus* effect of a woman's going out to work, since labour market status is a decision variable of the household. Several different biases are possible. It may be that women who choose to work have a lower level of interest in their children's development. On the other hand, those women who work may be those with higher marginal product, possibly reflected in superior child raising. There are also possible income effects. Absence of father has in general a negative effect on attainment[4], but is statistically significant only for reading. Social class and education of parents typically have the expected effect: the children of the educated and the higher social classes do better. Note that having parents who remained in education beyond the statutory minimum leaving age contributes about ten extra marks (out of 100) in both reading and maths, as does membership of the top socio-economic status (SES) group (proxied by father's occupation). We find that girls are better at reading at 7 but worse at maths. Why this should be so is beyond the scope of this article but we note that this has been found by other scholars in this area: for example, Harbison and Hanushek (1992) find similar results for rural Brazil. It is interesting that we are able to explain less than half the variance in maths scores that we explain for reading scores. It may be that reading skills are easier to impart than mathematical skills.

Reading and maths tests were re-administered at age 11. These are normalised to the range [0,100] and graphed in Figures 4 and 5. Our regressions explain the improvement in scores by this age. The influence of the peer group is measured by the percentage of the child's classmates who have fathers in the top socio-economic group at 7. There is a natural bias towards finding peer group effects in empirical studies since membership of a peer group is endogenous. If parents actively choose the schools their children attend, then one presumes ambitious parents will seek good schools for their children as well as helping them in other ways. Given that parental ambition is unobserved, empirical studies will tend mistakenly to allocate to the peer group the effect of the unobserved extra parental influence. A similar selection bias has been studied by Evans *et al* (1992) for the incidence of teenage pregnancy and school dropout behaviour. We shall deal with this bias, or at least reduce its effect, by using instrumental variables, taking as instruments 11 possible regions of birth in Britain. These are significantly correlated with our peer variable (eg better schools

are found in the south) but are open to the objection that ambitious parents might well choose to live in regions where there are good schools. This seems unlikely to influence the region of birth of the oldest child in a family, but could pose problems for younger siblings if parents have already moved prior to the younger child's birth to obtain a better peer group for an older child. This would contaminate the instruments. In our dataset, looking just at first-born children, we fitted a logit model giving the probability of a cross region move between birth and 7 as a function of the difference in the peer group variable between those regions as well as 'pseudo-distance' (0 for no move, 1 for a move to contiguous regions, 2 for moderate moves and 3 for moves across the country). The peer group effects were found to be correctly signed but insignificant. Thus there seems only a slight tendency for parents to move to better regions after the birth of a first child, so that region of birth for second and higher children can also be treated as exogenous. We infer that region of birth dummies are likely to be exogenous for all and are valid instruments for the peer group effect.

The exact functional form of the relationship between attainment and the peer group is also an issue of substance in this work because, with diminishing returns to average peer quality, it is optimal utilitarian policy to mix children together as much as possible. On the other hand if there were increasing returns, it would be best to segregate them. We have discussed this elsewhere (Robertson and Symons (1996)). In fact, we find decreasing returns in all the models studied. Henderson *et al* (1978) have found diminishing returns to the peer variable measured as average class IQ in a study of third grade children in the United States. We investigate diminishing returns to the peer group variable by including quadratic terms.

In these regressions we are able to investigate a further peer group effect that arises from *streaming*, the practice of sorting children within schools into ability groups for teaching. About one third of the children in our sample attended streamed schools and we study them separately[5]. The NCDS records three ability groups within streamed schools and each of these will have a potential peer-group effect on children placed in them. We shall find that being in the bottom stream has a powerful negative effect on improvement between ages 7 and 11 in both maths and reading. The information about the stream to which the child belongs is taken at age 11 and there is thus another endogeneity problem, in that children who do exceptionally well between 7 and 11 will tend to be placed in the top stream. We evade some of this problem by instrumenting the streaming variables by a set of variables recording teachers' opinions of the abilities of the children at age 7. This procedure is open to the objection that teachers may have superior knowledge about which children are likely to experience unusual development between 7 and 11 so that instruments are not orthogonal to equation error. This is fairly plausible and we regard any evidence for peer group effects derived from streaming as less convincing than that obtained from our socio-economic variables discussed above.[6]

For each improvement regression we include the test score at age 7. Instrumented by teachers' assessments, this measures initial ability and may also capture dynamic elements of human capital formation. This variable will enable us to judge the extent to which differences in measured ability at age 7 persist to age 11. We continue to use the parent quality and parent time variables from the model of Table 2. Finally we include variables measuring conventional school inputs: class-

size and the percentage of teachers in the school with over three years teaching experience. Results are given in Tables 3 and 4. To begin with, we treat peer group and school quality as exogenous variables.

By and large, these results are not much different for maths and reading and we concentrate on the maths results. For the un-streamed schools, the parental variables have a qualitatively similar effect on the improvement between 7 and 11 as they had on the attainment at age 7: the children of the educated and the successful do best. Family size is broadly as before. The absence of a father has a negative but insignificant effect. In contrast to early attainment, a working mother does not have an important effect on improvement between 7 and 11. For the school-input variables, teachers' experience is very weak and class size is incorrectly signed[7]. The peer group variables, in contrast, are strong. There appears to be diminishing returns to the peer group effect.

Turning to the streamed schools, note that **both** sets of peer group variables, membership of ability group and per cent of fathers in the top socio-economic groupings, are important. Most other variables however have smaller parameters and reduced significance compared to unstreamed schools. The exception is score-at-7, the effect of which is that students retain nearly their full 7-years-score at 11 years in a un-streamed school, but only about 60% of this in a streamed school.

There are some intriguing differences between streamed and unstreamed schools. In streamed schools, scores in maths and reading at 11 are significantly related to little other than the peer group variables, whilst for unstreamed schools, parental quality remains extremely strong. One interpretation of this is that, in streamed schools, the values inculcated by a peer group comprised of children of similar abilities are much more potent than those derived from the family. Thus early performance or early ability is much less important in a streamed school. One can interpret the differences between the parameters in Tables 3 and 4 as saying that the heightened peer-group effects of the streamed school replace the flow of improved attainment derived from parental variables in non-streamed schools.

As emphasised above, both school input variables and the peer group variables are potentially endogenous. As we have said, these can be instrumented by dummy variables for area of birth in 11 broad regions of the UK. The instruments are fairly weak, so to sharpen our estimates we cast out consistently poor variables; the peer group quadratic term and mother's labour force status from both streamed and unstreamed models as well as the parental SES and educational variables from the streamed models. Table 5 reports estimates for these three variables. While peer groups remain correctly signed and moderately significant three times out of four, the other two are correctly signed and significant only once: teacher's experience for maths improvement in the unstreamed school.

We now turn to a set of regressions relating the attainment variables to earnings at age 33. The dependent variable is the log of earnings per week for current or most recent job. We discard as outliers those earning less than £100 per week and those over £500 per week. We retain all right-hand side variables from the improvement regressions to capture any on-going effects from these variables. Table 6 reports results separately for boys and girls. Note first the effect of educational attainment is similar between males and females. Both maths at 7 and the improvement by age 11 are richly rewarded; the reading variables are similar but less

powerful. The coefficients are such that, in the determination of earnings, the scores at age 11 are nearly a sufficient statistic for what has gone before, especially for girls. For boys, there is some evidence of on-going effects beyond age 11 of parental SES variables, with little measured effect for girls. This differential effect is something of a mystery. There is some evidence that number of siblings has a negative effect on earnings. The evidence for school quality variables is mixed. For boys these variables are wrong-signed and insignificant. For girls, however, the results indicate some correctly signed effects. We experimented also here with instrumenting the class size variable: again this result was qualitatively unchanged. We find little evidence for the widely held belief that smaller classes offer a better environment for educational achievement.

4. Attrition in the NCDS

Of the original 18,359 individuals eligible for inclusion in the NCDS only some 12,000 remain at age 33. Additionally the full set of variables is not uniformly available for even this subsample, so that for example our earnings regressions are estimated only on about 2,000 individuals. Such severe attrition must give cause for concern. One could in principle model the probability of response and the regression models jointly, conditioned on exogenous variables. However in the NCDS this is not possible, as there is no set of exogenous variables that exist for all individuals in the original cohort. There thus seems no way formally to correct the regression results for possible non-random attrition bias. The extent of the problem can be gauged from Table 7 which summarises the distribution of maths scores at age 7 for those individuals included in regressions at age 7, 11 and 33. It is immediately apparent that we are undersampling the lower tail of the distribution as the cohort ages by this measure. At age 11, when our score improvement regressions are performed, this seems to be a relatively minor problem. At age 33 where the increase in mean is of greater significance, part of this increase is of course due to the fact that those employed tend to be drawn from the upper portion of the maths score distribution. In fact the mean maths score at age 7 of those included in NCDS but not reporting earnings at age 33 is about 2 points below the mean at age 7 recorded in Table 7. Thus the high mean for those included in the regressions at age 33 is likely to be due in large part to this effect rather than genuine attrition.

5. Peer Group Effects and Streaming

There are two peer group effects studied in this paper: the effect of classmates coming from higher socio-economic groups, and the effect caused within schools by streaming. Students benefit from being placed in the top stream of streamed schools, but suffer if they are placed in lower streams. Figure 6 graphs the expected maths score at age 11 as a function of score at age 7 for students in unstreamed schools and in the three possible streams of streamed schools.[8] Those placed in the top stream benefit from attending streamed schools at all initial maths scores with the greatest benefit seen at lower scores, clearly supportive of a peer group effect. Attendance at a streamed school and being placed in the lower ability stream actually provides a worse outcome than attending an unstreamed school at all initial maths score.

Again this supports the peer group effect, the average quality of peers in an unstreamed school presumably exceeding that of the lowest stream in a streamed school.

6. Conclusions

This paper has studied the determinants of academic attainment at age 11 and the links between this attainment and earnings at age 33. We have found strong evidence for the importance of peer groups. We have found also strong effects from parental social class and parental academic achievement. To see the strength of these effects it is instructive to calculate the implied increment in expected earnings at age 33 for a male with father in the top socio-economic group, whose parents both stayed at school beyond minimum leaving age , and who attended an unstreamed school in which half of fellow pupils have parents in the top social class. Using our estimated models, we find reading and maths scores at age 7 rise by 19.0 and 17.1 respectively. At age 11 this plus ongoing effects from parents and peers raise scores by 21.5 and 33.8 points in reading and maths respectively. The earnings regressions translate these increases into an extra 31% in income.

ENDNOTES

1. See also Altonji and Dunn (1995) for support, and Betts (1995) for a defence of the traditional view.

2. This issue of possible endogeneity is investigated below. We find that this is unlikely to cause problems here.

3. Throughout mother refers to mother or mother-figure, father to father or father-figure.

4. There is a complication here due to the fact that we are measuring the family's social class by the occupation of the father. This creates problems when no father is present. Accordingly we have fitted a three-outcome multinominal logit of father's socio-economic status (SES) as a function of the variable registering whether the mother stayed on for post-compulsory education; and used the predicted values for our SES variable when no father was present. This should ensure that the presence-of-a-father variable genuinely measures the extra effect of a father in a family of given social class.

5. The information on whether schools did stream their pupils was collected at child age 11. Those children who had entered secondary education by this age were censored from our sample so that this information does not refer to the practice of a school at which the child had just started.

6. Another reservation about interpreting the effects of streaming variables as due to the peer group is that curricula and teaching methods may depend on the ability of the class. Thus, for example, the top stream may do better merely because they have been taught more advanced material.

7. Experiments with a variety of non-linear forms of both these variables also failed to find any significant effects.

8. These graphs are derived from Table 3 setting all right-hand side variables except the streaming dummies to zero.

TABLE 1

Tobit Regressions for Reading Scores at Age 7

	Coef	t-statistic
oral ability	-1.1	-4.8
awareness	-1.8	-6.6
reading ability	-12.4	-45.9
creativity	-0.6	-2.4
maths ability	-1.9	-8.4
comprehension	-6.8	-47.8
constant	156.7	233.2

Number of Obs. 11955 uncensored observations.
2797 right-censored observations at reading score>=100.

Notes: (i) Dependent variable is reading score at age 7, scale 0-100, assumed truncated at 100.
(ii) The explanatory variables in this regression are teacher's assessments of child's ability on a scale 1 (highest) to 5 (lowest), except for comprehension marked 1 (highest) to 7 (lowest).

TABLE 2

Reading and Maths Scores at 7

		Reading Score		Corrected Reading Score		Maths Score	
Parent Quality	Top SES father	7.5	(9.6)	7.1	(10.9)	6.5	(7.4)
	Middle SES father	4.4	(7.9)	3.2	(6.8)	2.8	(4.4)
	Father stayed on	5.5	(8.9)	6.1	(11.7)	4.8	(7.4)
	Mother stayed on	5.1	(8.9)	5.8	(12.2)	5.8	(9.0)
Parent Time	Mother works FT	-3.1	(3.7)	-1.9	(2.8)	-2.8	(3.1)
	Mother works PT	-1.0	(1.7)	-0.7	(1.6)	-0.8	(1.2)
	No father	-5.3	(3.7)	-4.0	(3.4)	-1.6	(1.0)
	Family size	-2.7	(18.7)	-2.3	(18.7)	-0.8	(5.0)
	Sex (=1 if girl)	6.3	(14.0)	5.1	(13.7)	-2.5	(5.0)
	Constant	78.7	(104.6)	53.9	(86.4)	51.2	(60.9)
	R^2	0.13		0.16		0.056	
	Number of Obs	8749		8701		8733	

Notes: (i) Absolute t-statistics in parentheses.
(ii) The top SES group is defined as professional, managerial and skilled non-manual occupations; the middle SES group as semi-skilled and skilled manual occupations.
(iii) 'Stayed on' means remained at school beyond compulsory schooling age.
(iv) Family size is the number of children under 21 in the family.
(v) Mother works FT and PT are respectively dummy variables on whether mother worked full time or part time before the child started school.

TABLE 3
Maths Improvement 7 - 11

		Streamed School		Unstreamed School	
	Maths score at 7*	-0.4	(3.3)	0.0	(0.1)
	Top stream*	0.7	(0.0)	-	
Peer	Bottom stream*	-38.3	(2.1)	-	
Group	Per cent top SES	0.41	(4.2)	0.16	(2.9)
	(Per cent)^2	-0.0039	(3.3)	-0.0004	(0.6)
	Father in top SES	2.1	(1.0)	4.9	(3.1)
Parent	Father in middle SES	-0.3	(0.2)	3.2	(3.2)
Quality	Father stayed on	1.6	(1.0)	2.2	(1.8)
	Mother stayed on	2.4	(1.7)	2.8	(2.6)
	Mother full time work	0.6	(0.0)	-0.0	(0.0)
Parent	Mother part time work	1.9	(1.0)	1.0	(0.8)
Time	No father	-1.5	(0.5)	-2.3	(0.9)
	Family size	0.2	(0.4)	-1.7	(5.8)
School	Teachers' experience	-0.02	(0.5)	0.00	(0.2)
Quality	Class size	-0.05	(0.5)	0.14	(2.4)
	Sex (= 1 if girl)	-2.9	(1.9)	1.2	(1.4)
	Constant	21.6	(2.0)	-18.7	(5.7)
	R^2	0.32		0.05	
	Number of Obs	1454		3135	

Notes overleaf

Notes: (i) Absolute t-statistics in parentheses.

(ii) The variable names are largely as in Table 2. Exceptions are that here mother works FT and PT refer to mother working full time or part time when child aged 11. 'Teachers experience' is the percentage of teachers in the school with over three years teaching experience. 'Per cent top SES' is the percentage of the child's class with fathers in the

top SES class, measured 0-100, where the definitions of SES are as in Table 2.

(iii) Variables marked with an asterisk are treated as endogenous. Instruments are teacher's assessments of student's abilities at 7 by the six criteria used in the regression reported in Table 1.

TABLE 4
Reading Improvement 7 - 11

		Streamed School		Unstreamed School	
	Reading score at 7*	-1.1	(4.9)	-0.5	(38.0)
	Top stream*	0.2	(0.0)	-	
Peer	Bottom stream*	-54.8	(1.9)	-	
Group	Per cent top SES	0.21	(2.2)	0.09	(3.0)
	(Per cent)^2	-0.0011	(1.0)	-0.0003	(1.0)
	Father in top SES	0.2	(0.1)	4.8	(5.8)
Parent	Father in middle SES	-0.7	(0.5)	1.6	(2.9)
Quality	Father stayed on	0.7	(0.4)	1.3	(2.0)
	Mother stayed on	2.4	(1.6)	2.3	(3.8)
	Mother full time work	-2.1	(0.9)	-0.1	(0.1)
Parent	Mother part time work	-1.8	(0.7)	0.2	(0.3)
Time	No father	2.1	(0.7)	-1.7	(1.3)
	Family size	0.8	(1.0)	-1.0	(6.2)
School	Teachers' experience	0.00	(0.1)	0.01	(0.6)
Quality	Class size	-0.06	(0.6)	0.06	(1.9)
	Sex (= 1 if girl)	-3.5	(2.9)	-2.9	(6.3)
	Constant	67.5	(2.7)	15.9	(9.5)
	R^2	0.00		0.34	
	Number of Obs	1456		3142	

Notes: As for Table 3.

TABLE 5

Improvement 7-11:
treating peer group and school inputs as endogenous

	Maths Improvement		Reading Improvement	
	Streamed	Unstreamed	Streamed	Unstreamed
Per cent top SES	0.12 (1.6)	0.20 (2.0)	0.19 (3.1)	-0.01 (0.3)
Teachers' Exp	-0.18 (1.4)	0.57 (3.9)	0.13 (1.2)	-0.26 (3.3)
Class Size	-0.37 (1.1)	1.67 (4.4)	0.28 (1.0)	-0.05 (0.3)

Notes: Other variables from Table 4 were fitted but are not reported. Instruments were dummy variables for area of birth in 11 regions of the UK.

TABLE 6
Earnings Regressions

		Males		Females	
	Maths score at 7	0.003	(4.8)	0.004	(4.2)
	Maths Improvement	0.002	(2.6)	0.003	(3.1)
	Reading score at 7	0.003	(3.6)	0.003	(2.3)
	Reading Improvement	0.002	(1.9)	0.003	(2.0)
Peer	Per cent top SES	0.001	(0.7)	0.002	(1.3)
Group	$(\text{Per cent})^2$	-0.000	(0.2)	-0.000	(1.6)
	Father in top SES	0.07	(2.0)	0.00	(0.0)
Parent	Father in middle SES	0.08	(3.7)	0.00	(0.0)
Quality	Father stayed on	0.02	(0.8)	0.06	(1.7)
	Mother stayed on	0.04	(1.7)	0.01	(0.5)
	Mother full time work	-0.04	(1.4)	-0.1	(0.4)
Parent	Mother part time work	-0.01	(0.3)	-0.03	(0.7)
Time	No father	-0.01	(0.1)	0.07	(0.7)
	Family size	-0.01	(1.6)	-0.02	(1.7)
School	Teachers' experience	-0.000	(1.1)	0.001	(1.6)
Quality	Class size	0.001	(0.6)	-0.003	(1.7)
	Constant	5.24	(68.4)	5.04	(46.9)
	R^2	0.19		0.15	
	Number of Obs	1062		754	

Notes: As for Table 3.

TABLE 7

Attrition Distributions

Distributions of maths score at age 7
for regression samples at age 11 and 33

	Score							
	0-20	21-40	41-60	61-80	81-100	mean	SD	No.Obs.
Age 7	14.58	25.97	27.60	21.31	10.55	52.35	24.27	8733
Age 11	14.75	25.58	27.35	22.05	10.26	52.38	24.37	4589
Age 33	13.77	25.17	26.54	22.85	11.67	53.74	24.49	1816

Note: The first row gives the percentage distribution, the mean and the standard deviation of maths scores at age 7 for those individuals included in the regression of Table 2; the second row for those individuals included in the regressions of Table 3; the third row for those individuals included in the regressions of Table 6.

FIGURE 1

Reading Score at Age 7

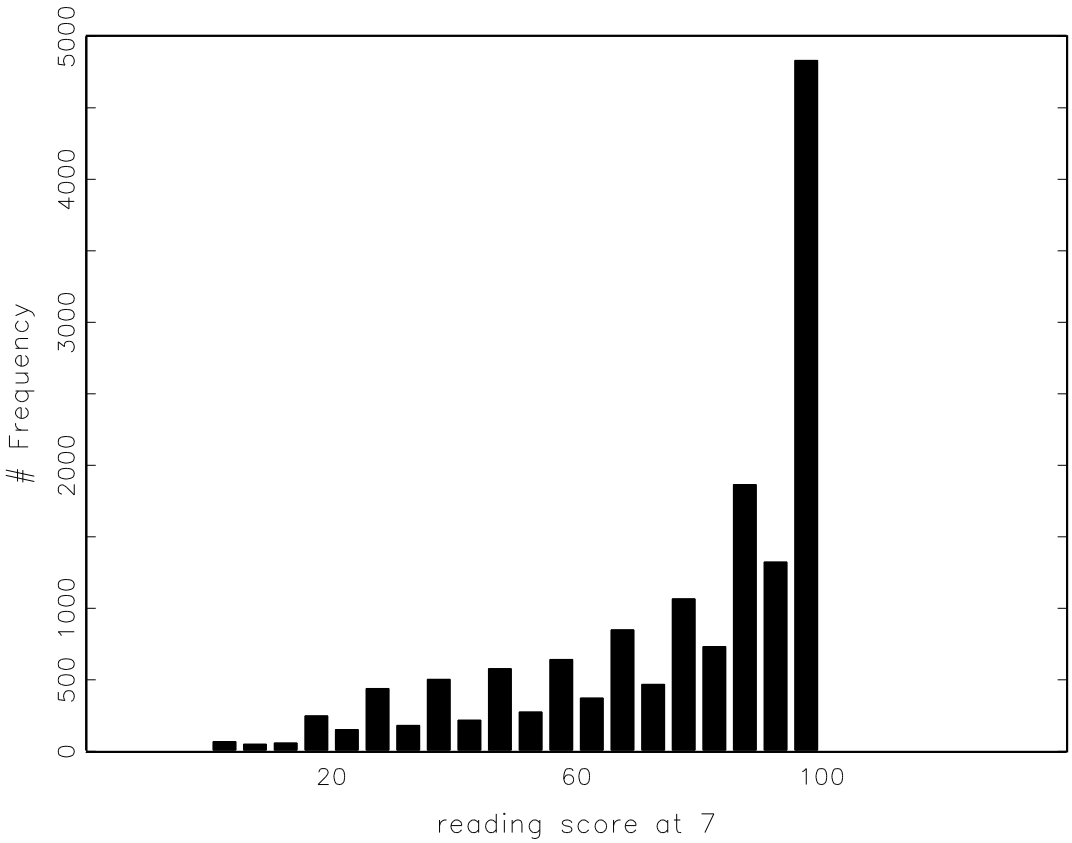


FIGURE 2

Tobit-Corrected Reading Score at Age 7

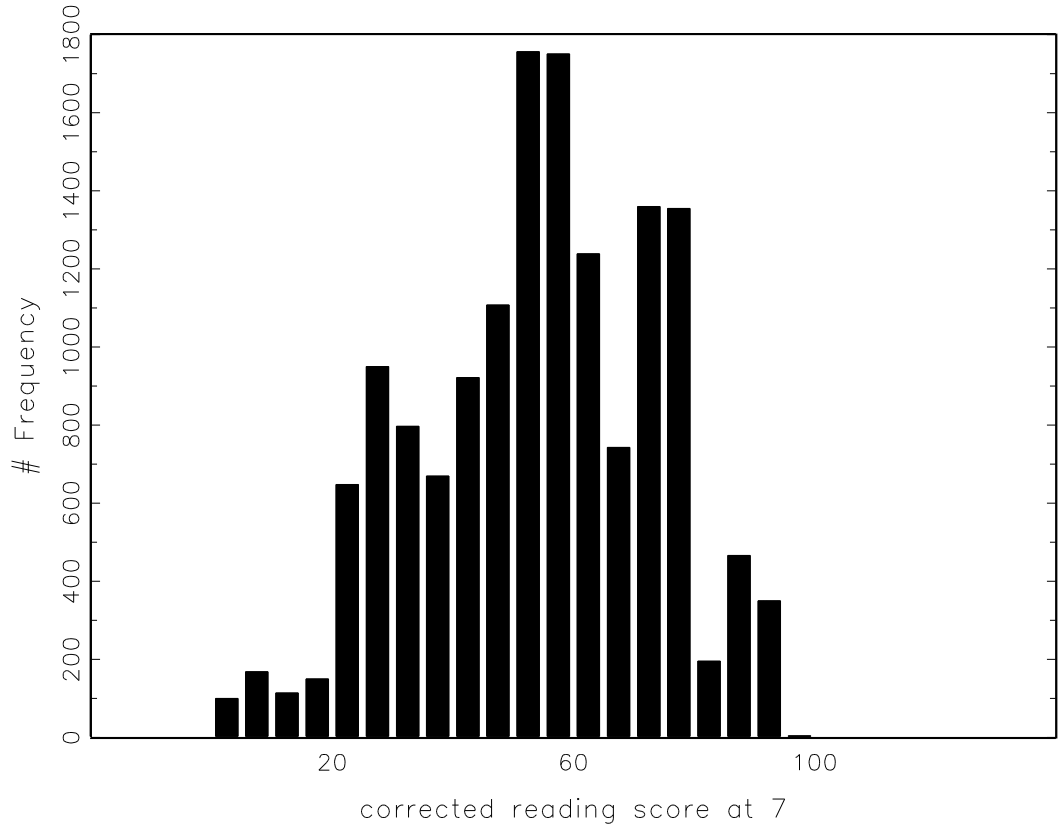


FIGURE 3

Maths Score at 7

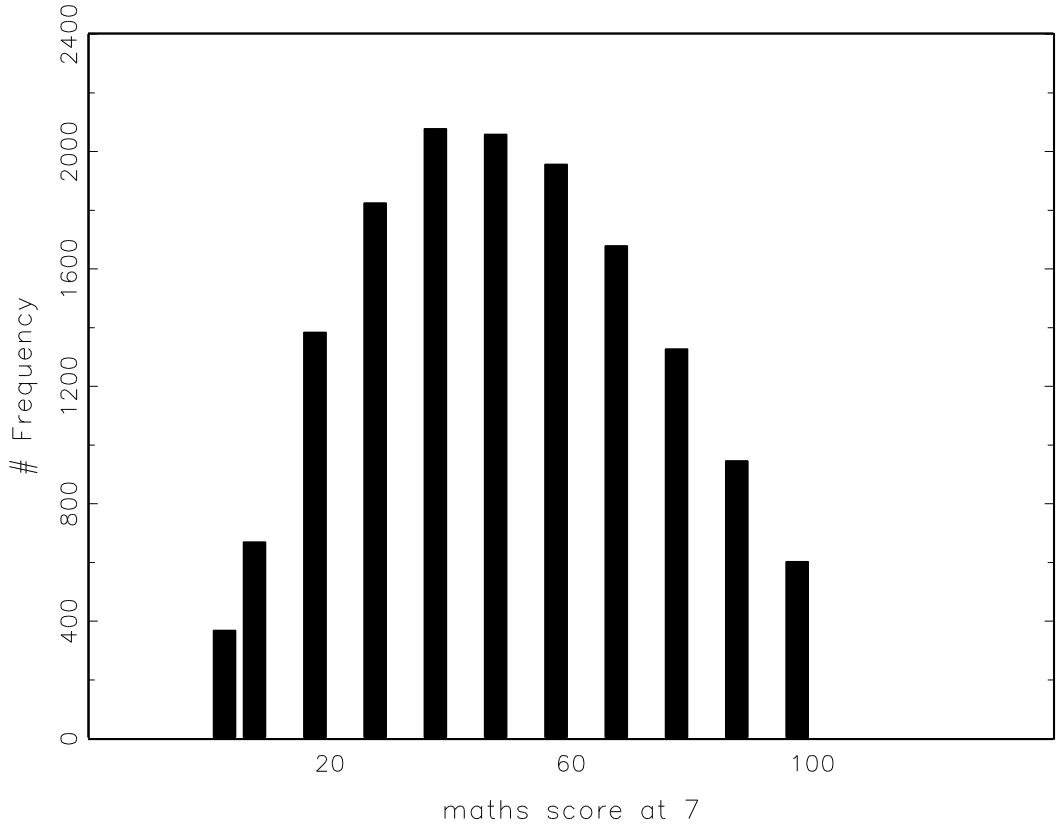


FIGURE 4

Reading Score at 11

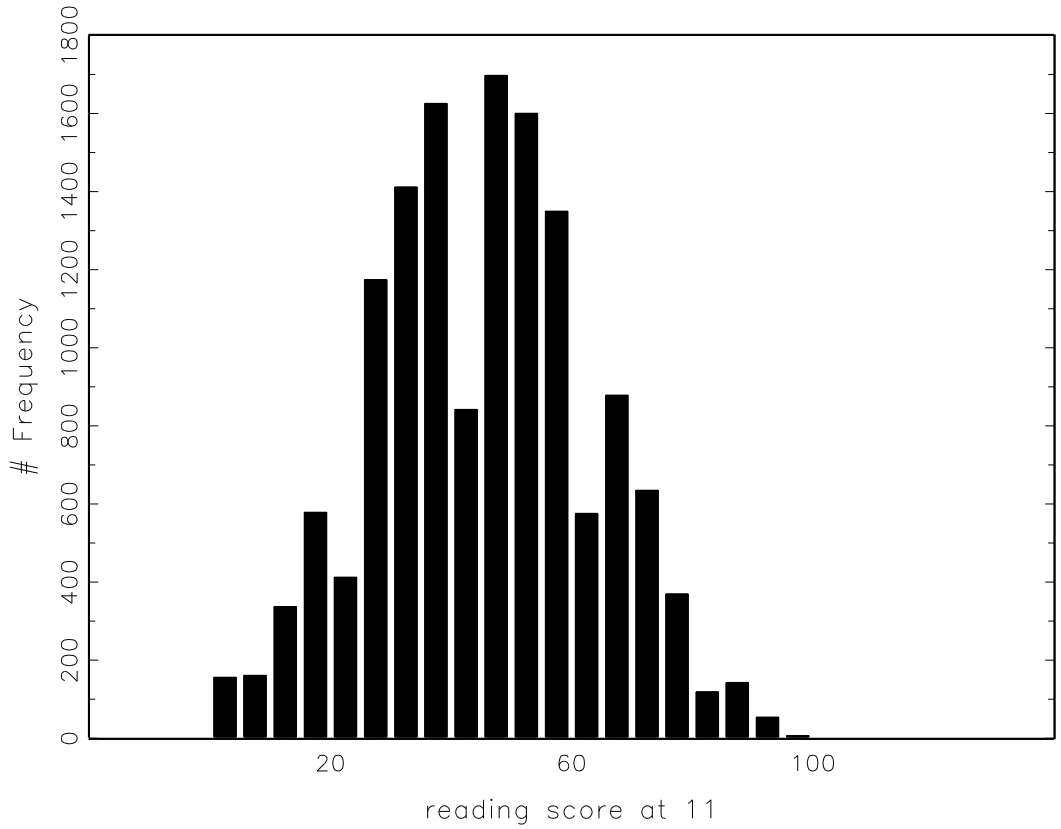


FIGURE 5

Maths Score at 11

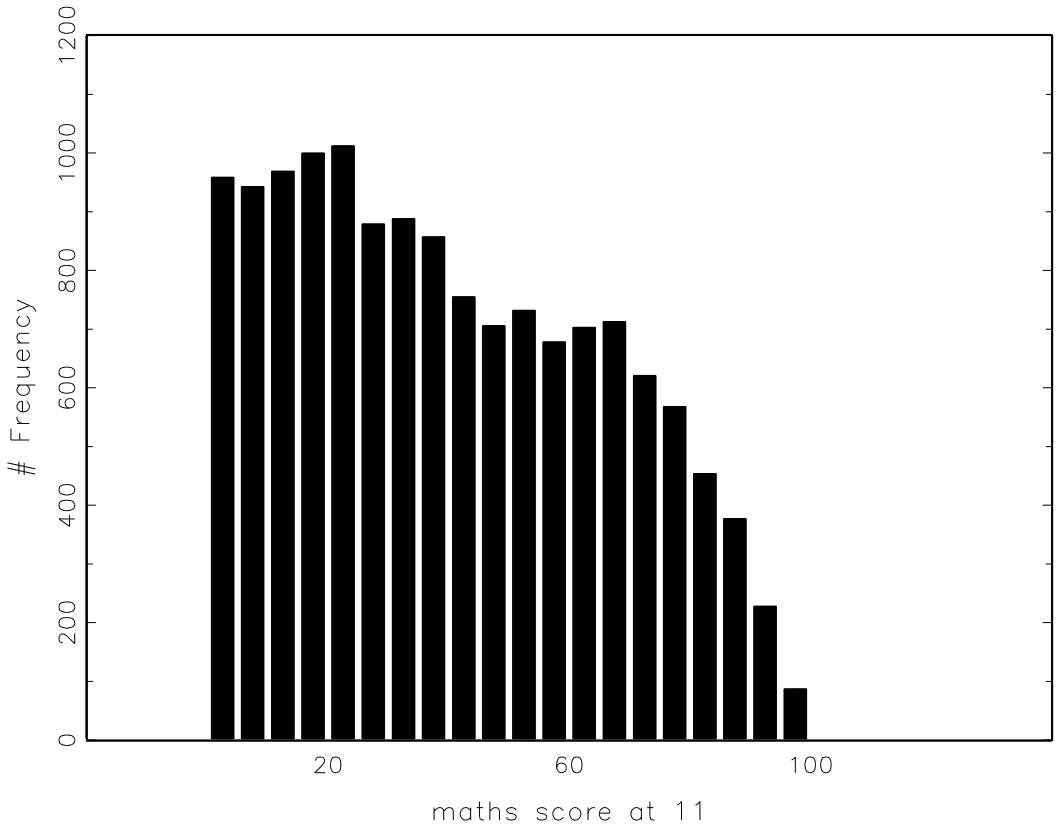
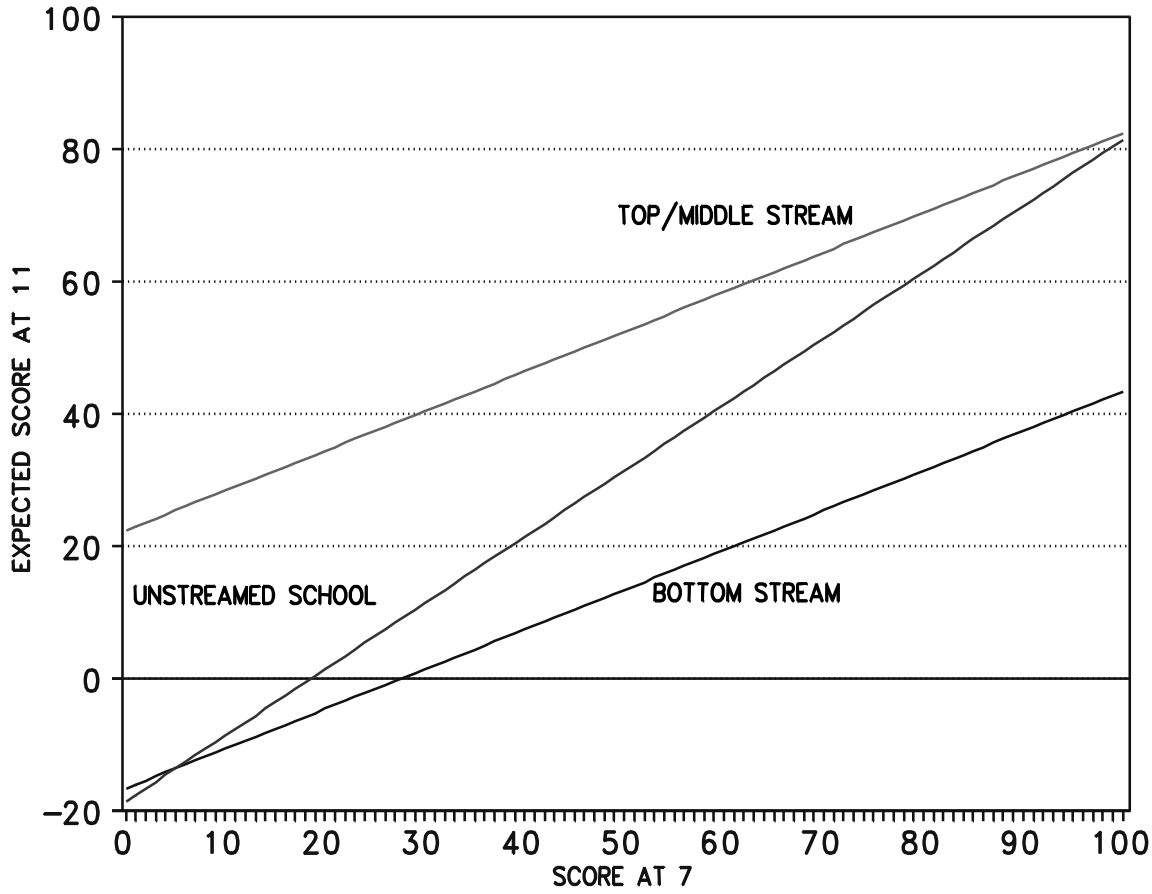


FIGURE 6

Expected Maths Score at 11 as a Function of Score at 7
(from regressions in Table 3)



REFERENCES

Altonji, J.G. and Dunn, T.A. (1995) 'The Effects of School and Family Characteristics on the Return to Education' NBER Working Paper No. 5072.

Becker, G.S. (1981) *A Treatise on the Family* Cambridge, MA: Harvard University Press.

Betts, J.R. (1995) 'Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth' Review of Economics and Statistics, 77(2), pp.231-50.

Borjas, G.J. (1992) 'Ethnic Capital and Intergenerational Mobility', Quarterly Journal of Economics, Vol.107 pp.123-150.

Card, D. and Krueger, A. (1992) 'Does School Quality Matter?', Journal of Political Economy, Vol.100, No.1, pp.1-40.

Coleman, J.S. *et al* (1966) *Equality of Educational Opportunity*, Washington D.C., US GPO.

Evans, W.N., Oates, W.E. and Schwab, R.M. (1992) 'Measuring Peer Group Effects: A Study of Teenage Behaviour', Journal of Political Economy, Vol.100, No.5.

Hanushek, E.A. (1986) 'The Economics of Schooling: Production and Efficiency in Public Schools', Journal of Economic Literature Vol.XXIV, pp.1141-1177.

Harbison, R.W. and Hanushek, E.A. (1992) *Educational performance of the poor: lessons from rural north-east Brazil*, Oxford University Press for The World Bank.

Henderson,V. Mieszkowski, P. and Sauvageau,Y. (1978) *Peer Group Effects and Educational Production Functions*, Ottawa,Canada: Economic Council of Canada.

Robertson, D. and Symons, J. (1989) 'The Occupational Choice of British Children', Economic Journal.

Robertson, D. and Symons, J. (1996) 'Self-Selection in the State School System', forthcoming Centre for Economic Performance discussion paper.

Summers, A.A. and Wolfe, B.L. (1977) 'Do Schools Make a Difference?' American Economic Review Vol.67, No.4, pp.639-52.