SCHOOL QUALITY AND THE LONGER-TERM EFFECTS OF HEAD START

Janet Currie
Duncan Thomas

Working Paper **6362**

NBER WORKING PAPER SERIES

SCHOOL QUALITY AND THE LONGER-TERM EFFECTS OF HEAD START

Janet Currie
Duncan Thomas

Working Paper 6362
http://www.nber.org/papers/w6362

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1998

We thank Larry Katz, Steve Rivkin, Seth Sanders and participants in the NBER's Program on Children for helpful comments. Support from the Alfred P. Sloan Foundation, the National Science Foundation grant #SBR-9512670 and the National Institute of Child Health and Human Development grant R01-HD3101A2 is gratefully acknowledged. We are solely responsible for the views expressed. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

School Quality and the Longer-Term Effects
of Head Start
Janet Currie and Duncan Thomas
NBER Working Paper No. 6362
January 1998
JEL Nos. I21, I38

ABSTRACT

Recent research on Head Start, an enriched preschool program for poor children, has shown that effects on test scores "fade out" more quickly for black children than for white children. This paper uses data from the 1988 wave of the National Educational Longitudinal Survey to show that black children who attended Head Start go on to attend schools of worse "quality" than other black children, in the sense that they attend schools in which most children have worse test scores. We do not see any similar pattern among white children, indicating that on average, white Head Start children attend schools similar to those attended by other white children. Moreover, when we stratify by school type, we find that gaps in test scores between Head Start and other children are very similar for blacks and whites. These patterns suggest that the effects of Head Start may fade out more rapidly among black students than among whites, at least in part because black Head Start children are more likely to subsequently attend bad schools.

Janet Currie
Department of Economics
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90095
and NBER
currie@econ.sscnet.ucla.edu

Duncan Thomas
Department of Economics
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90095
dt@ucla.edu

Early intervention programs for preschoolers have been suggested as a means of improving the educational attainment of disadvantaged children.[1] Children who lag behind their peers when they begin school often fall further behind -- they are more likely to repeat grades, and to eventually drop out of school. Quality preschool programs can help prepare children for school by providing cognitive stimulation, as well as health and nutritional benefits.

There are, however, large disparities in participation in these programs: in 1993, 80% of children in families earning more than $75,000 per year attended a preschool, compared to 45% of children in families with annual incomes less than $30,000. Moreover, programs attended by lower income children tend to be of lower quality. They typically have high child/staff ratios, more teacher turnover, and teachers with less training; the preschools attended by lower income children are also less likely to have formal curricula or guidelines for what students should learn, and are less likely to provide such services as health screenings (U.S. General Accounting Office, 1995).

One of the goals of the Head Start preschool program is to redress these inequities. It is a federal-local matching grant program intended to improve the skills of poor children so that they can begin schooling on an equal footing with their more advantaged peers. While program guidelines require that 90% of participants be from families below the poverty line, this constraint is seldom binding in practice. In 1992, for example, 95% of children served were poor, (U.S. Dept. of Health and Human Services, 1993).

Begun in 1964 as part of the "War on Poverty", the program now serves over 700,000 children in predominately part-day programs, at a cost of approximately $4,000 per child, per

[1] Part of the appeal of early intervention is that interventions aimed at improving the skills of teens and young adults generally have limited effects (c.f. Grossman, 1992; Lalonde, 1995; U.S. Dept. of Labor, 1995).

year (U.S. House of Representatives, 1993). This represents roughly 30% of eligible 3 to 5 year olds. By way of comparison, the average family with an employed mother spent a total of approximately $3,000 on child care in 1991, and poorer families spent even less (Casper, Hawkins and O'Connell, 1994). In order to serve greater numbers of poor children, both Democratic and Republican administrations have increased Head Start funding annually since 1989. In addition, many state governments began funding similar programs in the early 90s (Smith, Fairchild, and Groginsky, 1995).

Children who participate in a Head Start program reap many benefits during the program and immediately after completion. These benefits include improvements in future school attendance and a reduction in the probability that a child is placed in remedial education (McKey *et al.*, 1985). There is evidence suggesting that "Cadillac" interventions such as the famous Perry Preschool Program (which was funded at twice the level of regular Head Start programs) have lasting effects on a broader range of outcomes including high school completion and the avoidance of crime and teen pregnancy. When more typical Head Start programs have been evaluated, however, the evidence for long-term benefits has been less clear. For example, several studies of test scores report that benefits that emerge upon completion of the program have all but faded away three or four years later (Barnett, 1992; Berrueta-Clement *et al.*,1984). Zigler and Meunchow (1992) point out that it may be unrealistic to expect a relatively short-term intervention, like Head Start, to have long lasting effects on cognitive achievement and academic success.

Recent work, however, suggests that the rate at which Head Start benefits fade out differs among sub-populations and that there may, in fact, be relatively long-lasting gains for some children. Using a national sample of children, Currie and Thomas (1995) show that at age 6,

children who went to Head Start score better on vocabulary and reading tests[2] than their siblings who did not attend preschool. These initial gains are essentially *identical* for black and white children. However, among blacks, the beneficial effects associated with Head Start fade out relatively rapidly whereas among white children the benefits persist well into adolescence. Since most of the earlier evaluations of Head Start focussed on relatively small samples which consisted largely of minority children, earlier work had not uncovered the longer-term benefits for white children. Currie and Thomas' estimates are reproduced in Appendix Table 1.

There are several plausible explanations of differential fadeout. First, it is possible that Head Start programs do not serve black children as well as white children, perhaps because of differences in program quality. However, the fact that initial gains are identical for black and white children casts doubt on this explanation and suggests that other factors must play a role. A second hypothesis is that it is the child's experiences *after* leaving the program affect the ultimate gains that are obtained from Head Start. These experiences reflect resources available at home, in the neighborhood, and at school.[3]

Given continuing concerns about racial segregation in schooling, and recent evidence about the importance of school quality in affecting economic outcomes (c.f. Boozer *et al.*, 1992; Card and Krueger, 1992, 1994), we focus on the role of race-specific differences in school quality

---

[2]They use the Peabody Picture Vocabulary Test (PPVT), and the Peabody Individual Achievement Test (PIAT) of Reading Recognition in the National Longitudinal Survey of Youth, Child Assessments.

[3] A third possibility is that the long term gains differ because different sorts of children are selected into the program. As discussed in Currie and Thomas (1995), and demonstrated in a different context below, whites in Head Start appear to more negatively selected relative to other whites than blacks in Head Start are relative to other blacks. Thus, if the reason for differential fade out has to do with differences in selection, Head Start would have to be stunningly successful for whites but of much less value to blacks, even at the time of completion of the program.

as a possible explanation for different patterns of fade-out among black and white Head Start children.[4] Specifically, we start from the premise that the initial positive effects of Head Start may be undermined if Head Start children are subsequently exposed to inferior schools. We noted above that the beneficial effects of Head Start fade out more quickly for black students than for white students. This observation suggests a testable hypothesis--if it is subsequent school quality that determines the longer-term impact of Head Start, and if school quality has similar effects on blacks and whites, then the gap in school quality between Head Start and other students must be greater for blacks than for whites. We test this proposition using a sample of 8th graders from the National Educational Longitudinal Study of 1988 (NELS).

The analysis begins with a model that compares children who attended Head Start to other children of the same race, and shows that in 8th grade Head Start children are more likely to live in low income families, to have parents who have not graduated from high school, and to live in single-parent households. Using average scores on standardized tests taken by children in the same school as an indicator of the "quality" of the school (c.f. Hanushek, 1986), we show that children who attended Head Start are likely to attend schools of lower quality in the 8th grade. A similar result is reported by Lee and Loeb (1995) who show that the schools attended by Head Start children are of worse quality in several observable dimensions than schools attended by other children.[5] Since Head Start children are more likely to be poor than other children, this

---

[4] Of course, differences in school quality can reflect family or neighborhood characteristics that are themselves associated with variation in child outcomes, peer effects, or disparities in resources and programming. Interpreting the differences in school quality uncovered below remains an important avenue for future research.

[5] Specifically, they construct 5 indexes of quality: Average school SES, average achievement, safety, teacher-student relations, and academic climate. They then estimate OLS models of whether or not the student went to Head Start which include demographic variables and the quality measures entered one at a time. They find in this framework that Head Start children attend schools that are worse than those attended by other children in each of the 5 quality dimensions.

finding is not surprising.

What we find however, is that there are sharp differences between races: Among black 8th graders, those who attended Head Start are enrolled in schools of significantly lower quality than those who did not. Among whites, the quality gap is small in magnitude and not statistically significant. This result is consistent with the hypothesis that faster fade out among black Head Start children is related to the quality of the schools these children later attend.

Our next task is to ask whether the differences in school quality that we observe matter. Children who attended Head Start have significantly worse scores than their peers on tests of reading, mathematics, history, and science. A good part of this gap is explained by socio-economic differences (since Head Start children tend to be poor). But a significant part of the gap can be explained by school quality differentials (as measured by average test scores of the other children in the school) particularly among blacks.

Since this average test score measure is unlikely to capture all quality differences across schools, we also look within schools and compare the test scores of children who attended Head Start with those of other children in the same 8th grade school. In these regressions including school fixed effects, the gap between Head Start children and other children is virtually eliminated for blacks. But among whites, the inclusion of school fixed effects has no effect on the estimated differential between Head Start children and other children. Thus, black Head Start children appear to attend schools in which most students do poorly, whereas this is not true for whites.

We explore the relationship between Head Start and future school quality in further detail by adding interactions between the average test scores for other children in the same school and each child's Head Start status to models that also include school fixed effects. The estimates

indicate that Head Start children benefit from improved school quality but less than other children. Thus, white Head Start children, who attend schools of average "quality", perform better than they would in a bad school but more poorly than their peers; in contrast, the average black Head Start child goes on to attend a school with average test scores at least one standard deviation below the mean for all schools. These children and their peers within the school all perform relatively poorly on the standardized tests.

In summary, we show that black Head Start children are systematically more likely than other black children to subsequently attend schools of poor quality, in the sense that they attend schools in which all children tend to perform poorly. Moreover, among black children, some of this differential in quality can be shown to correspond to observable differences in average test scores between schools. These results suggest that improvements in the quality of schools attended by black Head Start children might help to safeguard the initial gains in cognitive achievement fostered by the Head Start program.

1. Methods

The eighth grade outcomes we observe depend on child and family characteristics, initial human capital investments made while the child is in the Head Start program and earlier, and subsequent investments in the child's human capital which will vary with the "quality" of the child's school.

Specifically,

$$(1)\ \text{Outcome}_{iT} = b_1\Sigma^{T}_{t=0}X_{it} + b_2\text{Hdst}_i + b_3f(T)\text{Hdst}_i + b_4\Sigma^{T}_{t=0}Q_{it} + e_{it}.$$

The first term on the right hand side captures the influence of all the time invariant and time varying characteristics of children and families that affect outcomes, from "time zero" up to

capital T, the age at which the outcome is measured. $Hdst_i$ is a dummy variable equal to one if a child attended Head Start. The coefficient b_2 is the estimated initial effect of Head Start on child outcomes. The next term captures the idea that the human capital acquired by the child at Head Start will depreciate over time, though not necessarily at a constant rate. Thus, f(T) is some function of the time that has elapsed, while b_3 is expected to be negative. Finally, outcomes will depend on the quality of schooling the child has been exposed to, the sum of the Q_{it}. If all the relevant characteristics of child and family background and school quality are included in the model, then the last term will be an error that is uncorrelated with the other variables included in the model.[6]

Equation (1) cannot be estimated using the NELS data, since the NELS begins with 8th grade children and subsequent waves measure children two and four years later. If fadeout occurs by the 3rd or 4th grade, then it will already have taken place by the time the children are first measured. Hence, we would not expect to be able to identify fadeout by looking at white children in grades 8, 10, and 12. Thus, we cannot identify $b_3f(T)$, although we can (and will) address the selection issue of whether children who attended Head Start went on to attend schools of inferior quality.

Due to data limitations, we are also constrained to estimate a version of (1) that omits X_{i0} to $X_{i,T-1}$, and Q_{i0} to $Q_{i,T-1}$. Even if we were able to observe past characteristics of children, families, and schools, it is unlikely that we would be able to adequately measure all of the characteristics that affect outcomes, selection into the Head Start program, and school choice. Thus, omitted variables are likely to bias our estimates of the effects of Head Start and school

[6] In principal, interactions between X variables and Q can also be included, but these have been suppressed from (1) in order to simplify the discussion.

quality.

It is however, possible to use prior knowledge to sign these biases. First, Haskins (1989) and Lee *et al.* (1990) find that given a limited number of places, program administrators tend to choose the most disadvantaged eligible children to participate in the Head Start program. Second, Head Start children are overwhelmingly poor, and we can control only for current income, not past income. Thus, past poverty is an omitted variable that is correlated both with Head Start participation and with poorer child outcomes (Korenman *et al.*, 1995). Finally, Currie and Thomas (1995) show that unless both observable and unobservable fixed characteristics of families are controlled for, the estimated effects of Head Start on test scores and schooling attainment are negative. These considerations all suggest that OLS estimates of the effects of Head Start will be biased downwards.

Note that if Head Start children attend schools of worse quality, then omitting the Q_{it} from (1) will also bias the estimated effects of Head Start downwards. Thus, while given the available data we can obtain only biased estimates of the effects of Head Start, *the difference* between estimates of the effects of Head Start obtained from models that include measures of school quality and those obtained from models that do not, can shed light on the extent to which omitting school quality changes the estimated returns to Head Start. And to the extent that current school quality is correlated with the unobserved variables (e.g. with past school quality and with family background), including it will reduce the bias from omitting these variables.

As discussed above, we measure school quality at time T in two different ways. First, we include a measure of the average test scores of the other children in the school. As discussed above, higher scores may reflect family or neighborhood characteristics or peer effects in addition to any effect of schools *per se*. It will not be possible to disentangle the importance of each

factor in this study. Second, we estimate models that include a fixed effect for each school. The fixed effect captures all observed and unobserved characteristics of schools, including school quality, and allows us to compare Head Start children to other children within the same school. In some models, we interact mean test scores with the child's Head Start status in addition to including school fixed effects. These models allow us to ask how Head Start children perform within schools of different qualities.

2. Data

The data for this study are taken from the National Educational Longitudinal Study of 1988. This is one of the few data sets available that combines information about past participation in Head Start with measures of school quality for large numbers of black and white children, although as discussed above, the fact that it begins in 8th grade limits the extent to which it can be used to study the fadeout of Head Start effects. Still, one benefit of using the NELS is that Murnane, Willett, and Levy, (1995) show that test scores measured in the 8th grade are significant predictors of future labor market outcomes.

A two-stage probability design was used to select a sample of students and schools. In the first stage, 1,734 schools were selected, while in the second, 26,435 students were randomly selected from these schools.[7] We excluded 700 students from a special, supplemental "hearing

[7] In principal, it is possible to weight in order to obtain a nationally representative sample. When we apply the weights available in the NELS, we obtain results that are, except where specifically noted, qualitatively similar to those discussed below. However, we have several reservations about the sampling weights. First, the weights vary from 2 to over 800, indicating that a few individuals who were unlikely to be surveyed have a lot of influence in the weighted regressions. Excluding the 1% of the sample with the highest weights produces point estimates that are in some cases quite different than those estimated using the weights with the full sample. Second, the major function of the weights is to control for the probability that a particular school was included in the sample. Since this probability should be subsumed in the school fixed effect, weighting should have little impact on the estimated coefficients in the fixed effects models. We have therefore decided to present unweighted statistics. The reader should keep in mind then that our sample, while large, is not necessarily nationally representative.

impaired" sample, and excluded a further 1,632 students whose parents did not return their questionnaires. Excluding children with missing test scores leaves 2531 black students and 14343 white students available for analysis.[8] Since the focus of our analysis is on racial differences in the way that Head Start children are selected into schools of differing quality levels, we estimate all our models separately for whites and blacks. The mean numbers of students per school included in the analyses of blacks and whites were 5 and 15, respectively.

Although the student is the base unit of analysis in the survey, the NELS surveyed parents, teachers, and school administrators as well as students. Information about student test scores is taken from the student's questionnaire; information about family background and the child's early educational experiences comes from the parent's questionnaire (which was completed by the parent "most knowledgeable" about the child's education); information about the classroom environment was provided by teachers in two of the four test areas (reading, mathematics, science, history); and information about the school was provided by the principal or another school administrator designated by the principal. Note that since a couple of teachers were surveyed for each child, we take the means of each child's teachers' responses. Further information about the sources and the coding of all our variables is shown in the data appendix. More information about the survey is available in Ingels *et al.* (1989).

Means of the variables used in our analyses are shown in Table 1, by race. In order to highlight potential differences between Head Start children and other children, the means are further stratified by the type of preschool the child attended: Head Start, another preschool, or

[8] The largest groups excluded from the current analysis are hispanics, asians, and "other race". In an earlier version of this paper, we presented separate estimates for the sample of 2676 hispanic students. We found that these students were more similar to whites than to blacks, in that controlling for school fixed effects eliminated little of the differences between Head Start and other hispanic children that were found in OLS.

no preschool. The column marked "Head Start" includes all children who ever attended Head Start, while the column marked "Other Preschool" includes children who only attended non Head Start preschools.

The first panel of Table 1 shows that children who attend other preschools typically come from more advantaged backgrounds than Head Start children: their mothers are more likely to be highschool graduates or to have attended college, their families are less likely to have single heads, and their family incomes are much higher. These differences are particularly pronounced among whites, suggesting that white Head Start children are more negatively selected in terms of family background than black Head Start children. In terms of child characteristics, Table 1 shows that Head Start children are less likely to be first born than children who attended other preschools, and they are also less likely to have attended other non-preschool day care programs.[9]

In general, the differences between children who attended Head Start and those who attended no preschool are less pronounced. However, even relative to these children, Head Start children have lower family incomes and are more likely to be from single-headed households. They are also more likely to have attended other day care programs.

Panel B of Table 1 shows that these differences in background are reflected in differences in child outcomes. The NELS tested students in four areas: reading, mathematics, science, and history.[10] The test scores are expressed in percentage terms so that in principal, scores may

---

[9] Although there is likely to be some noise in the data regarding the distinction between preschools and day care programs, the estimates shown below suggest that the two differ substantively. Children who attended programs described as preschools have significantly higher test scores than those who attended day care programs.

[10] The reading test (21 items, 21 minutes) consisted of 5 short passages followed by comprehension and interpretation questions. The mathematics test (40 items, 30 minutes) consisted of quantitative comparisons and other questions. The science test (25 items, 20 minutes) assessed scientific knowledge and reasoning. The history test (30 items, 14 minutes) asked about U.S. history, civics, and government.

vary from 0 to 100. Each score has a standard deviation of 15 to 20 points. Hence, the table shows that children who attended Head Start typically lag more than 1/2 a standard deviation behind those who attended other preschools. White Head Start children also tend to perform more poorly than children who attended no preschool, though these differences are less pronounced. Table 2 explores the relationship between Head Start status and current school quality. In the models shown, the school quality relevant to each child is measured using the average reading score taken over all the other children in the school. Very similar results were obtained using math, history, or science scores. For ease of interpretation, the scores are expressed as Z-scores, so that a school one standard deviation below the mean would have a score of -1 (call this a poor school), while a school one standard deviation above the mean would have a score of 1 (call this a good school).[11] Note that the standard deviation in school mean scores varies between 10 and 12 points.

Table 2 shows that among blacks, the probability that a given child attended Head Start rises as the quality of the current school declines, conditional on observable family background characteristics. No similar relationship exists for whites. Thus, Table 2 suggests that black Head Start children are concentrated in poor schools, relative to other black children, while the same is not true for whites.

The rest of Table 2 confirms what we saw in Table 1 but in a multi-variate context: Head Start children are more likely to be in poor families, and to have less educated parents. The coefficients on income and education are greater for blacks than for whites, which may reflect larger differences in these variables between Head Start and non-Head Start children among

[11] We do not use the mean test score taken over all the children in the school because this measure would be mechanically correlated with the child's own score. However, for the purposes of computing z-scores, we subtract out the mean of all the school's mean scores, and divide by the standard deviation of school mean scores.

whites.

3. Results

Initial OLS estimates of our model of test scores are shown in Panel 1 of Table 3. For clarity's sake, the discussion in Section 1 stressed Head Start enrollment, but there are a number of other indicators of early childhood educational experiences available in the NELS. Hence, in addition to $Hdst_i$ our empirical analog to (1) also includes controls for whether the child attended another preschool or a day care program. The coefficients on the X_{it} variables included in the regression model (measures of parent's education, income, and family structure as well as children's demographic characteristics similar to those shown in Table 2) have been suppressed in order to focus on the main results.

In view of the discussion in Sections 1 and 2, it is not surprising to find that the estimated effects of Head Start are negative[12], and somewhat larger for white children in absolute value (since these children appear to be most negatively selected in terms of parent's income and education). However, it must be recognized that when measured in 8th grade, the estimated effects of Head Start in these OLS models are quite small (on the order of 1/5 of a standard deviation) and not significantly different for blacks and whites.[13] Thus, much of the raw deficit between Head Start and other children appears to be accounted for by observable differences in family background. If unobservable differences between Head Start and other children could also

---

[12] The only result that is qualitatively different when we weight, is that the Head Start coefficient on the math score for blacks becomes statistically insignificant.

[13] It is possible that different mechanisms underlie these small effects: If black Head Start students are not very negatively selected and achieve large gains from Head Start which then fade out quickly, while white students are much more negatively selected and achieve large gains which fade out only slowly, we might see that black and white Head Start students would end up with similar deficits relative to other black and white students by the time they got to 8th grade. Unfortunately, we are not aware of any data that would allow a direct test of this hypothesis.

be controlled, then we might find positive effects of Head Start as in Currie and Thomas (1995).

In contrast to the effects of Head Start, the estimated effects of preschool are positive for blacks and whites -- a finding which presumably reflects both human capital investments and positive selection into these programs. The effects of attending other day care programs are generally negative, and significant for all four test scores among whites. These findings may reflect either the poor quality of many of these programs (discussed above), or negative selection.

a) Estimates from models including average school test scores

The second panel of Table 3 shows estimates that control for school quality measured using the average test scores. In these models we use average reading scores in the model for reading scores, average math scores in the model for math scores, etc. Table 3 shows that these average scores are highly significant determinants of individual test scores. Children who attend schools in which other children are doing well tend to have higher scores themselves -- a one standard deviation change in the scores of the other children is associated with a increase in individual scores of 4 to 6 percentage points.

Among blacks, the inclusion of this school quality measure reduces the size of the estimated Head Start effect between 27 and 41%, and the effects of Head Start become statistically insignificant in the equations for math and science scores. Among whites however, the estimated Head Start effects are reduced by only 5 to 10%, and all remain highly statistically significant. Thus, these estimates suggest that among blacks, conditioning on this measure of school quality reduces the gap in test scores between Head Start and other children considerably. Thus, black Head Start children attend schools of worse quality than other black children, while the same is not true among whites.

b) Estimates from models including school fixed effects

As discussed above, we do not expect our average scores to capture all of the relevant differences in quality between schools. Hence, the third panel of Table 3 shows estimates that control for school quality by including school-specific fixed effects. Among black children, the results of controlling for school quality in this way are striking: the coefficients on $Hdst_j$ all fall by 50 to 80% and become statistically insignificant. Including school fixed effects has little impact on the estimated coefficients on other preschools, or day care programs. Moreover, although they are not shown, it is interesting to note that the coefficients on the other family background variables included in the model (measures of parent's education, income, and family structure) hardly move when the school fixed effects are added. Among white children, controlling for school fixed effects has virtually no effect on *any* of the coefficients included in the model.

These results show that relative to other black children, black children who went to Head Start subsequently attend schools of low quality in the sense that the black children in these schools have poorer outcomes. The same is not true for white Head Start children. Thus, the results suggest that differential fadeout in the effects of Head Start could be caused by greater subsequent exposure to poor schools among black Head Start children.

Since these models were estimated separately for whites and blacks, the school fixed effects are not constrained to be equal for the two groups. In fact, if there were segregation *within* schools, then the same school might be of higher "quality" for whites than for blacks. Alternatively, if the quality of education was similar for whites and blacks within schools, then one might expect to obtain similar results whether or not one constrained the fixed effects to be equal across racial groups. That is, it would not matter whether the fixed effects were calculated

using the sample of whites, the sample of blacks, or using all the children in the school. In fact, models that constrain the fixed effects to be equal across the two groups produce estimates very similar to those shown in Table 3, suggesting that within schools, whites and blacks receive educations of similar quality.

c) Estimates from models including interactions of mean scores and preschool experience

A more direct way to examine the relationship between Head Start and subsequent school quality is to interact the dummy variable for Head Start attendance with an indicator of quality, while continuing to control for other unobserved characteristics of the school by including fixed effects. In this section we use the mean scores of all the other students in the school as a measure of quality. The idea is that children are likely to get higher scores in schools where the other children test better, because parents who value education prefer these schools, because of peer effects, or because the other children's high scores reflect direct school inputs into the educational process.

Mean scores are interacted with the Head Start dummy, as well as with the dummies for preschool and other day care. Aside from these three interactions, the specification is identical to the fixed effects specification discussed above. We use the mean reading score in the model for reading, the mean math score in the model for math, etc. All scores are expressed in Z-scores.

The results are shown in Table 4.[14] The interactions of test scores with Head Start and preschool are highly statistically significant, but somewhat difficult to interpret. Thus, in the second part of the table, we show the implied effects of attending a school with scores one

[14] There are slightly fewer observations in these tables because schools with only one black or white child could not be used.

standard deviation below the mean, a school with average scores, and a school with scores one standard deviation above the mean. For example, a black Head Start child who attended a poor school would have a reading score that was 1 point less (since -4.53 + 3.53 equals 1) than the school's average score, while if the same child attended a good school, his/her score would be 8 points less than the school's average score (since -4.53 - 3.53 equals 8.06).

There are several striking results in this table. First, when stratified by school type, the estimated size of the gaps between the Head Start children and the other children are very similar for blacks and whites. To continue with the example begun above, a white Head Start child who attended a poor school would have a reading score .7 points less than the average for his/her school, while one who attended a good school would score 7.7 points lower than his/her peers. This result is important because it suggests that school quality (as measured using deviations from average test scores) means similar things to blacks and whites.

Second, a comparison of these estimates with the simple fixed effects estimates shown in Table 3 suggests that white Head Start children attend schools of average quality, while black Head Start children attend schools of poor quality. To see this, recall that Table 3 indicated that within schools, white Head Start children scored 3 or 4 points below their peers. Table 4 shows that this is the gap that one would expect in a school of average quality. Table 3 also suggested that there was little difference in test scores between black Head Start children and other children within schools. Table 4 shows that this is what one would expect if the black children attended schools with scores one standard deviation or more below the mean score.

A third result is that although the gap between the Head Start children and the other children grows with school quality, Head Start children do better when they attend better schools. Recall that a good school is one that has scores a standard deviation, or 15 points, above the

mean for all schools, while a bad school is one with scores a standard deviation lower than this mean. Hence, a Head Start child in a poor school would score about 10 points below the mean score in an average school, while a similar child who attended a good school would score 3 to 5 points above the mean in an average school.

The estimates for preschools show greater differences between blacks and whites. For example, among whites, attending preschool is more protective for children who go on to attend poor schools than it is among blacks. However, once again a comparison between Tables 3 and 4 suggests that the black children are more likely to be in poor schools, while the white children are more likely to attend average schools. A result that may merit further study is that other preschools are protective for children in poor schools, have relatively little effect on children in average schools, and actually appear to be associated with worse outcomes for children who subsequently attend good schools.

4. Conclusions

This paper shows that black children who attended Head Start go on to attend schools of worse "quality" than other black children, in the sense that they attend schools in which all children have worse test scores. The same is not true among whites. We also find that within schools of similar quality (in terms of mean test scores), the estimated size of the gaps between the Head Start children and the other children are very similar for blacks and whites. However, our results suggest that white Head Start children go to schools of average quality, while black Head Start children often attend schools with test scores as much as one standard deviation below the mean for all students.

These results suggest that the effects of Head Start may fade out more rapidly among

blacks than whites because black Head Start children are more likely to subsequently attend schools of poor quality. Unfortunately, given the constraints imposed by the available data, we cannot directly document the effects of school quality on fadeout as it is taking place. However, the idea that poor schools can undermine early gains from Head Start is consistent with evidence from Head Start "Follow Through" and from more recent projects which show that continuing enrichment programs into the early grades can help students to retain their initial gains from Head Start (c.f. Rhine, 1981; Reynolds, 1997).

References

Barnett, Steven. "Benefits of Compensatory Preschool Education", Journal of Human Resources, Spring 1992, 27 #2, 279-312.

Boozer, Michael, Alan Krueger, and Sherry Wolkin. "Race and School Quality since *Brown v. Board of Education*", Brookings Papers on Economic Activity: Microeconomics 1992, Washington D.C.: Brookings Institution, 269-338.

Brown, J. Larry and Ernesto Pollitt, "Malnutrition, Poverty, and Intellectual Development", Scientific American, February 1996, 38-43.

Card, David and Alan Krueger. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States", Journal of Political Economy, #100, Feb. 1992, 1-40.

Card, David and Alan Krueger. "The Economic Returns to School Quality: A Partial Survey", Working Paper #334, Industrial Relations Section, Princeton University, 1994.

Casper, Lynne, Mary Hawkins and Martin O'Connell. "Who's Minding the Kids?", Washington D.C.: U.S. Dept. of Commerce, P70-36E, June 1994.

Currie, Janet and Duncan Thomas. "Does Head Start Make a Difference", American Economic Review, June, 1995.

Currie, Janet and Duncan Thomas. "Nature vs. Nurture? The Bell Curve and Children's Cognitive Achievement", NBER Working Paper #5240, August 1995b.

Currie and Thomas, 1996. "Does Head Start Help Hispanic Children", mimeo, National Bureau of Economic Research, Cambridge MA, Sept. 1996.

Ehrenberg, Ronald and Dominic Brewer. "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond?" Economics of Education Review, #13, March 1994, 1-17.

Grossman, Jean Baldwin. "The Evaluation of Youth Employment and Training Programs", Princeton NJ: Public Private Ventures, November 1992.

Hanushek, Eric. "The Economics of Schooling." Journal of Economic Literature, September 1986 24, pp. 1141-77.

Haskins, Ronald. "Beyond Metaphor: The Efficacy of Early Childhood Education", American Psychologist, February 1989, 44(2), pp. 274-82.

Hayes, Cheryl; John Palmer and Martha Zaslow. Who Cares for America's Children: Child Care

Policy for the 1990s, Washington D.C.: National Academy Press, 1990.

Ingels, Steven, Sameer Abraham, Kenneth Rasinski *et al.* National Educational Longitudinal Study User's Manual, Washington D.C.: U.S. Department of Education National Center for Education, December 1989.

Korenman, Sanford, Jane Miller, and J. Sjaastad. "Long-Term Poverty and Child Development in the United States -- Results from the NLSY", Children and Youth Services Review, 1995, v17 n1-2:127-155.

Lalonde, Robert. "The Promise of Public Sector Sponsored Training Programs", Journal of Economic Perspectives, 9 #2, Spring 1995, 149-168.

Lee, Valerie and Susanna Loeb. "Where do Head Start Attendees End Up? One Reason Why Preschool Effects Fade Out", Educational Evaluation and Policy Analysis, Spring 1995, v17 #1, 62-82.

Lee, Valerie, Jeanne Brooks-Gunn, Elizabeth Schnur, and Fong-Ruey Liaw. "Are Head Start Effects Sustained? A Longitudinal Follow-Up Comparison of Disadvantaged Children Attending Head Start, No Preschool, and Other Preschool Programs", Child Development, April 1990, 61 #2, 495-507.

Loeb, Susanna and John Bound. "The Effect of Measured School Inputs on Academic Achievement: Evidence from the 1920s, 1930s, and 1940s Birth Cohorts", NBER Working Paper #5331, November 1995.

McKey, Ruth, Larry Condell, Harriet Ganson, *et al.* The Impact of Head Start on Children, Families and Communities: Final Report of the Head Start Evaluation, Synthesis, and Utilization Project, Washington, D.C.: CSR Inc., 1985.

Murnane, Richard, John Willett, and Frank Levy. "The Growing Importance of Cognitive Skills in Wage Determination", Review of Economics and Statistics, May 1995.

Reynolds, Arthur. "Early Childhood Intervention and Adolescent Social Competence: Findings from the Chicago Child-Parent Centers", mimeo, University of Wisconsin-Madison, 1997.

Rhine, W. Ray. Making Schools More Effective: New Directions From Follow Through, New York: Academic Press, 1981.

Smith, Shelley, Mary Fairchild and Scott Groginsky. "Early Childhood Care and Education: An Investment that Works", Washington D.C.: National Conference of State Legislatures, February 1995.

U.S. Department of Health and Human Services, Head Start Program Performance Standards, Washington D.C.: Government Printing Office, 45-CFR 1304, June 1992.

U.S. Department of Health and Human Services, Creating a 21st Century Head Start: Final Report of the Advisory Committee on Head Start Quality and Expansion, Washington D.C.: U.S. Government Printing Office, 1993.

U.S. Department of Labor. What's Working (and What's Not): A Summary of Research on the Impact of Employment and Training Programs, Washington D.C.: Office of the Chief Economist, 1995.

U.S. General Accounting Office, Early Childhood Centers: Services to Prepare Children for School Often Limited, GAO/HEHS-95-21, March 1995.

U.S. General Accounting Office, Child Care: Promoting Quality in Family Child Care, GAO/HEHS-95-36, Dec. 7, 1994.

U.S. House of Representatives, Committee on Ways and Means, 1993 Green Book: Background Material and Data on Programs Within the Jurisdiction of the Committee on Ways and Means, WMCP-103-27, Washington D.C.: Government Printing Office, 1993.

Zigler, E.F. and S. Meunchow. Head Start: The Inside Story of American's Most Successful Educational Experiment, New York: Basic Books, 1992.

Table 1: Means of Key Variables by Race, Ethnicity and Type of Preschool

	Black			White		
	No Pre.	Hdst Start	Other Pre.	No Pre.	Hdst Start	Other Pre.
A: Child and Family Characteristics						
# Obs.	1026	873	632	6822	678	6843
Attended Day Care	.17	.36	.47	.10	.18	.25
First Born	.34	.41	.44	.36	.38	.43
# of Siblings	2.69	2.77	2.14	2.26	2.47	1.84
	(.06)	(.06)	(.06)	(.02)	(.06)	(.02)
Family Income (1000s)	22.42	18.15	37.26	35.67	26.15	63.40
	(.71)	(.62)	(1.26)	(.34)	(.98)	(.63)
Mother HS, GED, or Vocational Degree	.42	.39	.32	.51	.45	.32
Mother Any College	.26	.32	.57	.31	.25	.63
Father HS, GED, or Voc. (father present)	.22	.20	.19	.37	.33	.20
Father Any College (father present)	.16	.13	.36	.35	.24	.64
Single Headed HH	.41	.47	.37	.15	.22	.14
Foreign Language in Home	.08	.04	.05	.06	.06	.05
B: Child Outcomes						
Reading Score	46.59	46.03	57.62	62.09	56.65	70.72
	(.61)	(.67)	(.91)	(.26)	(.89)	(.25)
Math Score	39.13	38.64	49.65	55.24	49.75	65.55
	(.54)	(.56)	(.83)	(.25)	(.78)	(.25)
Science Score	41.28	40.81	49.10	55.00	50.49	61.75
	(.45)	(.50)	(.66)	(.21)	(.70)	(.21)
History Score	52.84	52.84	61.34	64.38	59.36	71.59
	(.50)	(.53)	(.70)	(.21)	(.70)	(.20)

Note: Standard errors in parentheses. Column labeled Head Start includes all children who ever attended Head Start, while column labeled "other preschool" includes only children who attended other preschools.

Table 2: Correlates of Past Participation in Head Start

	Black	White
Average School Reading Scores	-.031	-.004
	(3.05)	(1.74)
Child Male	-.006	-.004
	(.33)	(1.25)
Family Income ($1000s)	-1.87	-.026
	(3.91)	(5.37)
Family Income Missing	-.049	-.017
	(1.11)	(1.98)
Foreign Language Spoken at Home	-.090	.007
	(2.30)	(.86)
Child First Born	.048	.006
	(2.23)	(1.57)
Number Siblings	.010	.006
	(1.75)	(4.51)
Mother GED	.042	-.005
	(.67)	(.40)
Mother High School	-.094	-.062
	(2.90)	(8.37)
Mother Vocational Training	-.067	-.056
	(2.00)	(7.00)
Mother 2 Years College	-.039	-.054
	(1.14)	(6.65)
Mother 3 Years College	.011	-.067
	(.28)	(7.14)
Mother 4 Years College	-.210	-.077
	(4.21)	(8.70)
Mother Graduate School	-.180	-.070
	(3.11)	(6.83)
Mother Education Missing	-.092	-.028
	(2.35)	(2.52)
Father GED	.012	.008
	(.13)	(.59)
Father High School	-.028	-.026
	(.68)	(3.32)
Father Vocational Training	-.053	-.019
	(1.15)	(2.18)
Father 2 Years College	-.097	-.031
	(2.02)	(3.62)
Father 3 Years College	-.031	-.036
	(.59)	(3.79)
Father 4 Years College	-.110	-.031
	(1.78)	(3.49)
Father Graduate School	-.110	-.031
	(1.68)	(3.25)
Father Education Missing	-.021	.003

	(.50)	(.30)
No Father	-.007	-.008
	(.21)	(.93)
Intercept	.430	.124
	(9.70)	(13.72)
R-squared	.073	.033
# Observations	2484	14235
Mean of Dependent Variable	.34	.05

Notes: T-statistics in parentheses. Models also included dummy variables equal to one if first born or the number of siblings was missing.

Table 3: Effects of Early Childhood Education on 8th Grade Outcomes

	Black Reading Score	Black Math Score	Black Science Score	Black History Score	White Reading Score	White Math Score	White Science Score	White History Score
Panel 1: OLS								
Head Start	-2.86	-1.94	-2.24	-1.49	-3.48	-3.79	-2.90	-3.45
	(2.82)	(2.18)	(2.98)	(1.84)	(4.07)	(4.69)	(4.20)	(5.03)
Other Preschool	3.54	3.88	2.30	3.00	1.71	2.72	1.61	1.65
	(3.18)	(3.97)	(2.78)	(3.37)	(3.85)	(6.50)	(4.50)	(4.65)
Daycare	-.03	-1.79	-1.01	.23	-1.28	-1.53	-1.25	-1.18
	(.03)	(1.88)	(1.26)	(.26)	(2.55)	(3.23)	(3.10)	(2.95)
# Observations	2531	2532	2517	2516	14342	14340	14343	14283
R-squared	.15	.16	.13	.13	.18	.22	.16	.18
Panel 2: Same as 1 but including Average Test Scores Other Children in School								
Head Start	-1.99	-1.27	-1.63	-.88	-3.29	-3.56	-2.79	-3.08
	(2.02)	(1.49)	(2.26)	(1.12)	(3.90)	(4.52)	(4.13)	(4.64)
Other Preschool	2.81	3.26	1.98	2.56	.93	1.58	1.06	.85
	(2.59)	(3.48)	(2.49)	(2.96	(2.11)	(3.95)	(3.03)	(2.45)
Daycare	-.04	-1.60	-.74	.22	-1.08	-1.18	-1.03	-.92
	(.03)	(1.75	(.95)	(.26)	(2.19)	(2.57)	(2.61)	(2.38)
Average Test Score	5.20	6.03	4.44	4.65	4.43	5.81	4.54	5.08
	(12.39)	(15.47)	(14.44)	(13.27)	(20.13)	(28.15)	(26.02)	(31.19)
# Observations	2525	2527	2512	2511	14330	14327	14331	14270
R-squared	.20	.23	.20	.19	.20	.26	.20	.23
Panel 3: Same as 1 but with School Fixed Effects								
Head Start	-.46	-.57	-.90	-.20	-3.56	-3.90	-3.12	-3.27
	(.42)	(.61)	(1.13)	(.23)	(4.05)	(4.77)	(4.45)	(4.76)
Other Preschool	2.45	3.58	2.15	2.67	1.03	1.67	1.43	1.15
	(2.04)	(3.52)	(2.46)	(2.79)	(2.23)	(3.92)	(3.92)	(3.20)
Daycare	-.38	-1.89	-.79	-.56	-1.20	-1.22	-1.09	-.93
	(.33)	(1.93)	(.93)	(.61)	(2.32)	(2.55)	(2.66)	(2.30)
# Observations	2531	2532	2517	2516	14342	14340	14343	14283
# Schools	493	493	492	492	947	948	946	946
R-squared	.34	.39	.35	.34	.27	.33	.28	.31

Notes: T-statistics in parentheses. All regression models also include a dummy variable for the child's gender, family income; a dummy variable for whether a foreign language was spoken in the home; the number of siblings; dummy variables for whether or not the child is the firstborn, or whether the birth order is missing. Dummy variables for missing daycare, Head Start, preschool, or kindergarten and for missing family income were also included. Parental education is controlled for by including 16 dummy variables (for maternal and paternal completion of a GED, highschool, a vocational certificate, 2 year of college, 3 years of college, 4 years of college, and graduate work). Dummies were also included for missing education, and for whether or not the household was single-headed. Average reading scores included in equations for reading, average math scores included in equations for math, etc.

Table 4: Differential Effects of Early Education Within Schools

	Black Reading Score	Black Math Score	Black Science Score	Black History Score	White Reading Score	White Math Score	White Science Score	White History Score
Panel 1: Coefficient Estimates								
Head Start	-4.53	-3.49	-3.24	-2.85	-4.21	-4.62	-3.49	-3.86
	(3.02)	(2.58)	(2.82)	(2.44)	(4.77)	(5.61)	(5.00)	(5.58)
Head Start x Mean Score	-3.53	-2.41	-1.82	-2.46	-3.51	-4.37	-3.22	-3.26
	(3.65)	(2.60)	(2.52)	(2.96)	(3.44)	(4.73)	(4.05)	(4.60)
Other Preschool	-.30	-.20	-1.23	-.52	1.73	2.22	2.09	1.61
	(.20)	(.15)	(1.05)	(.44)	(3.77)	(5.20)	(5.71)	(4.51)
Preschool x Mean Score	-2.75	-3.82	-3.13	-3.68	-7.54	-6.00	-5.69	-5.20
	(2.77)	(4.23)	(4.25)	(4.39)	(16.51)	(13.87)	(15.66)	(15.29)
Daycare	-1.34	-.88	-1.24	-1.06	-1.19	-1.33	-1.46	-1.11
	(.93)	(.68)	(1.07)	(.91)	(2.26)	(2.71)	(2.18)	(2.72)
Daycare x Mean Score	-1.06	.99	-.44	-.54	-1.52	-.79	-.26	-.23
	(1.09)	(1.08)	(.58)	(.64)	(3.03)	(1.74)	(.66)	(.59)
# Observations	2525	2527	2512	2511	14330	14327	14331	14270
# Schools	488	488	487	487	935	935	934	933
R-squared	.40	.45	.41	.40	.29	.34	.29	.33
Panel 2: Implied Gaps in Poor, Average and Good Schools								
Between Head Start and No Preschool Children								
Poor (Mean Score = -1)	-1.00	-1.08	-1.42	-.39	-.70	-.25	-.27	-.60
Avg. (Mean Score = 0)	-4.53	-3.49	-3.24	-2.85	-4.21	-4.62	-3.49	-3.86
Good (Mean Score = 1)	-8.06	-5.90	-5.06	-5.31	-7.72	-8.99	-6.71	-7.12
Between Preschool and No Preschool Children								
Poor (Mean Score = -1)	2.40	3.62	1.90	3.16	9.27	8.22	7.78	6.81
Avg. (Mean Score = 0)	-.30	-.20	-1.23	-.52	1.73	2.22	2.09	1.61
Good (Mean Score = 1)	-3.05	-4.02	-4.36	-4.20	-5.81	-3.78	-3.60	-3.59

Notes: T-statistics in parentheses. Models similar to those in Panel 2 of Table 2 except for the inclusion of the interactions. We use mean reading scores in the models for reading, mean math scores in the models for math, etc.

Appendix Table 1: Examples of Differential Fadeout from Previous Research

Sample:	Black	White	Black	White
Dependent Variable:	PPVT Score	PPVT Score	PIAT-Reading Score	Piat-Reading Score
Effect of Head Start @ 5 Years Old	6.85	6.88	8.36	6.63
	(3.55)	(2.87)	(2.30)	(2.08)
Loss with Each Additional Year of Age	-1.28	-.19	-1.25	-.59
	(4.13)	(.46)	(2.38)	(1.10)

Notes: This table reports coefficient estimates from models in which the test scores were regressed on a dummy variable for participation in Head Start, as well as an interaction of this variable with the child's age less 5 years. It is the coefficients on these two variables that are shown in the table. Each column represents a separate regression--all models were estimated separately for blacks and whites. T-statistics appear in parentheses. The PPVT is the Peabody Picture Vocabulary Test, while PIAT-Reading refers to the PIAT Reading Recognition test. Scores are expressed in percentile terms. These estimates are based on a sample of siblings drawn from the National Longitudinal Survey of Youths Child-Mother file, waves 1986, 1988, and 1990. In addition to the variables shown, the models included a fixed effect for each mother, as well as controls for whether the child attended another preschool, interactions of other preschools and years of age since age 5, child age, child gender, whether the child was first born, and the household income at the time the child was aged 3. Adding controls for maternal marital and employment status at age 3 does not affect these estimates. The fact that the interaction term is statistically significant in the regressions for blacks but not in the regressions for whites shows that fadeout is concentrated among blacks. See Currie and Thomas (1995) for further details.

Data Appendix: Variable Definitions

This appendix gives the source of the question (i.e. student, parent, teacher, or school questionnaire); the NELS variable name; the question that was asked; and additional notes regarding the way that the variable was coded, if appropriate.

Outcomes

Test Scores (in percent). Student Base Year Questionnaire, bytxrstd, bytxmstd, bytxsstd, bytxhstd.

Early Education

Parent's Questionnaire, byp37a-d. "Did 8th grader attend day care program"?, "Did 8th grader attend nursery/preschool", "Did 8th grader attend Head Start?", "Did 8th grader attend Kindergarten?" Respondents answer yes, no or don't know. Dummy variables for missing values of these variables were defined and included in all regressions.

Child and Family Characteristics

Total Family Income from All Sources, 1987. Parent's Questionnaire, byp81. (1000s)

Foreign Language. Parent's Questionnaire, byp22c13. Coded 1 if a language other than english was spoken in the respondent's home.

 # Siblings. Parent's Questionnaire, byp3a. Total number of child's siblings including step-siblings and adoptive siblings.

First Born. Parent's Questionnaire, byp4. Coded 1 if none of the child's siblings counted in byp3a are older than the sample 8th grader.

Mother and Father's Education. Parent's Questionnaire, byp30, byp31. These variables record the highest level of education completed by the respondent, and the respondent's spouse. Note that the respondent may be the mother, the father, or some other relative. If some other relative answered this question, then the parent's education is missing. For both the mother and the father, we define separate dummy variables for the following educational categories: GED, High school, Vocational Certificate, 2 years of college, 3 years of college, 4 years of college, graduate work, and missing education.

Single Parent Household. Parent's Questionnaire, byp1a2. Question asks about partner's relationship to 8th grader. If the respondent replied that there was no partner than the dummy variable for a single parent household was coded as a one.